Journal of Sports Sciences
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjsp20

Childhood aerobic fitness predicts cognitive performance one year later

Laura Chaddock [a b], Charles H. Hillman [b c], Matthew B. Pontifex [c], Christopher R. Johnson [c], Lauren B. Raine [c] & Arthur F. Kramer [a b]

[a] Department of Psychology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[b] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[c] Department of Kinesiology and Community Health, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Available online: 19 Jan 2012

To cite this article: Laura Chaddock, Charles H. Hillman, Matthew B. Pontifex, Christopher R. Johnson, Lauren B. Raine & Arthur F. Kramer (2012): Childhood aerobic fitness predicts cognitive performance one year later, Journal of Sports Sciences, 30:5, 421-430

To link to this article: http://dx.doi.org/10.1080/02640414.2011.647706

Journal of Sports Sciences, March 2012; 30(5): 421–430

Childhood aerobic fitness predicts cognitive performance one year later

LAURA CHADDOCK[1,2], CHARLES H. HILLMAN[2,3], MATTHEW B. PONTIFEX[3], CHRISTOPHER R. JOHNSON[3], LAUREN B. RAINE[3], & ARTHUR F. KRAMER[1,2]

[1]*Department of Psychology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA,* [2]*Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA and* [3]*Department of Kinesiology and Community Health, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA*

(Accepted 5 December 2011)

Abstract

Aerobically fit children outperform less fit peers on cognitive control challenges that involve inhibition, cognitive flexibility, and working memory. The aim of this study was to determine whether, compared with less fit children, more fit 9- and 10-year-old pre-adolescents exhibit superior performance on a modified compatible and incompatible flanker task of cognitive control at the initial time of fitness testing and approximately one year later. We found that more fit children demonstrated increased flanker accuracy at both test sessions, coupled with a superior ability to flexibly allocate strategies during task conditions that required different amounts of cognitive control, relative to less fit children. More fit children also gained a speed benefit at follow-up testing. Structural MRI data were also collected to investigate the relationship between basal ganglia volume and task performance. Bilateral putamen volumes of the dorsal striatum and globus pallidus volumes predicted flanker performance at initial and follow-up testing one year later. The present findings suggest that childhood aerobic fitness and basal ganglia volumes relate to cognitive control at the time of fitness testing and may play a role in cognitive performance in the future. We hope that this research will encourage public health and educational changes that will promote a physically active lifestyle in children.

Keywords: *Basal ganglia, brain, cognition, development, exercise*

Introduction

A physically active lifestyle during childhood is positively associated with brain and cognitive health. A growing database of research suggests that aerobically fit children exhibit higher academic achievement scores, superior cognitive performance, more efficient neuroelectric activation underlying attentional processes, and larger brain volumes in the hippocampus and basal ganglia than less fit children (Buck, Hillman, & Castelli, 2008; Castelli, Hillman, Buck, & Erwin, 2007; Chaddock et al., 2010a, 2010b; Chaddock, Hillman, Buck, & Cohen, 2011; Chomitz et al., 2009; Hillman, Buck, Themanson, Pontifex, & Castelli 2009; Hillman, Castelli, & Buck, 2005; Kamijo et al., 2011; Pontifex et al., 2011; Sibley & Etnier, 2003). These findings have significant public health and educational implications, especially as childhood obesity and illness rates increase, and opportunities for physical activity during the school day are reduced or eliminated (Anderson, Crespo, Barlett, Cheskin, & Pratt, 1998; Baker, Olsen, & Sorensen, 2007; Department of Health and Human Services and Department of Education, 2000; Freedman, Khan, Dietz, Srinivasan, & Berenson, 2001; Sisson et al., 2009).

Since most studies focusing on childhood fitness and neurocognition employ cross-sectional designs, conclusions cannot be made regarding the duration of the positive effects of fitness on children's cognition. Nevertheless, in a large study of 1,221,727 young adults, cardiovascular fitness at age 18 years predicted intelligence and educational achievement later in life (Aberg et al., 2009). Similarly, the physical activity level of older adults has been shown to predict cognitive function and dementia 5–6 years later (Barnes, Yaffe, Satariano, & Tager, 2003; Laurin, Verreault, Lindsay, MacPherson, & Rockwood, 2001; Middleton et al., 2011). Accordingly, we applied this prospective study design to a youth population by examining whether

Correspondence: L. Chaddock, Department of Psychology, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, IL 61801, USA. E-mail: lchaddo2@illinois.edu

ISSN 0264-0414 print/ISSN 1466-447X online © 2012 Taylor & Francis
http://dx.doi.org/10.1080/02640414.2011.647706

childhood aerobic fitness levels can predict cognitive performance approximately one year later.

Researchers have demonstrated that more fit pre-adolescent children outperform their less fit peers on tests of cognitive control that involve inhibition, cognitive flexibility, and working memory (Chaddock et al., 2010b; Diamond, 2006; Hillman, Erickson, & Kramer, 2008; Pontifex et al., 2011). These cognitive operations are important for academic achievement, learning, and activities of daily living. Specifically, inhibition includes the ability to filter irrelevant environmental information, override a prepotent response, and stop an ongoing response, abilities that are central to attention and action control (Barkley, 1997; Davidson, Amso, Anderson, & Diamond, 2006). Cognitive flexibility refers to the ability to restructure knowledge and information based on changing situational demands (Diamond, 2006). Working memory is often defined as an ability to temporarily store and manage information while learning and responding to cognitive challenges (Baddeley, 1992).

We extended previous research involving cross-sectional designs by examining whether 9- and 10-year-old pre-adolescent children classified as more fit exhibited superior performance compared with less fit peers on a test of cognitive control at both the initial time of fitness testing as well as approximately one year later. At both visits, we administered the same modified flanker paradigm, which has been shown to yield performance differences as a function of childhood fitness (Pontifex et al., 2011). The modified flanker paradigm consisted of both compatible (congruent and incongruent) and incompatible (congruent and incongruent) trials. Compatibility refers to the relationship between target direction and response direction, while congruency refers to the relationship between target and distractors. Each compatibility condition included both congruent and incongruent trials.

During the compatible (congruent and incongruent) task condition, which is the condition used in most flanker research (e.g. Chaddock et al., 2010b), participants were required to attend and respond to the direction of a centrally presented arrow amid laterally presented flanking arrows (Eriksen & Eriksen, 1974; Kramer, Humphrey, Larish, Logan, & Strayer, 1994). The compatible incongruent task condition (i.e. < < > < <) required greater attentional and interference control to filter potentially misleading flankers that were mapped to incorrect behavioural responses, relative to the compatible congruent task condition (i.e. < < < < <). We extend the compatible flanker task (Chaddock et al., 2010b) to include a second condition that consisted of an incompatible (congruent and incongruent) stimulus–response condition at initial test (Pontifex et al., 2011) and follow-up test. During incompatible trials, cognitive control demands increased as participants were instructed to respond to the direction *opposite* to that of the centrally presented target arrow (Friedman, Nessler, Cycowicz, & Horton, 2009; Pontifex et al., 2011). While the task conditions varied in cognitive control demands, the paradigm generally required all elements of cognitive control: (1) inhibition, via the filtering of interfering flanking arrows; (2) cognitive flexibility, via the need to switch between instructional demands regarding the correct response to the direction of the central arrow; and (3) working memory, important for managing and remembering task goals.

Structural magnetic resonance imaging (MRI) data were also collected at initial testing. We recently reported that, compared with their less fit peers, more fit children have larger volumes in the dorsal striatum of the basal ganglia, which is associated with a greater ability to inhibit distraction during a compatible congruent and incongruent flanker task (Chaddock et al., 2010b). These results support the claim that the dorsal striatum is involved in cognitive control, motor integration, response resolution, and the execution of learned behaviours (Aron, Poldrack, & Wise, 2009), processes that are amenable to fitness across the lifespan (Hillman et al., 2008; Kramer et al., 1999). The findings are also consistent with rodent research indicating that wheel running positively influences the striatum (Aguiar, Speck, Prediger, Kapczinski, & Pinho, 2008; Li et al., 2005; Marais, Stein, & Daniels, 2009; Marques et al., 2008; Shi, Luo, Woodward, & Chang, 2004). In addition, Chaddock et al. (2010b) demonstrated an association between aerobic fitness, inhibition, and volume of the globus pallidus, an output region of the basal ganglia (Aron et al., 2009; Di Martino et al., 2008; Draganski et al., 2008). No relationship was found between fitness, task performance, and volume of the ventral striatum, a region involved in limbic and reward processes, functions that are less involved in selective attention and inhibition (Aron et al., 2009; Casey, Getz, & Galvan, 2008; Graybiel, 2005, 2008).

Given the reported association between childhood fitness, specific basal ganglia volumes, and cognitive control (Chaddock et al., 2010b), in the present study we wished to determine whether this relationship existed approximately one year later. It was hypothesized that children classified as more fit at initial testing would show superior flanker task performance at time of fitness testing and approximately one year later. Furthermore, we predicted that the volume of the dorsal striatum and globus pallidus would relate to aerobic fitness and modified flanker task performance at the initial visit to our laboratory (Chaddock et al., 2010b), as well as task performance at the follow-up visit.

Methods

Participants

Pre-adolescent 9- and 10-year-old children were recruited from East-Central Illinois. At the first visit, children were screened for several factors that influence physical activity participation or cognitive function. The Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990) was administered to each child to obtain a composite intelligence quotient (IQ) score including both crystallized and fluid intelligence measures. Participants were excluded if their scores were more than one standard deviation below the mean (85%). Next, a guardian of the child completed the Attention-Deficit Hyperactivity Disorder (ADHD) Rating Scale IV (DuPaul, Power, Anastopoulos, & Reid, 1998) to screen for the presence of attentional disorders. Participants were excluded if they scored above the 85th percentile. Pubertal timing was also assessed using a modified Tanner Staging System (Tanner, 1962; Taylor et al., 2001) with all included pre-pubescent participants at or below a score of 2 on a 5-point scale of developmental stages. In addition, socio-economic status was determined by creating a trichotomous index based on three variables: participation in a free or reduced-price meal programme at school; the highest level of education obtained by the child's mother and father; and the number of parents who worked full-time (Birnbaum et al., 2002).

Furthermore, eligible participants were required to: (1) qualify as higher-fit or lower-fit (see next subsection on "Aerobic fitness assessment"); (2) demonstrate right handedness (as measured by the Edinburgh Handedness Questionnaire; Oldfield, 1971); (3) report no adverse health conditions, physical incapacities, or neurological disorders; (4) report no use of medications that influence central nervous system function; (5) successfully complete a mock MRI session to screen for claustrophobia in an MRI machine; and (6) sign an informed assent approved by the University of Illinois at Urbana-Champaign. A legal guardian provided written informed consent in accordance with the Institutional Review Board of the University of Illinois at Urbana-Champaign. Participants were compensated for their participation.

Aerobic fitness assessment

The aerobic fitness of each child was determined at the first visit by measuring maximal oxygen uptake ($\dot{V}O_{2max}$) using a computerized indirect calorimetry system (ParvoMedics True Max 2400) during a modified Balke protocol (American College of Sports Medicine, 2006). Specifically, participants ran on a motor-driven treadmill at a constant speed with increases in grade of 2.5% every 2 min until volitional exhaustion. Mean values for oxygen uptake ($\dot{V}O_2$) and respiratory exchange ratio (RER) (the ratio between carbon dioxide and oxygen) were assessed every 30 s. In addition, heart rate was measured throughout the fitness test using a heart rate monitor (Polar WearLink® + 31, Polar Electro, Finland), and ratings of perceived exertion were assessed every 2 min using the children's OMNI scale (Utter, Roberson, Nieman, & Kang, 2002).

Maximal oxygen uptake was considered to have been reached when oxygen consumption remained at a steady state despite an increase in workload. Relative peak $\dot{V}O_2$ was based upon maximal effort as evidenced by: (1) a peak heart rate greater than 185 beats · min^{-1} (American College of Sports Medicine, 2006) accompanied by a heart rate plateau, i.e. an increase in work rate without a concomitant increase in heart rate (Freedson & Goodman, 1993); (2) RER greater than 1.0 (Bar-Or, 1983); and/or (3) ratings on the children's OMNI scale of perceived exertion greater than 8 (Utter et al., 2002). Relative peak $\dot{V}O_2$ was expressed in mL · kg^{-1} · min^{-1}.

Fitness group assignments (i.e. higher-fit and lower-fit) were based on whether a child's $\dot{V}O_{2max}$ fell above the 70th percentile or below the 30th percentile according to normative data provided by Shvartz and Reibold (1990). Children who did not qualify as higher-fit or lower-fit were excluded from participation.

Modified flanker task

At initial and follow-up testing, participants completed a modified version of the Eriksen flanker task, a paradigm often employed to measure cognitive control (Eriksen & Eriksen, 1974; Kramer et al., 1994). The modified version of the task included both a compatible condition with the original congruent and incongruent flanker task arrays (Chaddock et al., 2010b; Hillman et al., 2009) and an incompatible condition that added additional conflict and cognitive control requirements (Friedman et al., 2009; Pontifex et al., 2011) (Figure 1).

First, participants completed the compatible stimulus–response task condition, in which they were instructed to respond both quickly and accurately to the direction of a centrally presented target arrow amid either congruent (e.g. <<<<< or >>>>>) or incongruent (e.g. <<><< or >><>>) flanking arrows. The incongruent

Compatible	**Incompatible**
Congruent	Congruent
> > > > >	> > > > >
Response Right	Response Left
Compatible	**Incompatible**
Incongruent	Incongruent
> > < > >	> > < > >
Response Left	Response Right

Figure 1. The modified flanker task. Compatible congruent: right hand response (arrow points right, respond to direction of central arrow). Compatible incongruent: left hand response (arrow points left, respond to direction of central arrow, amid distractors). Incompatible congruent: left hand response (arrow points right, respond to direction opposite of central arrow). Incompatible incongruent: right hand response (arrow points left, respond to direction opposite of central arrow, amid distractors).

condition required the processing of both the correct response (elicited by the target arrow) and the incorrect response (elicited by the flanking stimuli) before stimulus evaluation was complete (Spencer & Coles, 1999). Thus, greater amounts of cognitive control are required to inhibit the flanking stimuli and to execute a correct response in the incongruent condition relative to the congruent condition.

Next, participants completed the incompatible stimulus–response condition, wherein they were instructed to respond both quickly and accurately in the direction *opposite* to that of the centrally presented target arrow (Friedman et al., 2009, Pontifex et al., 2011). The incompatible condition manipulated task difficulty through perceptual and response conflict such that the incompatible, incongruent condition required the greatest amount of inhibitory control.

For each compatibility condition, two blocks of 100 trials were presented with equally probable congruency and directionality. The order of the compatibility blocks was fixed (compatible, compatible, incompatible, incompatible) to ensure response strategy conflict for incompatible blocks. Across both conditions, 3 cm tall white arrows were presented focally on a black background (16.5 cm wide). A 200 ms stimulus duration, 1600 ms response window, and 1700 ms inter-stimulus interval (i.e. 1500 ms of blank screen between each array of arrows) were programmed.

Magnetic resonance imaging protocol and image processing

High-resolution (1.3 mm × 1.3 mm × 1.3 mm) T1-weighted structural brain images were acquired for all participants at initial testing using a 3D MPRAGE (Magnetization Prepared Rapid Gradient Echo Imaging) protocol with 144 contiguous axial slices, collected in ascending fashion parallel to the anterior and posterior commissures [echo time (TE) = 3.87 ms, repetition time (TR) = 1800 ms, field of view (FOV) = 256 mm, acquisition matrix = 192 mm × 192 mm, slice thickness = 1.3 mm, and flip angle = 8°). All images were collected on a 3T head-only Siemens Allegra MRI scanner.

Segmentation and volumetric analysis of the left and right dorsal striatum (i.e. caudate nucleus and putamen), ventral striatum (i.e. nucleus accumbens), and globus pallidus were performed using a semi-automated, model-based subcortical tool (FMRIB's Integrated Registration and Segmentation Tool; FIRST) in FMRIB's Software Library (FSL) version 4.1.4 (Patenaude 2007; Patenaude, Smith, Kennedy, & Jenkinson, 2007a, 2007b). A two-stage affine registration to a standard space template (MNI space) with 1 mm resolution using 12 degrees of freedom and a subcortical mask to exclude voxels outside the subcortical regions was performed on each participant's MPRAGE. Next, the caudate nucleus, putamen, nucleus accumbens, and globus pallidus were segmented with 30, 40, 50, and 40 modes of variation for each structure, respectively. For a full description of the FIRST methodology, refer to Chaddock et al. (2010a) and Patenaude et al. (2007a, 2007b). Previously, researchers have reported high test–retest reliability of this segmentation algorithm (Erickson et al., 2010).

Segmentations were visually checked for errors. Finally, boundary correction was run, a process that classifies boundary voxels as belonging to the structure (or not) based on a statistical probability (z-score > 3.00; $P < 0.001$). The volume of each participant's caudate nucleus, putamen, nucleus accumbens, and globus pallidus was measured in cubic millimeters (mm^3), and bilateral brain volume values were used in subsequent analyses.

Procedure

Participants visited the laboratory on four separate testing days, at two different time points (i.e. initial testing and approximately one year later). Thirty-two participants completed both test sessions. The sample included 14 higher-fit children (7 boys, 7 girls) and 18 lower-fit children (8 boys, 10 girls). On average, the children returned to the laboratory for follow-up cognitive testing approximately one year after initial testing (mean = 1.28 years, $s = 0.46$).

Initial visit. Informed assent and consent were obtained together with information about IQ, ADHD, pubertal timing, socioeconomic status, handedness, and health. The $\dot{V}O_{2max}$ test of aerobic fitness and the mock MRI session were also administered. Participants returned on a separate day to complete the modified flanker task. During another visit, participants completed the MRI session to acquire a structural brain image.

Follow-up visit. Participants returned to the laboratory approximately one year following initial testing. Participants and their guardians again completed informed assent and consent forms, and an identical modified flanker task was administered. No demographic, $\dot{V}O_{2max}$ or MRI data were collected.

Statistical analysis

Task performance (response accuracy, reaction time (RT)) was assessed using a 2 (aerobic fitness group: higher-fit, lower-fit) × 2 (compatibility: compatible, incompatible) × 2 (congruency: congruent, incongruent) × 2 (test session: initial test, follow-up test) repeated-measures analysis of variance (ANOVA). If the omnibus ANOVA reached significance, *post-hoc* comparisons were performed (with Bonferroni-corrected *t*-tests). The family-wise alpha level was set at $P < 0.05$.

Analyses of covariance (ANCOVAs) were conducted to compare bilateral basal ganglia volumes as a function of aerobic fitness group, with total intracranial volume (mm^3) as a covariate to control for variation in head size. Spearman correlations were performed to examine the hypothesized relationship between specific basal ganglia volumes and flanker task performance at initial testing and follow-up testing.

Results

Participant demographics

Participant demographic and fitness data from the initial visit are provided in Table I. Demographic variables (i.e. age, IQ, ADHD, Tanner, socioeconomic status) did not differ between fitness groups at the initial visit and were not correlated with follow-up flanker test performance (all $r < 0.2$, all $P > 0.2$). Higher-fit participants had higher $\dot{V}O_{2max}$ scores than lower-fit children ($t_{30} = 10.7$, $P < 0.001$, effect size = 4.0), confirming the aerobic fitness groupings. Higher-fit (mean = 1.3 years, $s = 0.5$ years) and lower-fit (mean = 1.3 years, $s = 0.4$ years) participants did not differ in the length of time between initial and follow-up testing ($t_{30} = 0.5$, $P > 0.6$).

Aerobic fitness and flanker task performance

Table II provides mean values (*s*) for response accuracy and reaction time for compatible and incompatible flanker task conditions (which consist of both congruent and incongruent trials) for higher-fit and lower-fit groups at initial testing and follow-up testing.

Table I. Participant mean (*s*) demographic and fitness data by aerobic fitness group at initial testing.

Variable	Lower-fit	Higher-fit
Gender	10 F, 8 M	7 F, 7 M
Age (years)	10.1 (0.6)	10.0 (0.6)
$\dot{V}O_{2max}$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	35.8 (3.8)*	52.1 (4.4)*
$\dot{V}O_{2max}$ percentile (%)	12.7 (18.0)*	83.1 (4.6)*
K-BIT composite score (IQ)	118.7 (14.2)	114.6 (7.6)
K-BIT crystallized score (Vocabulary)	112.2 (12.1)	110.7 (7.0)
K-BIT fluid score (Matrices)	121.4 (15.9)	115.7 (8.9)
ADHD[a]	6.6 (3.3)	6.3 (3.3)
Tanner[b]	1.6 (0.5)	1.6 (0.5)
Socioeconomic status[c] (median)	2.8 (0.6)	2.5 (0.8)

Note: K-BIT = Kaufman Brief Intelligence Test (Kaufman & Kaufman, 1990).
[a]Scores on the ADHD Rating Scale V (DuPaul et al, 1998).
[b]Pubertal timing assessed using a modified Tanner staging system (Tanner, 1962; Taylor et al, 2001).
[c]Socioeconomic status was determined by the creation of a trichotomous index based on three variables: child participation in a free or reduced-price lunch programme at school, the highest level of education obtained by the child's mother and father, and the number of parents who worked full-time (Birnbaum et al, 2002).
*Significantly different at $P < 0.001$.

Response accuracy. A main effect of compatibility indicated that participants had lower accuracy in the incompatible (mean = 79.1%, $s_x = 2.0\%$) than in the compatible condition (mean = 82.8%, $s_x = 1.6\%$) ($F_{1,30} = 19.2$, $P < 0.0001$). A main effect of congruency revealed lower accuracy on incongruent trials (mean = 77.4%, $s_x = 2.1\%$) than on congruent trials (mean = 84.5%, $s_x = 1.6\%$) ($F_{1,30} = 46.4$, $P < 0.0001$). Decomposition of the compatibility × -congruency interaction ($F_{1,30} = 28.5$, $P < 0.0001$) indicated lower accuracy for incongruent, compatible trials (mean = 77.6%, $s = 12.0\%$) compared with congruent, compatible trials (mean = 88.0%, $s = 8.4\%$) ($t_{32} = 7.4$, $P < 0.0001$), and for incongruent, incompatible trials (mean = 77.2%, $s = 13.9\%$) compared with congruent, incompatible trials (mean = 80.7%, $s = 11.3\%$) ($t_{32} = 3.6$, $P = 0.001$). Participants had lower accuracy for congruent, incompatible trials than for congruent, compatible trials ($t_{32} = 6.3$, $P < 0.0001$), while no differences in accuracy were observed between incongruent trials for compatible and incompatible conditions ($t_{32} = 0.42$, $P = 0.68$).

A main effect of fitness ($F_{1,30} = 5.1$, $P = 0.03$, effect size = 0.6) revealed that higher-fit children had higher accuracy (mean = 85.0%, $s_x = 2.7\%$) than

Table II. Mean (s) task performance (i.e. accuracy and reaction time) by aerobic fitness group for compatible and incompatible conditions of the modified flanker task.

Test	Initial test		Follow-up test	
	Lower-fit	Higher-fit	Lower-fit	Higher-fit
Compatible congruent accuracy (%)	82.5 (15.9)	88.6 (8.6)	88.6 (7.3)	92.9 (6.3)
Compatible incongruent accuracy (%)	69.6 (20.5)	77.9 (7.4)	78.6 (10.7)	83.7 (10.9)
Incompatible congruent accuracy (%)	72.7 (16.9)*	85.6 (10.7)*	79.9 (12.1)	85.3 (11.2)
Incompatible incongruent accuracy (%)	67.6 (22.1)*	80.4 (11.9)*	75.9 (13.2)*	85.6 (8.3)*
Compatible congruent RT (ms)	523.3 (128.0)	486.4 (73.4)	566.9 (98.3)*	461.2 (94.5)*
Compatible incongruent RT (ms)	587.3 (143.6)	567.4 (113.2)	621.9 (110.8)*	508.6 (94.4)*
Incompatible congruent RT (ms)	540.2 (113.8)	514.0 (130.0)	614.3 (136.4)*	485.3 (98.9)*
Incompatible incongruent RT (ms)	555.5 (122.5)	547.4 (162.5)	664.7 (150.9)*	522.2 (112.9)*

Note: RT = reaction time.
*Significantly different at $P < 0.05$.

lower-fit children (mean = 76.9%, $s_x = 2.4\%$), across compatibility task conditions and test sessions. A fitness × compatibility interaction ($F_{1,30} = 6.4$, $P = 0.017$) demonstrated that higher-fit children maintained accuracy across compatible (mean = 85.8%, $s_x = 2.4\%$) and incompatible (mean = 84.2%, $s_x = 3.0\%$) task conditions ($t_{13} = 1.1$, $P = 0.30$), whereas lower-fit children showed lower accuracy on the incompatible condition (mean = 74.0%, $s_x = 2.7\%$) relative to the compatible condition (mean = 79.8%, $s_x = 2.1\%$) ($t_{13} = 2.3$, $P = 0.032$).

A main effect of time ($F_{1,30} = 5.6$, $P = 0.024$) indicated higher accuracy at follow-up testing (mean = 83.8%, $s_x = 1.6\%$) compared with initial testing (mean = 78.1%, $s_x = 2.6\%$). A congruency × time interaction ($F_{1,30} = 8.3$, $P = 0.007$) indicated significant differences between incongruent accuracy at initial testing (mean = 73.9%, $s_x = 2.9\%$) and follow-up testing (mean = 81.0%, $s_x = 1.9\%$) ($t_{32} = 2.9$, $P = 0.008$) but no differences between congruent trials at initial testing (mean = 82.4%, $s_x = 2.3\%$) and follow-up testing (mean = 86.7%, $s_x = 1.5\%$) ($t_{32} = 2.0$, $P = 0.053$).

Reaction time. A main effect of congruency indicated longer reaction times for incongruent trials (mean - 571.9 ms, $s_x = 16.2$ ms) than for congruent trials (mean = 524.0 ms, $s_x = 14.4$ ms) ($F_{1,30} = 74.0$, $P < 0.0001$). This main effect was superseded by a compatibility × congruency interaction ($F_{1,30} = 13.5$, $P < 0.0001$), which further demonstrated longer reaction times for incongruent, compatible trials (mean = 571.3 ms, $s_x = 16.4$ ms) compared with congruent, compatible trials (mean = 509.4 ms, $s_x = 14.1$ ms) ($t_{32} = 9.7$, $P < 0.0001$) as well as for incongruent, incompatible trials (mean = 572.5 ms, $s_x = 18.2$ ms) compared with congruent, incompatible trials (mean = 538.5 ms, $s_x = 16.9$ ms) ($t_{32} = 4.9$, $P < 0.0001$). Participants showed longer reaction times for congruent, incompatible trials compared with congruent, compatible trials ($t_{32} = 2.7$, $P = 0.01$), and no differences between incongruent reaction times for compatible and incompatible trials ($t_{32} = 0.18$, $P = 0.86$).

A main effect of fitness ($F_{1,30} = 5.8$, $P = 0.02$, effect size = 0.6) indicated that higher-fit children (mean = 511.6 ms, $s_x = 22.6$ ms) had shorter reaction times than lower-fit children (mean = 584.3 ms, $s_x = 19.9$ ms). A fitness × time interaction ($F_{1,30} = 4.2$, $P = 0.05$) superseded the fitness main effect and indicated that higher-fit children and lower-fit children only differed in reaction time at the follow-up test, with higher-fit children (mean - 494.3 ms, $s_x = 28.3$ ms) demonstrating shorter reaction times than lower-fit children (mean = 617.0 ms, $s_x = 24.9$ ms) ($t_{32} = 3.4$, $P = 0.002$, effect size = 0.8). Furthermore, lower-fit children became slower and higher-fit children faster over the course of the year between initial and follow-up testing on the flanker task.

Aerobic fitness, basal ganglia volume, and flanker task performance

Higher-fit children (mean = 11536.3 mm^3, $s_x = 411.1$ mm^3) had larger bilateral putamen volumes of the dorsal striatum than lower-fit children (mean = 9681.1 mm^3, $s_x = 362.0$ mm^3) ($F_{1,29} = 11.4$, $P = 0.002$, effect size = 0.8). Higher-fit children (mean = 4006.7 mm^3, $s_x = 119.9$ mm^3) also had larger bilateral globus pallidus volumes than lower-fit children (mean = 3571.9 mm^3, $s_x = 105.6$ mm^3) ($F_{1,29} = 7.3$, $P = 0.011$, effect size = 0.7). There were no fitness group differences in bilateral caudate volumes or bilateral nucleus accumbens volumes (both $F < 1.2$, both $P > 0.2$).

Table III presents all Spearman correlations between basal ganglia volumes and flanker task performance at initial and follow-up testing. In

Table III. Spearman correlations (*r*) between bilateral basal ganglia volumes and modified flanker task performance (i.e., accuracy and reaction time).

Test	Caudate	Putamen	Globus pallidus	Nucleus accumbens
Initial test				
Compatible congruent accuracy (%)	0.14	0.09	0.13	0.18
Compatible incongruent accuracy (%)	0.23	0.19	0.17	0.18
Incompatible congruent accuracy (%)	0.17	0.35*	0.31	0.23
Incompatible incongruent accuracy (%)	0.33	0.39*	0.38*	0.22
Compatible congruent RT (ms)	0.21	−0.04	0.02	0.07
Compatible incongruent RT (ms)	0.10	−0.13	−0.08	−0.04
Incompatible congruent RT (ms)	0.25	0.05	0.06	0.11
Incompatible incongruent RT (ms)	0.15	−0.02	−0.01	0.06
Follow-up test				
Compatible congruent accuracy (%)	−0.03	0.42*	0.25	0.38*
Compatible incongruent accuracy (%)	−0.10	0.28	0.09	0.23
Incompatible congruent accuracy (%)	−0.23	0.18	0.11	0.33
Incompatible incongruent accuracy (%)	−0.08	0.41*	0.18	0.23
Compatible congruent RT (ms)	−0.09	−0.43*	−0.49*	−0.24
Compatible incongruent RT (ms)	0.01	−0.38*	−0.43*	−0.24
Incompatible congruent RT (ms)	−0.10	−0.37*	−0.37*	−0.14
Incompatible incongruent RT (ms)	−0.13	−0.41*	−0.39*	−0.20

Note: RT = reaction time.
*Significantly different at $P < 0.05$.

particular, at initial test, putamen and globus pallidus volumes were positively correlated with incompatible accuracy. At follow-up test, putamen volumes were related to compatible and incompatible accuracy and reaction time, and globus pallidus volumes were associated with compatible and incompatible reaction times.

Discussion

We extended previous research by demonstrating that children classified as aerobically fit still outperform their lower-fit peers on the cognitive challenge approximately one year later. These results parallel similar observations in young adults (Aberg et al., 2009) and older adults (Barnes et al., 2003; Laurin et al., 2001; Middleton et al., 2011) and suggest that aerobic fitness in a pre-adolescent population can predict later cognitive abilities.

In the present study, higher-fit children exhibited increased accuracy across compatibility conditions during both test sessions, coupled with shorter compatible and incompatible response times at follow-up testing, compared with lower-fit children. Furthermore, across test sessions, higher-fit participants maintained accuracy during compatible and incompatible trials while lower-fit participants showed performance decrements during the task conditions requiring greater amounts of inhibition. The results suggest that higher-fit participants have a superior capability to flexibly allocate cognitive control processes and alter their strategies to effectively meet task demands both at the time of fitness classification as well as approximately one year later. Moreover, while all participants showed increased accuracy across test sessions, possibly due to practice, higher-fit children maintained greater accuracy and gained a speed benefit at follow-up testing, relative to their lower-fit peers. It is possible that higher-fit children gained additional task performance benefits (i.e. in terms of accuracy and reaction time) with learning and/or maturation. In support of this idea, the aerobic fitness ($\dot{V}O_{2max}$) of each child at initial test predicted his or her follow-up reaction time.

Our results suggest that aerobic fitness predicts future cognitive performance on a test of cognitive control and that volumetric measurements of the basal ganglia are associated with flanker task performance at both test sessions. In particular, bilateral putamen volume was associated with incompatible task performance at initial testing and compatible and incompatible performance at follow-up testing. Bilateral globus pallidus volume predicted incompatible, incongruent accuracy at initial testing as well as compatible and incompatible response times at follow-up testing. The associations were coupled with the finding that higher-fit children had larger volumes of the bilateral putamen and globus pallidus compared with lower-fit children at initial testing. No relationship was found between aerobic fitness and ventral striatum volume. The MRI data support Chaddock and colleagues' (2010b) observation of a relationship between fitness, specific basal ganglia volumes, and inhibitory control and extend these results to suggest that basal ganglia volume can predict children's cognitive control abilities approximately one year later. We extend Chaddock and colleagues' (2010b) compatible-only flanker task findings by showing an association between basal ganglia volumes and incompatible flanker task trials that require increased cognitive control. It is possible that larger basal ganglia volumes in higher-fit children are a factor in the ability to flexibly employ cognitive control strategies approximately one year later. Moreover, the brain–performance correlations raise the possibility that basal ganglia volume plays a role in the reaction time benefit specifically for higher-fit children at follow-up test.

Our study was designed to test whether baseline aerobic fitness and brain volume measurements in children can predict later cognitive performance. However, this prospective framework raises the possibility that changes in other variables (e.g. development, fitness, puberty, dexterity, nutrition) may play a role as well. Although our demographic variables were not correlated with follow-up testing performance, future investigations should measure additional developmental, brain, and environmental factors to clarify the mechanisms underlying the fitness–brain–cognition relationships. It would be interesting to classify the children as higher-fit and lower-fit at follow-up testing to see how the maintenance and/or change in fitness relate to performance.

In the future, researchers should continue to explore the level of response conflict in which lower-fit children are unable to maintain performance levels and allocate attentional resources effectively in response to increased task demands. Our results support studies using a compatible-only flanker task that showed an association between aerobic fitness, task performance (Hillman et al., 2009), and brain volume (Chaddock et al., 2010b). In addition, our study further illuminates fitness-related differences in cognitive control via the incompatible stimulus–response task manipulation (Friedman et al., 2009; Pontifex et al., 2011). Nevertheless, while participants showed significantly lower accuracy rates and longer response times for incompatible trials compared with compatible trials at both test sessions (Friedman et al., 2009; Pontifex et al., 2011), future investigations are needed to decompose the relationship between compatibility and congruency as a function of cognitive control requirements and childhood fitness.

In conclusion, we have provided additional support for the need to modify public health and educational policy to encourage a physically active lifestyle in children. Increased aerobic fitness and physical activity have been shown to relate to increased academic performance in mathematics and reading (Castelli et al., 2007; Chomitz et al., 2009; Coe, Pivarnik, Womack, Reeves, & Malina, 2006; Trudeau &, Shephard, 2008). It is possible that higher cognitive control, as exhibited by higher-fit children in the present study, helps to optimize important skills needed inside and outside the classroom. That is, children with higher cognitive control may show superior selective attention, inhibition of inappropriate or interfering responses, flexible thinking, and maintenance of information in working memory. Furthermore, our results suggest that aerobic fitness and specific brain volumes are associated with cognitive health at the initial time of fitness testing as well as play a role in future cognitive performance.

References

Aberg, M. A., Pedersen, N. L., Toren, K., Svartengren, M., Backstrand, B., Johnsson, T. et al. (2009). Cardiovascular fitness is associated with cognition in young adulthood. *Proceedings of the National Academy of Sciences USA*, *106*, 20906–20911.

Aguiar, A. S., Speck, A. E., Prediger, R. D., Kapczinski, F., & Pinho, R. A. (2008). Downhill training upregulates mice hippocampal and striatal brain-derived neurotrophic factor levels. *Journal of Neural Transmission*, *115*, 1251–1255.

American College of Sports Medicine (2006). *ACSM's guidelines for exercise testing and prescription* (7th edn.). New York: Lippincott Williams & Wilkins.

Anderson, R. E., Crespo, C. J., Barlett, S. J., Cheskin, L. J., & Pratt, M. (1998). Relationship of physical activity and television watching with body weight and level of fatness among children. *Journal of the American Medical Association*, *279*, 938–942.

Aron, A. R., Poldrack, R. A., & Wise, S. P. (2009). Cognition: Basal ganglia role. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (Vol. 2, pp. 1069–1077). Oxford: Academic Press.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559.

Baker, J. L., Olsen, L. W., & Sorensen, T. I. A. (2007). Childhood body-mass index and risk of coronary heart disease in adulthood. *New England Journal of Medicine*, *357*, 2329–2337.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions. *Psychological Bulletin*, *121*, 65–94.

Barnes, D. E., Yaffe, K., Satariano, W. A., & Tager, I. B. (2003). A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. *Journal of the American Geriatrics Society*, *51*, 459–465.

Bar-Or, O. (1983). *Pediatric sports medicine for the practitioner: From physiologic principles to clinical applications.* New York: Springer-Verlag.

Birnbaum, A. S., Lytle, L. A., Murray, D. M., Story, M., Perry, C. L., & Boutelle, K. N. (2002). Survey development for assessing correlates of young adolescents' eating. *American Journal of Health Behavior*, *26*, 284–295.

Buck, S. M., Hillman, C. H., & Castelli, D. M. (2008). The relation of aerobic fitness to Stroop task performance in preadolescent children. *Medicine and Science in Sports and Exercise*, *40*, 166–172.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review*, *28*, 62–77.

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 and Exercise Psychology*, *29*, 239–252.

Chaddock, L., Erickson K. I., Prakash, R. S., Kim, J. S., Voss, M. W., VanPatter, M. et al. (2010a). A neuroimaging investigation of the association between aerobic fitness, hippocampal volume and memory performance in preadolescent children. *Brain Research*, *1358*, 172–183.

Chaddock, L., Erickson, K. I., Prakash, R. S., VanPatter, M., Voss, M. V., Pontifex, M. B. et al. (2010b). Basal ganglia volume is associated with aerobic fitness in preadolescent children. *Developmental Neuroscience*, *32*, 249–256.

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*, 344–349.

Chomitz, V. R., Slining, M. M., McGowan, R. J., Mitchell, S. E., Dawson, G. F., & Hacker, K. A. (2009). Is there a relationship between physical fitness and academic achievement? Positive results from public school children in the northeastern United States. *Journal of School Health*, *79*, 30–37.

Coe, D. P., Pivarnik, J. M., Womack, C. J., Reeves, M. J., & Malina, R. M. (2006). Effect of physical education and activity levels on academic achievement in children. *Medicine and Science in Sports and Exercise*, *38*, 1515–1519.

Davidson, M. C., Amso, D., Anderson, L. C., & 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*, 2037–2078.

Department of Health and Human Services and Department of Education (2000). *Promoting better health for young people through physical activity and sports.* A Report to the President from the Secretary of Health and Human Services and the Secretary of Education. Silver Spring, MD: Centers for Disease Control.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). New York: Oxford University Press.

Di Martino, A., Scheres, A., Marguiles, D. S., Kelly, A. M., Uddin, L. Q., Shehzad, Z. et al. (2008). Functional connectivity of human striatum: A resting state fMRI study. *Cerebral Cortex*, *18*, 2735–2747.

Draganski, B., Kherif, F., Kloppel, S., Cook, P. A., Alexander, D. C., Parker, G. J. et al. (2008). Evidence for segregated and integrative connectivity patterns in the human basal ganglia. *Journal of Neuroscience*, *28*, 7143–7152.

DuPaul, G. J., Power, T. J., Anastopoulos, A., & Reid, R. (1998). *ADHD rating scale IV: Checklists, norms, and clinical interpretation.* New York: Guilford Press.

Erickson, K. I., Boot, W. R., Basak, C., Neider, M. R., Prakash, R. S., Voss, M. W. et al. (2010). Striatal volume predicts level of video game skill acquisition. *Cerebral Cortex*, *20*, 2522–2530.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters on the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*, 143–149.

Freedman, D. S., Khan, L. K., Dietz, W. H., Srinivasan, S. R., & Berenson, G. S. (2001). Relationship of childhood obesity to coronary heart disease risk factors in adulthood: The Bogalusa heart study. *Pediatrics*, *108*, 712–718.

Freedson, P. S., & Goodman, T. L. (1993). Measurement of oxygen consumption. In T. W. Rowland (Ed.), Pediatric laboratory exercise testing: Clinical guidelines (pp. 91–113). Champaign, IL: Human Kinetics.

Friedman, D., Nessler, D., Cycowicz, Y. M., & Horton, C. (2009). Development of and change in cognitive control: A comparison of children, young adults, and older adults. *Cognitive, Affective and Behavioral Neuroscience*, *9*, 91–102.

Graybiel, A. M. (2005). The basal ganglia: Learning new tricks and loving it. *Current Opinion in Neurobiology*, *15*, 638–644.

Graybiel, A. M. (2008). Habits, rituals and the evaluative brain. *Annual Review of Neuroscience*, *31*, 359–387.

Hillman, C. H., Buck, S. M., Themanson, J. R., Pontifex, M. B., & Castelli, D. M. (2009). Aerobic fitness and cognitive development: Event-related brain potential and task performance of executive control in preadolescent children. *Developmental Psychology*, *45*, 114–129.

Hillman, C. H., Castelli, D. M., & Buck, S. M. (2005). Aerobic fitness and neurocognitive function in healthy preadolescent children. *Medicine and Science in Sports and Exercise*, *37*, 1967–1974.

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews Neuroscience*, *9*, 58–65.

Kamijo, K., Pontifex, M. B., O'Leary, K. C., Scudder, M. R., Wu, C., Castelli, D. M. et al. (2011). The effects of an afterschool physical activity program on working memory in preadolescent children. *Developmental Science*, *14*, 1046–1058.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman brief intelligence test.* Circle Pines, MN: AGS.

Kramer, A. F., Hahn, S., Cohen, N., Banich, M., McAuley, E., Harrison, C. et al. (1999). Aging, fitness, and neurocognitive function. *Nature*, *400*, 418–419.

Kramer, A. F., Humphrey, D. G., Larish, J. F., Logan, G. B., & Strayer, D.L. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging*, *9*, 491–512.

Laurin, D., Verreault, R., Lindsay, J., MacPherson, K., & Rockwood, K. (2001). Physical activity and risk of cognitive impairment and dementia in elderly persons. *Archives of Neurology*, *58*, 498–504.

Li, J., Ding, Y. H., Rafols, J. A., Lai, Q., McAllister, J. P., & Ding, Y. (2005). Increased astrocyte proliferation in rats after running exercise. *Neuroscience Letters*, *386*, 160–164.

Marais, L., Stein, D. J., & Daniels, W. M. (2009). Exercise increases BDNF levels in the striatum and decreases depressive-like behavior in chronically stressed rats. *Metabolic Brain Disease*, *24*, 587–597.

Marques, E., Vasconcelos, F., Rolo, M. R., Pereira, F. C., Silva, A. P., Macedo, T. R. et al. (2008). Influence of chronic exercise on the amphetamine-induced dopamine release and neurodegeneration in the striatum of the rat. *Annals of the New York Academy of Sciences*, *1139*, 222–231.

Middleton, L. E., Manini, T. M., Simonsick, E. M., Harris, T. B., Barnes, D. E., Tylavsky, F. et al. (2011). Activity energy expenditure and incident cognitive impairment in older adults. *Archives of Internal Medicine*, *171*, 1251–1257.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Patenaude, B. (2007). *Bayesian statistical models of shape and appearance for subcortical brain segmentation.* DPhil thesis, University of Oxford.

Patenaude, B., Smith, S. M., Kennedy, D., & Jenkinson, M. (2007a). FIRST-FMRIB's integrated registration and segmentation tool. *Human Brain Mapping Conference*.

Patenaude, B., Smith, S. M., Kennedy, D., & Jenkinson, M. (2007b). *Bayesian shape and appearance models.* Technical report TR07BP1, FMRIB Center, University of Oxford.

Pontifex, M. B., Raine, L. B., Johnson, C. R., Chaddock, L., Voss, M. W., Cohen, N. J. et al. (2011). Cardiorespiratory fitness and the flexible modulation of cognitive control in preadolescent children. *Journal of Cognitive Neuroscience*, *23*, 1332–1345.

Shi, L. H., Luo, F., Woodward, D. J., & Chang, J. Y. (2004). Neural responses in multiple basal ganglia regions during spontaneous and treadmill locomotion tasks in rats. *Experimental Brain Research*, *157*, 303–314.

Shvartz, E., & Reibold, R. C. (1990). Aerobic fitness norms for males and females aged 6 to 75 years: A review. *Aviation, Space, and Environmental Medicine*, *61*, 3–11.

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

Sisson, S. B., Church, T. S., Martin, C. K., Tudor-Locke, C., Smith, S. R., Bouchard, C. et al. (2009). Profiles of sedentary behavior in children and adolescents: The US national health and nutrition examination survey, 2001–2006. *International Journal of Pediatric Obesity*, *4*, 353–359.

Spencer, K. M., & Coles, M. G. H. (1999). The lateralized readiness potential: Relationship between human data and response activation in a connectionist model. *Psychophysiology*, *36*, 364–370.

Tanner, J. M. (1962). *Growth at adolescence*. Oxford: Blackwell Scientific.

Taylor, S. J. C., Whincup, P. H., Hindmarsh, P. C., Lampe, F., Odoki, K., & Cook, D. G. (2001). Performance of a new pubertal self-assessment questionnaire: A preliminary study. *Paediatric and Perinatal Epidemiology*, *15*, 88–94.

Trudeau, F., & Shephard, R. J. (2008). Physical education, school physical activity, school sports and academic performance. *International Journal of Behavioral Nutrition and Physical Activity*, *5*, 5–10.

Utter, A. C., Roberson, R. J., Nieman, D. C., & Kang, J. (2002). Children's OMNI scale of perceived exertion: Walking/running evaluation. *Medicine and Science in Sports and Exercise*, *34*, 139–144.