ORIGINAL ARTICLE

WILEY | **PSYCHOPHYSIOLOGY** SPR

# Effects of the FITKids physical activity randomized controlled trial on conflict monitoring in youth

Eric S. Drollette[1] | Matthew B. Pontifex[2] | Lauren B. Raine[3] | Mark R. Scudder[4] | R. Davis Moore[5] | Shih-Chun Kao[6] | Daniel R. Westfall[3] | Chien-Ting Wu[7] | Keita Kamijo[8] | Darla M. Castelli[9] | Naiman A. Khan[6] | Arthur F. Kramer[3,10] | Charles H. Hillman[3,11]

[1] Department of Kinesiology, University of North Carolina at Greensboro, Greensboro, North Carolina, USA

[2] Department of Kinesiology, Michigan State University, East Lansing, Michigan, USA

[3] Department of Psychology, Northeastern University, Boston, Massachusetts, USA

[4] Department of Health and Physical Activity, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[5] Department of Exercise Science, University of South Carolina, Columbia, South Carolina, USA

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

[7] Exercise and Sport Science, University of South Carolina Upstate, Spartanburg, South Carolina, USA

[8] Faculty of Sport Sciences, Waseda University, Tokorozawa, Saitama, Japan

[9] Department of Kinesiology and Health Education, University of Texas at Austin, Austin, Texas, USA

[10] Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[11] Department of Health Sciences, Northeastern University, Boston, Massachusetts, USA

**Correspondence**
Charles H. Hillman, Ph.D., Northeastern University, Department of Psychology, 635 ISEC, 360 Huntington Avenue, Boston, MA 02115, USA.
Email: c.hillman@northeastern.edu

## Abstract

The present study investigated the effect of a 9-month physical activity (PA) intervention on children's cardiorespiratory fitness levels and neuroelectric indices of conflict monitoring (i.e., error-related negativity, ERN). Four hundred twenty-eight preadolescent children (8–9 years old) were randomized into a PA intervention or wait-list control group, and completed a fitness and cognitive control assessment (i.e., modified flanker task) at pre- and posttest. Following exclusion criterion, 308 children were included in the analyses (PA intervention: $n = 139$; wait-list control: $n = 169$). Children in the intervention displayed greater improvements in fitness and response accuracy, which were accompanied by stability of ERN amplitude from pre- to posttest. In contrast, the control group revealed increased ERN amplitude at posttest compared to pretest, despite no change in fitness or task performance. These findings demonstrate the efficacy of daily PA for promoting children's fitness and underlying neural processes associated with effective conflict monitoring. Such findings have significant implications for promoting organized PA programs intended to foster overall physical and brain health in school age children.

**KEYWORDS**
children, ERN, ERPs, exercise, fitness

**PSYCHOPHYSIOLOGY** SPR SOCIETY FOR PSYCHOPHYSIOLOGICAL RESEARCH

# 1 | INTRODUCTION

Research over the past decade is consonant in identifying physical activity (PA) as a marker for enhanced brain health and cognitive control performance (i.e., top-down goal-directed behavior) in developing youth (Erickson, Hillman, & Kramer, 2015). The implications surrounding these findings have considerable momentum for altering educational practice, yet current trends demonstrate that PA opportunities have been obviated from the school day (e.g., > 44% of school districts in the United States report reductions in physical education classes; Centers for Education Policy, 2007). Such trends may be maladaptive for underlying brain processes that support academic success in the classroom. However, only a small number of randomized controlled trials have utilized neuroimaging techniques to better understand functional brain mechanisms underlying these PA-induced changes in cognition (Chaddock-Heyman et al., 2013; Davis et al., 2011; Hillman et al., 2014; Kamijo et al., 2011; Krafft et al., 2014). Therefore, additional evidence is needed to establish strong empirical support for the importance of PA to cognitive and brain health, and provide the necessary rationale for driving public health change in developing youth.

Here, ERPs recorded during a modified flanker task were utilized to investigate the effect of a PA intervention on the neural underpinnings of conflict monitoring—one aspect of cognitive control (Larson, Clayson, & Clawson, 2014). Cognitive control encompasses a collection of higher-order cognitive operations responsible for guiding evaluative (i.e., conflict monitoring) and regulative (i.e., inhibition, attention) operations necessary for successful goal-directed behavior (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004). Regulative control includes the ability to allocate attention and adapt to a shifting environment by overriding prepotent responses to robust but irrelevant information (Bunge, Dudukovic, Thomason, Vaidya, & Gabrielli, 2002; Davidson, Amso, Anderson, & Diamond, 2006). Such processes have been found amenable to PA intervention in children (Hillman et al., 2014). For example, in a previous ERP investigation, preadolescent children assigned to a 9-month PA intervention demonstrated improved performance on a flanker task, with concomitant increases in neuroelectric measures of attentional resource allocation (i.e., P3-ERP component amplitude; Hillman et al., 2014). Such findings indicate that daily PA improves functional indices of cognitive control (specifically, attentional resource allocation) with associated behavioral improvements. However, this prior report focused primarily on regulative cognitive control operations and associated underlying functional processes (i.e., the P3 component). Accordingly, the present study sought to assess the effect of

PA-induced changes in evaluative control processes via neuroelectric indices of conflict monitoring. Such findings will provide new insight into the link between PA engagement with frontally mediated cognitive development and brain health in preadolescent children.

Effective cognitive control depends on the ability to detect conflict and make appropriate adjustments to improve subsequent behavior (i.e., conflict monitoring; Botvinick et al., 2001, 2004; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999). The anterior cingulate cortex (ACC) represents a main cortical contributor of these processes with neuroimaging investigations identifying the error-related negativity (ERN) as a neural marker of ACC conflict monitoring (Brazdil, Roman, Daniel, & Rektor, 2005). The ERN is a response-locked negative deflection (most prominent over the ACC or front of the scalp) often engendered by an error of commission. While prior research has suggested that the ERN component represents reinforcement learning of error detection (Holroyd & Coles, 2002), other work has indicated that the ERN represents the concurrent activation of multiple competing response options (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Coles, Meyer, & Donchin, 1990). For instance, when conflict awareness is heightened during task instructions that emphasize response accuracy over speed (Gehring, Goss, Coles, Meyer, & Donchin, 1993; Yeung, Botvinick, & Cohen, 2004), ERN amplitude is upregulated suggesting greater salience of an error and the allocation of cognitive control to effectively manage sources of conflict (Larson et al., 2014). Other modifications in task parameters that increase awareness to conflicting response options (i.e., spacing of flanking arrows or increasing size of stimuli; Danielmeier, Wessel, Steinhauser, & Ullsperger, 2009; Maier, di Pellegrino, & Steinhauser, 2012) also necessitate upregulation of cognitive control as reflected by larger ERN amplitude (Larson et al., 2014). Conversely, under conditions of greater conflict, individual differences revealing a decrease in ERN amplitude may suggest a lower threshold for conflict detection and subsequent error signaling (Botvinick et al., 2001; Carter et al., 2000).

Pertinent to this investigation, Themanson and colleagues (Themanson & Hillman, 2006; Themanson, Pontifex, & Hillman, 2008) evaluated the relation of cardiorespiratory fitness to ERN amplitude and flanker performance in young adults under conditions that modulated error salience. In a cross-sectional study, they observed smaller ERN amplitude for higher-fit compared to their lower-fit peers when task parameters emphasized speed over accuracy. In contrast, when accuracy was emphasized, higher fitness levels were associated with increased ERN amplitude and improvements in task performance. Additionally, findings revealed greater change in ERN amplitude associated with higher fitness when comparing between task conditions (accuracy vs.

speed). Collectively, these data suggest that higher fitness levels are associated with more efficient adjustments in conflict monitoring. That is, when error salience was heightened through task instruction, greater fitness was associated with efficient adjustments in top-down control necessary for appropriate corrective behavior.

However, recent investigations in children appear contrary to the ERN outcomes observed in young adults (Themanson et al., 2006, 2008). Specifically, findings in youth samples indicate that higher-fit children exhibit smaller ERN amplitude relative to lower-fit children when task instructions emphasize response accuracy (Pontifex et al., 2011). However, cognitive development research suggests that adults are more effective at modulating response selection (i.e., reaction time) across difficult trial types to maintain response accuracy. Conversely, children demonstrate a more impulsive response selection; thus, response accuracy measures more precisely reflect upregulation in inhibitory control processes in younger populations (Christakou et al., 2009; Davidson et al., 2006). Therefore, not surprising when accuracy is emphasized during task instruction in children, the effects of fitness on ERN demonstrate similar modulation patterns to that of previous adult investigations (emphasizing speed; Themanson et al., 2006), revealing smaller ERN amplitude for higher-fit compared to lower-fit children (Pontifex et al., 2011). These findings in children, together with the adult findings, affirm the relation of fitness to efficient adjustments in top-down control and further suggest that aerobic fitness may serve to influence underlying neural processes responsible for effective conflict monitoring and cognitive control operations across the lifespan. However, the directionality of the fitness-ERN relationship appears to be susceptible to multiple complex factors including stimulus characteristics, task instructions, and age.

Developmental factors also warrant consideration in the present investigation given the protracted development of the ERN and underlying cortical contributors (i.e., ACC) in young populations. Prior research demonstrates marked ACC development and greater increases in ERN amplitude occurring during late adolescent periods and into adulthood (Adleman et al., 2001; Crone, Zanolie, Van Leijenhorst, Westenberg, & Rombouts, 2008; Davies, Segalowitz, & Gavin, 2004; Segalowitz et al., 2010; Van Bogaert, Wikler, Damhaut, Szliwowski, & Goldman, 1998). However, the relation of larger ERN amplitude with age is not linear among prepubertal children and typically reveals maintenance of amplitude during the few years prior to pubertal onset (Meyer, Weinberg, Klein, & Hajcak, 2012; Van Meel, Heslenfeld, Rommelse, Oosterlaan, & Sergeant, 2012). For example, Meyer and colleagues (2012) evaluated ERN in children 8–13 years and found a nonsignificant trend for larger ERN amplitude with age only among older cohorts of

children. Similar findings revealed no difference in ERN amplitude between younger (6–9 years old) and older (10–12 years old) prepubertal children (Van Meel et al., 2012). Thus, based on these data, developmental factors in the present investigation are expected to be negligible given the young, prepubescent age (8–9 years old) and relatively short period between pre- and posttest assessments (9 months). However, a robust literature base of individual differences research reveals greater increases in ERN amplitude in children with obsessive-compulsive disorder (Hajcak, Franklin, Foa, & Simons, 2008) and clinical anxiety disorder compared to healthy cohorts (Ladouceur, Dahl, Birmaher, Axelson, & Ryan, 2006; Meyer et al., 2012), suggesting a maladaptive hyperactive error monitoring system during development (McDermott et al., 2009; Olvet & Hajcak, 2008). Taken together, younger children appear to maintain ERN amplitude with further evidence suggesting that greater increases in ERN may serve as a marker of atypical cognitive or psychological development. However, the investigation of other individual differences of health factors (i.e., physical activity) remain mostly unexplored.

The present investigation employed a 9-month afterschool PA intervention (Fitness Improves Thinking in Kids [FITKids]) to assess its impact on changes in children's cardiorespiratory fitness, neuroelectric indices of conflict monitoring (i.e., ERN amplitude), and associated task performance during a flanker task that modulated inhibitory control. We predicted that improvements in fitness engendered by the PA intervention would foster similar differences in task performance and ERN amplitude observed in earlier cross-sectional investigations in children. That is, given task instructions emphasizing accuracy and speed (i.e., increased error salience), we predicted improvements in task performance while the ERN would coincide with developmental and fitness research demonstrating maintenance or a reduction in amplitude associated with greater increases in fitness following the PA intervention compared to a wait-list control group. Such findings may further promote public health programming regarding the effects of physical activity on cognitive and brain health in young children.

## 2 | METHOD

The present investigation constitutes the evaluation of a subset of children who participated in either the FITKids (ClinicalTrials.gov: NCT01619826) or the FITKids2 (ClinicalTrials.gov: NCT01334359) clinical trials, and for simplicity's sake will be identified as the FITKids intervention throughout. Accordingly, this investigation represents a secondary analysis of the previously published FITKids dataset (Hillman et al., 2014). In the previous paper, correct trial response time (RT) and response accuracy were reported

along with the P3-ERP measure outcomes. In this secondary analysis, our aim was to assess conflict monitoring using the ERN and error-trial task performance. Task performance data reported herein from the subset of participants that overlap with Hillman et al. (2014) are presented only to better inform the novel ERN findings.

## 2.1 | Study design

The purpose of the FITKids intervention was to provide occasion for children to engage in moderate-to-vigorous PA for at least 70 min per weekday for one school year (9-months) with the intention of improving cardiorespiratory fitness. This was accomplished by randomizing all consenting participants into either the PA intervention or a wait-list control group following pretest assessment. A trained staff member accomplished the randomization procedure following baseline pretesting. Pairs of participants were matched on baseline demographics (age, gender, race, socioeconomic status [SES], IQ, and $VO_2$) and then assigned to respective groups based on a coin flip. Additionally, group assignment was blinded for analyses by indicating 1 or 0 associated with each participant ID number. Those randomized to the wait-list group were guaranteed placement in the after-school intervention the following school year (but not tested). Group assignment was blinded to all staff members involved with data collection. Pretest and posttest measures were collected over a 2-day period across the academic years 2009 to 2016.

## 2.2 | After-school physical activity intervention

The FITKids intervention used a modified version of the American Academy of Pediatrics Community Access to Child Health (CATCH; www2.aap.org/catch/) program developed by the National Institutes of Health. This program is structured such that daily lessons provide intermittent bouts of PA integrated with educational lessons intended to facilitate and encourage healthy lifestyle behaviors outside the program. Additionally, the efficacy of the CATCH program has been demonstrated in prior research to reduce unhealthy weight status (Coleman et al., 2005) and increase PA behaviors among young children (Luepker et al., 1996).

Utilizing the CATCH curriculum, the FITKids intervention occurred every weekday (in accordance with the annual academic calendar for local public school districts) following completion of regular school hours. Children were either bused or guardian-driven to a designated gymnasium rented from the University of Illinois Campus Recreation. During the intervention, children were provided at least 70 min of moderate-to-vigorous PA during a 2-hr period per day with intensity monitored utilizing Polar heart rate (HR) monitors

(Polar E600, Polar Electro, Finland) worn for the duration of the program. Wait-list control children were instructed to maintain daily/normal activities between the 9-month pre- and posttest assessments, thus providing unintervened representation of physical and cognitive development in our sample.

## 2.3 | Participants

Nine hundred twenty-nine children were screened and 428 children were randomly assigned to either the intervention or wait-list (see Figure 1). All participants provided written informed assent and their legal guardians provided written informed consent in accordance with the Institutional Review Board of the University of Illinois at Urbana-Champaign. Prior to testing, legal guardians completed a battery of health history and demographic questionnaires on behalf of the participant. Based on these questionnaires, all participants included in the analyses had not received special educational services from their school in connection with learning or attentional disorders, were free of neurological or physical disabilities, and had (corrected-to-)normal vision based on the minimal 20/20 standard.

## 2.4 | Procedure

Pre- and posttest protocols were identical with each assessment divided across two separate days. All participants, on their first visit to the laboratory, completed informed assent/consent, preliminary screening, demographic information, IQ assessment, and the cardiorespiratory fitness assessment. On the second laboratory visit, participants were fitted with an EEG cap and seated in a quiet testing chamber to complete the flanker task. All participants were instructed on appropriate completion of the flanker task and afforded practice prior to testing. Upon completion of posttesting, participants were briefed on the purpose of the study and received remuneration at a rate of $10/hour for their participation.

## 2.5 | Demographic assessment

A dichotomous SES index was used based on highest level of education obtained by the female guardian (Ogunwole, Drewery, & Rios-Vargas, 2012). Low SES was determined if the female guardian had no high school degree, a high school degree, or some college experience. High SES was determined if the female guardian obtained a bachelor or advanced degree. Hand dominance was determined by the Edinburgh Handedness Inventory (Oldfield, 1971). Pubertal timing was determined using the Tanner Staging System Questionnaire (Tanner, 1962; Taylor et al., 2001), which was completed by a legal guardian in collaboration with the
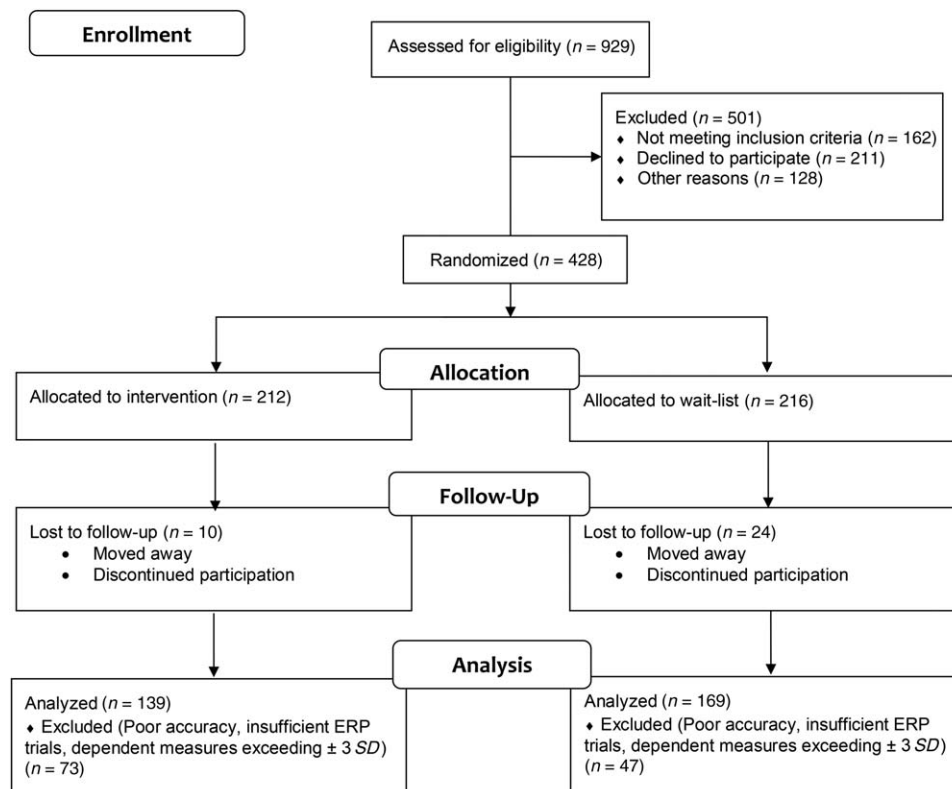
**FIGURE 1** Flow diagram of the FITKids intervention

participant. Body mass index (BMI) was calculated as weight divided by the square of height (i.e., $kg/m^2$). Per the Centers for Disease Control and Prevention BMI-for-age growth charts (Kuczmarski et al., 2002), children who were $\geq$ 85th percentile were classified as overweight, children who were $\geq$ 95th percentile were classified as obese. Lastly, an estimate of IQ was collected based on an age-normed standardized IQ examination, which was administered by a trained experimenter. IQ was assessed using the composite subtests of the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990) and the Woodcock-Johnson III Test of Cognitive Abilities (Woodcock, McGrew, & Mather, 2001) with a standard score of 100 representing average IQ with a standard deviation of 15.

## 2.6 | Cardiorespiratory fitness assessment

Each participant was fitted with a Polar HR monitor (Polar WearLink 31, Polar Electro), and measurements of height and weight were recorded (stadiometer and a Tanita WB-300 Plus digital scale) prior to the fitness assessment test. A modified Balke protocol (American College of Sports Medicine [ACSM], 2014) was used while participants ran on a motor-driven treadmill (Life Fitness 93T classic, Brunswick Corporation, Schiller Park, IL) at a constant speed. Sequentially, at 2-min intervals, treadmill incline increased 2.5%, and the OMNI scale was used to collect ratings of perceived

exertion (Utter, Robertson, Nieman, & Kang, 2002). Running was maintained until volitional exhaustion was achieved. Maximal aerobic capacity ($VO_{2peak}$) was measured during treadmill running by a computerized indirect calorimetry system (ParvoMedics True Max 2400) with averages of respiratory exchange ratio (RER) and oxygen uptake ($VO_2$) assessed every 20 s. Relative peak oxygen consumption was expressed in milliliters of oxygen consumed per kilogram of body weight per minute. $VO_{2peak}$ was based upon maximal effort as evidenced by a minimum of two of the following four criteria: (a) a plateau in oxygen uptake corresponding to an increase of less than 2 ml/kg·min$^{-1}$ despite an increase in exercise workload, (b) a peak heart rate $\geq$ 185 beats per minute (ACSM, 2014) and a heart rate plateau (Freedson & Goodman, 1993), (c) RER $\geq$ 1.0 (Bar-Or, 1983), and/or (d) ratings on the children's OMNI scale of perceived exertion $\geq$ 8 (Utter et al., 2002).

## 2.7 | Flanker task

Participants completed a modified version of the Eriksen flanker task (Eriksen & Eriksen, 1974). All stimuli were presented focally on a computer screen at a distance of approximately 1 m using Neuroscan STIM software (Compumedics NeuroScan, Charlotte, NC). Participants were instructed to respond with a thumb press on a response pad (Neuroscan STIM system) to the direction of the centrally presented

target (3-cm tall goldfish on a blue background) amid either congruent (e.g., [fish images] or [fish images]) or incongruent (e.g., [fish images] or [fish images]) flanking nontargets (identical goldfish). Task instructions and encouragement to participants prior to and following each task block emphasized response accuracy (i.e., "It is important that you respond as accurately as possible …") with secondary instructions encouraging response speed to maintain responding within the allotted response window (i.e., "… but we also want you to respond quickly so please make sure you respond before the next set of fish appears on the screen"). Prior to testing, participants were given 40 practice trials to control for potential practice effects. Two blocks of 75 trials were presented randomly with equiprobable congruency (i.e., 50% congruent trials) and directionality (50% of targets facing right) for 200 ms with a fixed intertrial interval (ITI) of 1,700 ms. These parameters were slightly modified during FITKids2 intervention such that trials were presented for 250 ms with random variable ITI of 1,600, 1,800, and 2,000 ms.

## 2.8 | ERP recording

EEG activity was recorded from 64 Ag/AgCl sintered electrode sites (FPz, Fz, FCz, Cz, CPz, Pz, POz, Oz, FP1/2, F7/5/3/1/2/4/6/8, FT7/8, FC3/1/2/4, T7/8, C5/3/1/2/4/6, M1/2, TP7/8, CB1/2, P7/5/3/1/2/4/6/8, PO7/5/3/4/6/8, O1/2) arranged in an extended montage according to the International 10-10 system (Chatrian, Lettich, & Nelson, 1985) using a Neuroscan Quickcap (Compumedics Neuroscan). Prior to EEG recordings, electrode impedance was maintained at < 10 kΩ. Online, continuous data were referenced to a midline electrode placed at midpoint between Cz and CPz with AFz serving as the ground electrode. Additional electrodes were placed above and below the left orbit and outer canthus of each eye to monitor electrooculographic (EOG) activity with separate bipolar recordings to monitor vertical (VEOG) and horizontal activity. Continuous online data were digitized at a sampling rate of 500 Hz, amplified 500 times with a DC to 70 Hz filter, and a 60-Hz notch filter was applied using Neuroscan SynAmps2 amplifier.

Offline, continuous EEG data were processed utilizing MATLAB (R2012b) and various toolbox plug-ins including EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2010). EEG data were rereferenced to averaged mastoids (M1, M2). Independent component analyses (ICA) were conducted to identify stereotypical eyeblink artifact (Comon, 1994) followed by an autocorrelation procedure developed for rejecting ICA components related to VEOG activity. This was accomplished by correlating point-by-point raw VEOG data with separate ICA activation waveforms (i.e., EEG.icaact matrix generated by the ICA

procedure). No more than two ICA components with a correlation coefficient greater than 0.30 were removed.

Response-locked epochs were extracted for correct and commission-error trials (−600 to 1,000 ms relative to response onset), baseline corrected (−400 to −200 ms), and band-pass filtered (zero phase shift 1 to 12 Hz with 24 dB/octave roll-off). Epochs were rejected if a moving window peak-to-peak amplitude exceeded 100 μV (100-ms window width and 50-ms window step) and if the overall variance of the epoch exceeded ± 3 $SD$s of the mean of local (by electrode site) and global (all electrode sites) accepted epochs. ERP average waveforms (averaged across congruent and incongruent trial types) were computed as the difference between commission-error (ERN) and correct (CRN) response-locked averages ($ERN_{error - correct}$) to account for artifacts that may be due to activity beyond the confines of error-specific activity (Torpey, Hajcak, Kim, Kujawa, & Klein, 2012). The CRN, ERN, and $ERN_{error - correct}$ components were characterized using a peak interval technique that measured the mean amplitude (30-ms window) encompassing the largest negative-going peak identified within a −50 to 100 ms latency window relative to response onset.

## 2.9 | Statistical analysis

Prior to statistical analyses, participants were excluded for insufficient response-locked trials ($n = 75$) necessary to characterize ERP components (< 6 commission error trials; Pontifex et al., 2010), poor performance on the flanker task (< 50% accuracy; $n = 20$), and dependent measures exceeding ± 3 $SD$ ($n = 25$). Thus, statistical and multiple imputation procedures, utilizing SPSS (v. 22, SPSS, Chicago, IL), were conducted on the remaining participants ($n = 308$; $n = 151$ stem from previously published work, Hillman et al., 2014) following these exclusion criteria.

To account for missing data at posttest (∼12%), all missing values were imputed with 20 iterations for each data point. Pretest demographic measures, cardiorespiratory fitness, and BMI change scores (i.e., post − pre), were evaluated between wait-list and intervention groups with independent $t$ tests. Multivariate repeated measures multivariate analyses of variance (MANOVAs) were performed to evaluate the efficacy of the intervention with group as the between-subjects factor. Findings are reported with a family-wise alpha threshold for all tests set at $p = .05$. Subsidiary univariate ANOVAs were used for post hoc procedures with contrasting $p$ values adjusted using the Sidak correction method for multiple comparisons (i.e., a reported $p$ value of .05 is significant). Further reporting included partial $\eta^2$ and estimated effect size for main effects and interactions. Task performance (accuracy, median reaction time) were separately analyzed utilizing a 2 (Group: intervention, wait-list)

× 2 (Time: pretest, posttest) × 2 (Type: congruent, incongruent) statistical model. Remaining behavioral analyses (i.e., commission and omission errors and error runs, post-error accuracy, post-error latency) were separately analyzed utilizing a 2 (Group: intervention, wait-list) × 2 (Time: pretest, posttest) statistical model with decomposition of non-parametric interactions accomplished utilizing Mann-Whitney (between-subjects) and Wilcoxon signed-rank tests (within-subject). ERP analyses were performed at single electrode sites based on scalp distribution and previous research indicating component maximum at frontocentral (FCz) site (Clayson & Larson, 2012; Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003). ERN, CRN, and $ERN_{error - correct}$ were analyzed separately utilizing a 2 (Group: treatment, wait-list) × 2 (Time: pretest, posttest) statistical model. Lastly, planned Pearson partial correlations were performed to evaluate the relation of change in fitness from pre- to posttest to task performance and ERN changes observed separately for each treatment group. This was accomplished by performing within-group partial correlations on change scores (i.e., post − pre) while controlling for statistically correlated demographic and pretest confounding factors including age, IQ, sex, pretest BMI, and pretest fitness. Lastly, previous analyses of the FITKids intervention (Hillman et al., 2014) demonstrated significant differences in flanker performance at baseline between groups, such that children in the intervention demonstrated poorer performance compared to children in the wait-list group. Thus, the present investigation performed confirmatory analyses on a subsample of participants matched on flanker accuracy at baseline to more accurately represent PA-induced changes unbiased by pretest differences. This was accomplished by utilizing a case-controlled matching technique in SPSS pairing treatment and wait-list groups per overall flanker accuracy with a match tolerance set at five and treatment as the group indicator. Subsample analyses (treatment: $n = 139$; wait-list: $n = 139$) are reported post hoc.

## 3 | RESULTS

### 3.1 | Participants

Means and standard error of the mean ($\pm 1$ *SEM*) are reported for demographic measures provided in Table 1. Analyses of baseline demographics revealed no significant differences between groups on measures of age, cardiorespiratory fitness, BMI, IQ, and pubertal timing, $ts(306) \leq 1.43$, $ps \geq .15$. Children in the FITKids intervention attended 122 ($\pm 1.7$) sessions, which equated to an attendance rate of 85.4% ($\pm 0.8$%). Analyses were conducted on the number of ERP error-trials to ensure that group differences were not influenced by trial count included in response-locked averages. Results revealed

**TABLE 1** Mean ($\pm 1$ *SEM*) values for participant demographic and fitness measures

| Measure | Intervention | Wait-list |
|---|---|---|
| *n* | 139 | 169 |
| Female, *n* (%) | 71 (51) | 86 (51) |
| Race, *n* (%) | | |
|   Asian | 17 (12) | 15 (9) |
|   African American | 30 (22) | 37 (22) |
|   White | 63 (46) | 88 (52) |
|   Other or mixed | 28 (20) | 26 (15) |
|   Hispanic | 13 (9) | 12 (7) |
| Low SES (%) | 55 (40) | 70 (41) |
| Overweight/obese (%) | 62 (45) | 77 (46) |
| Age (years) | $8.72 \pm 0.05^a$ | $8.75 \pm 0.04^a$ |
| IQ | $108.04 \pm 1.18^a$ | $109.24 \pm 1.01^a$ |
| Puberty timing | $1.42 \pm 0.04^a$ | $1.45 \pm 0.04^a$ |
| ΔBMI (kg/m$^2$) | $0.15 \pm 0.11^a$ | $0.70 \pm 0.19^b$ |
| $\Delta VO_{2max}$ (mL/kg/min) | $1.8 \pm 0.40^a$ | $0.61 \pm 0.35^b$ |
| Days attended | $122 \pm 1.68$ | n/a |
| Attendance rate (%) | $85.4 \pm 0.84$ | n/a |

*Note.* Low SES was determined if the female guardian had no high school degree, a high school degree, or some college experience. Data are presented as mean ($\pm 1$ *SEM*) unless noted otherwise. Values sharing a common superscript are not statistically different at $p \leq .05$. SES = socioeconomic status; IQ = intelligent quotient; BMI = body mass index; n/a = not applicable.

no main effect or interaction involving group, $Fs(1, 306) \leq 2.1$, $ps \geq .16$, $\eta_p^2 \leq .01$, indicating an equivalent trial count across groups with an average of 21 ($\pm 0.6$) error trials per participant. Lastly, analysis of change scores revealed a three-fold increase in fitness from pre- to posttest for the intervention group (pre: $39.1 \pm 0.6$ mL/kg/min, post: $40.9 \pm 0.6$ mL/kg/min; $\Delta 1.8 \pm 0.4$ mL/kg/min; 5.4% change) compared to the wait-list group (pre: $39.9 \pm 0.6$ mL/kg/min; post: $40.5 \pm 0.6$ mL/kg/min; $\Delta 0.6 \pm 0.3$ mL/kg/min; 2.1% change; see Table 1), $t(306) = 2.3$, $p = .02$, demonstrating the efficacy of the PA intervention.

### 3.2 | Flanker performance

#### 3.2.1 | Reaction time

The omnibus analysis for RT revealed a main effect of time, $F(1, 306) = 51.0$, $p \leq .01$, $\eta_p^2 = .14$, demonstrating shorter RT at posttest ($467.2 \pm 5.2$ ms) compared to pretest ($505.2 \pm 6.5$ ms) for all participants. In addition, a main effect of type was observed, $F(1, 306) = 385.2$, $p \leq .01$,

**PSYCHOPHYSIOLOGY** SPR

**TABLE 2** Mean ($\pm 1$ *SEM*) values for participant flanker task performance measures

| Measure | Intervention | | Wait-list | |
|---|---|---|---|---|
| | Pretest | Posttest | Pretest | Posttest |
| Reaction time (ms) | $498.8 \pm 9.9^a$ | $468.4 \pm 7.5^b$ | $507.0 \pm 8.2^a$ | $460.5 \pm 7.3^b$ |
| Accuracy (%) | $71.4 \pm 1.3^a$ | $80.5 \pm 1.0^b$ | $74.4 \pm 1.0^c$ | $80.4 \pm 1.0^b$ |
| Subsampling accuracy (%) | $71.4 \pm 1.3^a$ | $80.5 \pm 1.0^b$ | $74.1 \pm 1.1^a$ | $80.3 \pm 1.0^b$ |
| Commission errors (#) | $24.2 \pm 1.0^a$ | $17.3 \pm 0.8^b$ | $23.0 \pm 1.0^a$ | $17.2 \pm 0.9^b$ |
| Commission error runs (#) | $4.2 \pm 0.3^a$ | $2.6 \pm 0.3^b$ | $3.8 \pm 0.3^a$ | $2.6 \pm 0.2^b$ |
| Omission errors (#) | $19.6 \pm 1.3^a$ | $10.7 \pm 0.9^b$ | $16.1 \pm 1.2^a$ | $11.7 \pm 0.8^b$ |
| Omission error runs (#) | $3.7 \pm 0.3^a$ | $1.6 \pm 0.2^b$ | $2.6 \pm 0.3^a$ | $1.9 \pm 0.2^b$ |
| Post-error reaction time (ms) | $554.8 \pm 11.9^a$ | $496.5 \pm 9.0^b$ | $560.0 \pm 10.6^a$ | $492.4 \pm 8.9^b$ |
| Post-error accuracy (%) | $69.8 \pm 1.5^a$ | $78.3 \pm 1.5^b$ | $71.8 \pm 1.4^a$ | $78.1 \pm 1.6^b$ |

*Note.* Values for each measure are collapsed across congruent and incongruent trials. Subsampling accuracy represents participants matched on flanker accuracy at baseline (treatment: $n = 139$, wait-list: $n = 139$); Values sharing a common superscript are not statistically different at $p \leq .05$.

$\eta_p^2 = .56$, indicating shorter RT for congruent ($465.4 \pm 5.0$ ms) compared to incongruent ($506.9 \pm 5.7$ ms) trials. No other main effects or interactions were observed for RT, $Fs$ $(1, 306) \leq 2.1$, $ps \geq .15$, $\eta_p^2 \leq .01$.

### 3.2.2 | Accuracy

The omnibus analysis for response accuracy revealed a main effect of type, $F(1, 306) = 502.1$, $p \leq .01$, $\eta_p^2 = .62$, indicating higher accuracy for congruent ($80.5 \pm 0.7\%$) compared to incongruent ($72.8 \pm 0.7\%$) trials. In addition, a main effect of time was observed, $F(1, 306) = 118.1$, $p \leq .01$, $\eta_p^2 = .28$, which was superseded by a Group $\times$ Time interaction, $F(1, 306) = 5.3$, $p = .02$, $\eta_p^2 = .02$. Decomposition of this interaction revealed increased response accuracy for both the wait-list and intervention groups at posttest (wait-list: $80.4 \pm 0.9\%$, intervention: $80.5 \pm 1.0\%$) compared to pretest (wait-list: $74.4 \pm 1.1\%$, intervention: $71.3 \pm 1.2\%$; see Table 2), $Fs(1, 306) \geq 40.7$, $ps \leq .01$. Between-groups comparisons at pre- and posttest revealed higher accuracy for the wait-list group relative to the intervention group only at pretest, $F(1, 306) = 3.9$, $p = .05$. Although posttest measures were not significant between groups, $F(1, 306) = 0.0$, $p = .93$, change score analysis demonstrated greater change from pre- to posttest for the intervention group ($9.3 \pm 1.0\%$) compared to the wait-list group ($6.0 \pm 1.0\%$), $F(1, 306) = 5.1$, $p = .02$. No other main effects or interactions were observed for accuracy, $Fs(1, 306) \leq 1.6$, $ps \geq .21$, $\eta_p^2 \leq .01$. Lastly, subsampling results that matched participants between groups based on pretest response accuracy scores were consonant with the above observed effects (i.e., greater change from pre- to posttest for the intervention group, $F(1, 276) = 4.3$, $p = .04$). Confirming

the success of the matching procedure, no significant differences in performance were observed at pretest between groups, $F(1, 276) = 2.9$, $p = .09$, (see Table 2), suggesting that greater changes in flanker performance for the intervention group remain regardless of pretest behavior differences.

### 3.2.3 | Error performance

The omnibus analysis for commission errors, commission error runs, omission errors, and omission error runs revealed a main effect of time, $Fs(1, 306) \geq 38.0$, $ps \leq .01$, $\eta_p^2 \geq 0.11$, with fewer errors and error runs at posttest ($16.4 \pm 0.6$ commission errors, $2.5 \pm 0.2$ commission error runs, $10.7 \pm 0.6$ omission errors, $1.6 \pm 0.1$ omission error runs) compared to pretest ($22.5 \pm 0.7$ commission errors, $3.8 \pm 0.2$ commission error runs, $17.0 \pm 0.8$ omission errors, $3.0 \pm 0.2$ omission error runs). Additionally, only omission errors and omission error runs revealed an interaction of Group $\times$ Time, $Fs(1, 306) \geq 6.1$, $ps \leq .02$, $\eta_p^2 \geq .02$. Decomposition of these interactions utilizing Mann-Whitney (between-subjects) and Wilcoxon signed-rank tests (within-subject) for nonparametric comparisons revealed that the wait-list and intervention groups both exhibited greater omission errors and error runs at pretest (intervention: $19.6 \pm 1.3$ omission errors, $3.7 \pm 0.3$ omission error runs; wait-list: $16.1 \pm 1.2$ omission errors, $2.6 \pm 0.3$ omission error runs) compared to posttest (intervention: $10.7 \pm 0.9$ omission errors, $1.6 \pm 0.2$ omission error runs; wait-list: $11.7 \pm 0.8$ omission errors, $1.8 \pm 0.2$ omission error runs; see Table 2) $Ws \geq 4,527$, $zs \geq 3.7$, $ps \leq .01$, with no differences observed between groups at pre- and posttest, $Us \leq 12,754$, $zs \leq 1.5$, $ps \geq .13$. However, analyses of change scores revealed a

greater reduction in omission errors and error runs from pre- to posttest for the intervention group ($-8.5 \pm 1.3$ omission errors, $-2.0 \pm 0.3$ omission error runs) compared to the wait-list group ($-4.1 \pm 1.2$ omission errors, $-0.7 \pm 0.3$ omission error runs), $Fs(1, 306) \geq 6.1$, $ps \leq .02$, $\eta_p^2 \geq .02$.

The omnibus analysis for post-error accuracy and RT revealed a main effect of time, $Fs(1, 306) \geq 42.3$, $ps \leq .01$, $\eta_p^2 \geq .12$, demonstrating increased accuracy and shorter RT at posttest (accuracy: $78.3 \pm 1.0\%$, RT: $495.2 \pm 6.2$ ms) compared to pretest (accuracy: $70.8 \pm 1.0\%$, RT: $557.4 \pm 8.0$ ms; see Table 2). No main effect of group or interaction of Group $\times$ Time were revealed for post-error accuracy or RT, $Fs(1, 306) \leq 1.2$, $ps \geq .30$, $\eta_p^2 \leq .01$.

## 3.3 | ERPs

The omnibus analysis for CRN amplitude revealed a main effect of time, $F(1, 306) = 9.9$, $p \leq .01$, $\eta_p^2 = .03$, demonstrating larger negative-going amplitude at pretest ($0.4 \pm 0.2$ μV) compared to posttest ($1.0 \pm 0.2$ μV). No main effect of group or interaction of Group $\times$ Time was observed, $Fs(1, 306) \leq 0.2$, $ps \geq .70$, $\eta_p^2 \leq .01$. The omnibus analyses for ERN and ERN$_{error - correct}$ amplitude revealed a main effect of time, $Fs(1, 306) \geq 5.6$, $ps \leq .04$, $\eta_p^2 \geq .02$, that was superseded by a Group $\times$ Time interaction, $Fs(1, 306) \geq 4.4$, $ps \leq .05$, $\eta_p^2 \geq .01$. Decomposition of the interaction revealed larger negative-going ERN and ERN$_{error - correct}$ amplitude at posttest (ERN: $-6.3 \pm 0.4$ μV; ERN$_{error - correct}$: $-9.9 \pm 0.5$ μV) compared to pretest (ERN: $-5.0 \pm 0.3$ μV; ERN$_{error - correct}$: $-7.7 \pm 0.3$ μV), only for the wait-list group, $Fs(1, 306) \geq 9.6$, $ps \leq .01$. Conversely, the intervention group demonstrated stability of ERN amplitude with no differences observed from pretest (ERN: $-5.7 \pm 0.3$ μV; ERN$_{error - correct}$: $-8.5 \pm 0.4$ μV) to posttest (ERN: $-5.8 \pm 0.4$ μV; ERN$_{error - correct}$: $-9.0 \pm 0.4$ μV; see Figure 2), $Fs(1, 306) \leq 0.6$, $ps \geq .46$. Additionally, analyses of change scores revealed increased negative-going ERN amplitude from pre- to posttest only for the wait-list group ($\Delta$ERN: $-1.1 \pm 0.5$ μV; $\Delta$ERN$_{error - correct}$: $-2.1 \pm 0.5$ μV) with no such effect observed for the intervention group ($\Delta$ERN: $-0.1 \pm 0.4$ μV; $\Delta$ERN$_{error - correct}$: $-0.5 \pm 0.4$ μV; see Figure 2a), $Fs(1, 306) \geq 10.2$, $ps \leq .01$. Lastly, subsampling results revealed no effect for $\Delta$ERN, $F(1, 276) = 3.39$, $p = .13$; however, $\Delta$ERN$_{error - correct}$ results were consonant with the initially observed effects, $F(1, 276) = 8.2$, $p = .02$, indicating greater changes in ERN$_{error - correct}$ from pre- to posttest for the wait-list group regardless of task performance differences at pretest.

## 3.4 | Correlation

Planned Pearson partial correlations for the intervention group revealed a significant association with increases in

fitness related to decreases in ERN$_{error - correct}$ amplitude from pre- to posttest, $r = .20$, $p = .02$ (see Figure 3). No further correlations of task performance measures or ERN with $\Delta$fitness were observed for the intervention, $rs \leq .09$, $ps \geq .26$, or wait-list, $rs \leq .12$, $ps \geq .10$, groups. Lastly, a one-tailed Fisher's $z$ test for difference of correlation strength was performed to determine if the relation of $\Delta$fitness to $\Delta$ERN$_{error - correct}$ observed for the intervention group was meaningfully different from the null results observed for the wait-list group. Results were significant, $z = 2.4$, $p \leq .01$, indicating that improving fitness by means of PA intervention has significant effect on reductions in ERN amplitude.

## 4 | DISCUSSION

This investigation demonstrated that a 9-month PA intervention improved fitness and behavioral indices of cognitive control in preadolescent children relative to children randomized to a wait-list control group. Additionally, children randomized to the PA intervention exhibited maintenance of ERN amplitude following the intervention, while those assigned to the wait-list group exhibited larger ERN amplitude 9 months later. Lastly, for those enrolled in the intervention group, greater increases in fitness were related to greater reductions in ERN amplitude, with no such effect observed for the wait-list group. Collectively, these results support the promotion of organized PA interventions in school age children, with positive implications for brain health and function that underlie effective school-based learning.

As expected, behavioral results are consonant with previous analyses of the FITKids intervention indicating greater improvements in task performance for the intervention compared to the wait-list group (Chaddock-Heyman et al., 2013; Hillman et al., 2014). However, unique to the present investigation, error analyses revealed greater reductions in omission errors and omission error runs for the intervention group. These findings align closely with prior cross-sectional results wherein higher-fit children demonstrated fewer omission errors and error runs on a flanker task compared to lower-fit children, suggesting poorer attentional vigilance among lower-fit children (Pontifex, Scudder, Drollette, & Hillman, 2012). Previous research suggests that these performance patterns occur in response to attentional resource depletion and subsequent temporal remedial measures for regeneration of neural resources and reengagement of attention (Fisk & Schneider, 1981). Accordingly, the present data further indicate that changes in fitness engendered by a daily PA intervention are associated with enhanced sustained attention necessary for superior cognitive control performance.

The ERN patterns observed in the current study provide further support for the a priori hypothesis regarding the effect of PA on adjustments in conflict monitoring. That is, larger
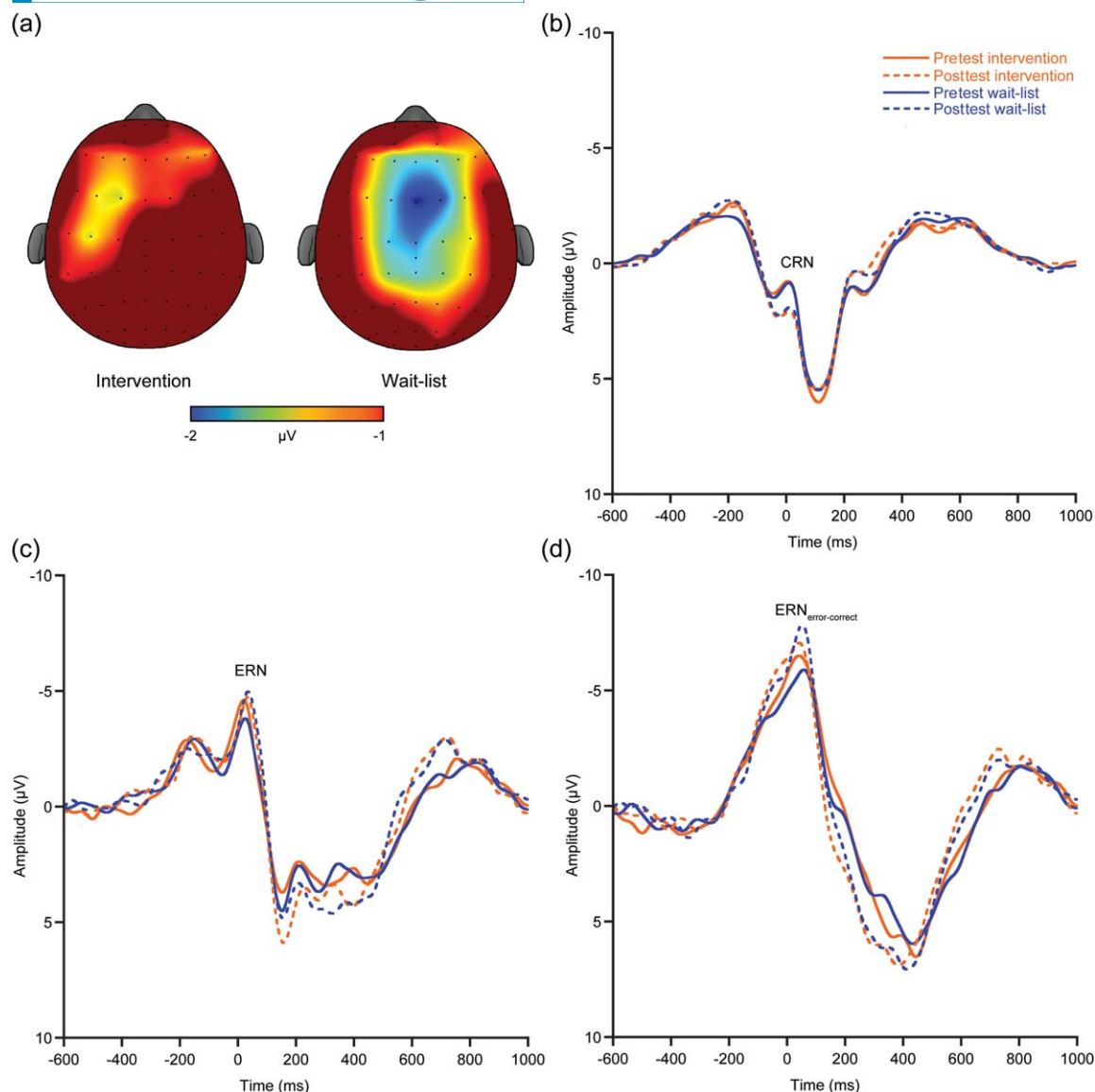
**PSYCHOPHYSIOLOGY** SPR

**FIGURE 2** (a) Topographic scalp distribution of change in ERN$_{error-correct}$ amplitude from pre- to posttest (i.e., posttest − pretest) by intervention (left) and wait-list (right) groups (time window plotted between −2 and −1 μV) for commission errors committed during the flanker task. Response-locked grand-averaged waveforms at site FCz for (b) CRN, (c) ERN, and (d) ERN$_{error-correct}$ separated by group and pre- and posttest

increases in ERN amplitude were observed for the wait-list group, with no change observed for children randomized to the intervention. The maintenance of ERN amplitude among the intervention group, along with larger improvements in task performance, suggest greater neural efficiency as evidenced by stable levels of conflict monitoring despite improved performance at posttest. Further, correlational analyses demonstrated a dose-response relationship with change in fitness, such that greater improvements in fitness from pre- to posttest were related to greater reductions in ERN amplitude—a finding that was selective to children in the intervention. These results, together with prior cross-sectional fitness investigations in children (Pontifex et al., 2011) and young adults (Themanson et al., 2006, 2008), are in accordance with conflict monitoring theory (Botvinick

et al., 2001), which suggests that reductions in ACC activation and ERN amplitude reflect a relative reduction in error-salient conflict facilitating greater increases in top-down cognitive control (Carter et al., 2000). Similar reductions in neural response patterns are observed in separate neuroimaging interventions in children indicating decreased hemodynamic activation in the frontal cortex (Chaddock-Heyman et al., 2013) and decreased resting state synchrony (i.e., task-free resting state activation) associated with the default mode network (i.e., predisposition for goal-directed cognitive operations; Krafft et al., 2014) for children randomized to PA interventions. These prior reports, in conjunction with our present findings, suggest that PA participation at a young age influences frontally mediated brain networks during pre-adolescent development in a manner that appears more
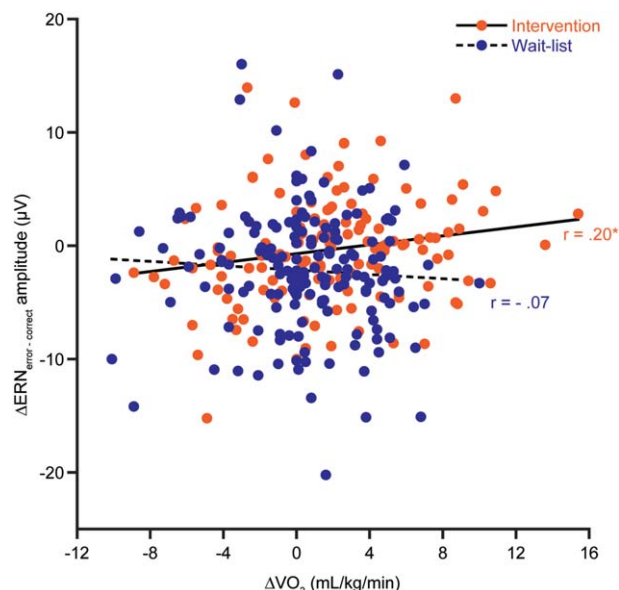
**FIGURE 3** Scatter plot for the bivariate relation between changes in cardiorespiratory fitness and changes in ERN amplitude separated by groups. This figure illustrates the partial correlation results (controlling for age, IQ, sex, pretest fitness, and pretest BMI) demonstrating a positive relation of increased fitness with reductions in ERN amplitude from pre- to posttest for the intervention ($r = .20$), with no such relation observed for the wait-list group ($r = -.07$)

efficient (i.e., downregulation in activation) for effective cognitive control and performance monitoring behavior.

Following this line of reasoning, the present results further suggest that physical inactivity may be a marker for atypical cognitive development among young children. Although literature investigating between-subjects variation in ERN and aspects of cognitive control is limited, previous research investigating ERN as a biomarker of abnormal personality and psychological disorders may shed light on this speculation such that increased ERN was observed in youth and young adults with symptoms of obsessive-compulsive disorder (Hajcak & Simons, 2002, 2008; Santesso, Segalowitz, & Schmidt, 2006), heightened levels of negative affect (Hajcak, McDonald, & Simons 2004; Luu, Collins, & Tucker, 2000), and anxiety (Ladouceur et al., 2006; Meyer et al., 2012) compared to typically developing cohorts. However, assessments of psychological disorders were beyond the scope of the present investigation. Thus, additional longitudinal studies are needed to evaluate the influence of PA and fitness on the developmental trajectory of the ERN and its relations with atypical psychological development in children. Regardless, this area of research warrants further consideration relative to how inactivity might influence brain and behavior, especially in developing children.

Further, our ERN results provide compelling evidence that links changes in brain function to changes in fitness stemming from an after-school PA program. However, no

PA-induced changes in post-error task performance were observed. Prior investigations suggest that post-error slowing is a behavioral marker of remedial adjustments associated with ERN production (see Larson et al., 2014, for review). However, such relations are not consistent in the literature, suggesting that conflict signaled by the ERN may not be directly related to remedial behavior measures (Hajcak & Simons, 2008). Therefore, it is possible that behavioral responses to conflict are necessitated by additional and separable neural systems (Larson et al., 2014), and that dynamic changes in ERN amplitude may not relate to sustained adjustments in cognitive control following errors (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001). Thus, the absence of significant findings for post-error performance in the present investigation may highlight the selectivity of PA participation on conflict monitoring particularly for neural processes involved with conflict monitoring and ERN production.

One possible mechanism for the observed PA-induced changes may include the upregulation and coexpression of neurotrophic growth factors (i.e., brain-derived neurotrophic factors [BDNF], insulin-like growth factor-1 [IGF-1]) and dopamine. Specifically, research reveals codependency of these two factors in nonhuman animal models (Hyman et al., 1991; Kuppers & Beyer, 2001; Williams & Undieh, 2009), demonstrating that prolonged PA engagement upregulates production and binding of dopamine (Foley & Fleshner, 2008; Marques et al., 2008) and BDNF (see Knaepen, Goekint, Heyman, & Meeusen, 2010, for review). Interestingly, increased dopamine in the brain is associated with ERN modulation in humans (see Jocham & Ullsperger, 2009, for review). Given that these neurochemicals facilitate proliferation and cell survival in various brain structures associated with cognitive control and memory operations (van Praag, Christie, Sejnowski, & Gage, 1999), the current ERN findings suggest that PA may facilitate neurochemical activation necessary for appropriate neurocognitive function. However, this remains speculative at this time given that BDNF and dopamine measures were not assessed in the present investigation. Regardless, there remains untapped avenues for future investigations to extend these findings and potentially isolate mechanistic factors associated with PA and brain development in young children.

As such, certain limitations are worth noting. Most importantly, a control group that received a nonphysical activity experience was not included in the study design. Thus, it is difficult to attribute the observed effects solely to PA because other factors may have differed between groups including social interaction, physical education instruction, and continual positive affirmation from adult figures. Additionally, the present investigation restricted the age range to 8- and 9-year-old children; therefore, the extent to which the

present results are observed in other domains of cognition, as well as across developmental age groups, remains unknown. Future research encompassing a developmental perspective may provide additional insight into the interaction of PA behaviors and the developmental time course of brain and cognitive operations.

In summary, this investigation provides further evidence for the utility of PA interventions as a means of optimizing cognition in preadolescent children, with selective effects observed for frontally mediated performance monitoring systems associated with cognitive control. Such findings provide valuable information for future investigations aimed at isolating and identifying cognitive operations and cortical mechanisms modified by PA participation. Further, these patterns of improved cognitive performance, combined with fitness changes, advocate for the importance of interventions aimed at all preadolescent children, not only for overall health but also for improving cognitive faculties necessary for effective learning.

## ORCID

*Matthew B. Pontifex* http://orcid.org/0000-0001-9819-9625

## REFERENCES

Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., & Reiss, A. L. (2001). A developmental fMRI study of the Stroop color-word task. *NeuroImage*, *16*, 61–75. https://doi.org/10.1006/nimg.2001.1046

American College of Sports Medicine. (2014). *ACSM's guidelines for exercise testing and prescription* (9th ed.). New York, NY: Loppincott Williams & Wilkins.

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

Botvinick, M. M., Braver, T. S., Barch, D., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652. https://doi.org/10.1037//0033-295X.108.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*, 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection for action in anterior cingulate cortex. *Letters to Nature*, *402*, 179–181.

Brazdil, M., Roman, R., Daniel, P., & Rektor, I. (2005). Intracerebral error-related negativity in a simple go/no-go task. *Journal of Psychophysiology*, *19*, 244–255. https://doi.org/10.1027/0269-8803.19.4.244

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrielli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*, 301–311. https://doi.org/10.1016/S0896-6273(01)00583-9

Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, A., Noll, D., & Cohen, J. D. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *PNAS*, *97*, 1944–1948. https://doi.org/10.1073/pnas.97.4.1944

Centers for Education Policy. (2007). *Choices, changes, and challenges: Curriculum and instruction in the NCLB era*. Washington DC: Center for Education Policy.

Chaddock-Heyman, L., Erickson, K. I., Voss, M. W., Knecht, A. M., Pontifex, M. B., Castelli, D. M., … Kramer, A. F. (2013). The effects of physical activity on functional MRI activation associated with cognitive control in children: A randomized controlled intervention. *Frontiers in Human Neuroscience*, *7*, 1–13. https://doi.org/10.3389/fnhum.2013.00072

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity. *American Journal of EEG Technology*, *25*, 83–92.

Christakou, A., Halari, R., Smith, A. B., Ifkovits, E., Brammer, M., & Rubia, K. (2009). Sex-dependent age modulation of frontostriatal and temporo-parietal activation during cognitive control. *NeuroImage*, *48*, 223–236. https://doi.org/10.1016/j.neuroimage.2009.06.070

Clayson, P. E., & Larson, M. J. (2012). Cognitive performance and electrophysiological indices of cognitive control: A validation study of conflict adaptation. *Psychophysiology*, *49*, 627–637. https://doi.org/10.1111/j.1469-8986.2011.01345.x

Coleman, K. J., Tiller, C. L., Sanchez, J., Heath, E. M., Sy, O., Milliken, G., & Dzewaltowski, D. A. (2005). Prevention of the epidemic increase in child risk of overweight in low-income schools: The El Paso coordinated approach to child health. *Archives of Pediatric Adolescent Medicine*, *159*, 217–224. https://doi.org/10.1001/archpedi.159.3.217

Comon, P. (1994). Independent component analysis—a new concept? *Signal Processing*, *36*, 287–314. https://doi.org/10.1016/0165-1684(94)90029-9

Crone, E., Zanolie, K., Van Leijenhorst, L., Westenberg, P., & Rombouts, S. (2008). Neural mechanisms supporting flexible performance adjustment during development. *Cognitive, Affective, &*

*Behavioral Neuroscience*, 8, 165–177. https://doi.org/10.3758/CABN.8.2.165

Danielmeier, C., Wessel, J. R., Steinhauser, M., & Ullsperger, M. (2009). Modulation of the error-related negativity by response conflict. *Psychophysiology*, 46, 1288–1298. https://doi.org/10.1111/j.1469-8986.2009.00860.x

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. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology*, 25, 355–376. https://doi.org/10.1207/z15326942dn2503 6

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, 91–98. https://doi.org/10.1037/a0021766

Delorme, A., & Makeif, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Erickson, K. I., Hillman, C. H., & Kramer, A. F. (2015). Physical activity, brain, and cognition. *Current Opinion in Behavioral Sciences*, 4, 27–32. https://doi.org/10.1016/j.cobeha.2015.01.005

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

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78, 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Fisk, A., & Schneider, W. (1981). Control and automatic processing during tasks requiring sustained attention: A new approach to vigilance. *Human Factors*, 23(6), 737–750.

Foley, T. E., & Fleshner, M. (2008). Neuroplasticity of dopanine circuits after exercise: Implications for central fatigue. *NeuroMolecular Medicine*, 10, 67–80. https://doi.org/10.1007/s12017-008-8032-3

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.

Gehring, W. J., Coles, M. G., Meyer, D. E., & Donchin, E. (1990). The error-related negativity: An event-related potential accompanying errors. *Psychophysiology*, 27, S34.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, 4, 385–390.

Hajcak, G., Franklin, M. E., Foa, E. B., & Simons, R. F. (2008). Increased error-related brain activity in pediatric obsessive-cumpulsive disorder before and after treatment. *American Journal of Psychiatry*, 165, 116–123. https://doi.org/10.1176/appi.ajp.2007.07010143

Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition*, 56, 189–197. https://doi.org/10.1016/j.bandc.2003.11.001

Hajcak, G., & Simons, R. F. (2002). Error-related brain activity in obsessive-compulsive undergraduates. *Psychiatry Research*, 110, 63–72. https://doi.org/10.1016/S0165-1781(02)00034-3

Hajcak, G., & Simons, R. F. (2008). Oops! . . . I did it again: An ERP and behavioral study of double-errors. *Brain and Cognition*, 68, 15–21. https://doi.org/10.1016/j.bandc.2008.02.118

Hillman, C. H., Pontifex, M. B., Castelli, D. M., Khan, N. A., Raine, L. B., Scudder, M. R., . . . Kamijo, K. (2014). Physical activity intervention improves cognitive and brain health in children. *Pediatrics*, 134, e1063–e1071. https://doi.org/10.1542/peds.2013-3219

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Hyman, C., Hofer, M., Barde, Y., Juhasz, M., Yancopoulos, G. D., Squinto, S. P., & Lindsay, R. M. (1991). BDNF is a neurotrophic factor for dopaminergic neurons of the substantia nigra. *Nature*, 350, 230–232. https://doi.org/10.1038/350230a0

Jocham, G., & Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neuroscience & Biobehavioral Reviews*, 33, 48–60. https://doi.org/10.1016/j.neubiorev.2008.08.011

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, 1046–1058. https://doi.org/10.1111/j.1467-7687.2011.01054.x

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test manual*. Circle Pines, MN: American Guidance Service.

Knaepen, K., Goekint, M., Heyman, E. M., & Meeusen, R. (2010). Neuroplasticity—Exercise-induced response of peripheral brain-derived neurotrophic factor. *Sports Medicine*, 40, 765–801.

Krafft, C. E., Pierce, J. E., Schwarz, N. F., Chi, L., Weinberger, A. L., Schaeffer, D. J., . . . McDowell, J. E. (2014). An eight month randomized controlled exercise intervention alters resting state synchrony in overweight children. *Neuroscience*, 256, 445–455. https://doi.org/10.1016/j.neuroscience.2013.09.052

Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., . . . Johnson, C. L. (2002). 2000 CDC growth charts for the United States: Methods and development. *Vital & Health Statistics*, 246, 1–190.

Kuppers, E., & Beyer, C. (2001). Dopamine regulates brain-derived neurotrophic factor (BDNF) expression in cultured embryonic mouse striatal cells. *NeuroReport*, 12, 1175–1179.

Ladouceur, C. D., Dahl, R. E., Birmaher, B., Axelson, D. A., & Ryan, N. D. (2006). Increased error-related negativity (ERN) in childhood anxiety disorders: ERP and source localization. *Journal of Child Psychology and Psychiatry*, 47, 1073–1082. https://doi.org/10.1111/j.1469-7610.2006.01654.x

Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: A theoretical review and critique of conflict-related ERPS. *International Journal of Psychophysiology*, *93*, 283–297. https://doi.org/10.1016/j.ijpsycho.2014.06.007

Lopez-Calderon, J., & Luck, S. J. (2010). *ERPLAB* (version 1.0.0.33a) (Computer software). UC-Davis Center for Mind & Brain. Retrieved from http://erpinfo.org/erplab/erplab-download

Luepker, R. V., Perry, C. L., McKinlay, S. M., Nader, P. R., Parcel, G. S., Stone, E. J., . . . Wu, M. (1996). Outcomes of a field trial to improve children's dietary patterns and physical activity. The Child and Adolescent Trial for Cardiovascular Health. CATCH collaborative group. *JAMA*, *275*, 768–776.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology: General*, *129*, 43–60. https://doi.org/10.1037/0096-3445.129.1.43

Maier, M. E., di Pellegrino, G., & Steinhauser, M. (2012). Enhanced error-related negativity on flanker errors: Error expectancy or error significance? *Psychophysiology*, *49*, 899–908. https://doi.org/10.1111/j.1469-8986.2012.01373.x

Marques, E., Vasconcelos, F., Rolo, M. R., Pereira, F. C., Silva, A. P., Macedo, T. R., & Ribeiro, C. F. (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. https://doi.org/10.1196/annals.1432.041

McDermott, J. M., Perez-Edgar, K., Henderson, H. A., Chronis-Tuscano, A., Pine, D. S., & Fox, N. A. (2009). A history of childhood behavioral inhibition and enhanced response monitoring in adolescence are linked to clinical anxiety. *Biological Psychiatry*, *65*, 445–448. https://doi.org/10.1016/j.biopsych.2008.10.043

Meyer, A., Weinberg, A., Klein, D., & Hajcak, G. (2012). The development of the error-related negativity (ERN) and its relationship with anxiety: Evidence from 8 to 13 year-olds. *Developmental Cognitive Neuroscience*, *2*, 152–161. https://doi.org/10.1016/j.dcn.2011.09.005

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*, 752–760. https://doi.org/10.1111/1469-8986.3850752

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects of response conflict and trial type frequency. *Cognitive, Affective & Behavioral Neuroscience*, *3*, 17–26.

Ogunwole, S. U., Drewery, M. P., & Rios-Vargas, M. (2012). *The population with a bachelor's degree or higher by race and Hispanic origin: 2006–2010*. United States Census Bureau.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathologh: toward an endophenotype. *Clinical Psychology Review*, *28*, 1343–1354. https://doi.org/10.1016/j.cpr.2008.07.003

Pontifex, M. B., Raine, L. B., Johnson, C. R., Chaddock, L., Voss, M. W., Cohen, N. J., . . . Hillman, C. H. (2011). Cardiorespiratory fitness and the flexible modulation of cognitive control in preadolescent children. *Journal of Cognitive Neuroscience*, *23*, 1332–1345. https://doi.org/10.1162/jocn.2010.21528

Pontifex, M. B., Scudder, M. R., Brown, M. L., O'Leary, K. C., Wu, C., Themanson, J. R., & Hillman, C. H. (2010). On the number of trials necessary for stabilization of error-related brain activity across the life span. *Psychophysiology*, *47*, 767–773. https://doi.org/10.1111/j.1469-8986.2010.00974.x

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*, 407–413. https://doi.org/10.1037/a0028795

Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2006). Error-related electrocortical responses are enhanced in children with obsessive-compulsive behaviors. *Developmental Neuropsychology*, *29*, 431–445. https://doi.org/10.1207/s15326942dn2903_3

Segalowitz, S. J., Santesso, D. L., Murphy, T. I., Homan, D., Chantziantoniou, D. K., & Khan, S. (2010). Retest reliability of medial frontal negativities during performance monitoring. *Psychophysiology*, *47*, 260–270. https://doi.org/10.1111/j.1469-8986.2009.00942.x

Tanner, J. M. (1962). *Growth at adolescence: With a general consideration of the effects of hereditary and environmental factors upon growth and maturation from birth to maturity*. Oxford, UK: Blackwell Scientific Publications.

Taylor, S. J., 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.

Themanson, J. R., & Hillman, C. H. (2006). Cardiorespiratory fitness and acute aerobic exercise effects on neuroelectric and behavioral measures of action monitoring. *Neuroscience*, *141*, 757–767. https://doi.org/10.1016/j.neuroscience.2006.04.004

Themanson, J. R., Pontifex, M. B., & Hillman, C. H. (2008). Fitness and action monitoring: Evidence for improved cognitive flexibility in young adults. *Neuroscience*, *157*, 319–328. https://doi.org/10.1016/j.neuroscience.2008.09.014

Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A., & Klein, D. N. (2012). Electrocortical and behavioral measures of response monitoring in young children during a go/no-go task. *Developmental Psychobiology*, *54*, 139–150. https://doi.org/10.1002/dev.20590

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

Van Bogaert, P., Wikler, D., Damhaut, P., Szliwowski, H. B., & Goldman, S. (1998). Regional changes in glucose metabolism during brain development from the age of 6 years. *NeuroImage*, *8*, 62–68. https://doi.org/10.1006/nimg.1998.0346

Van Meel, C. S., Heslenfeld, D. J., Rommelse, N. N., Oosterlaan, J., & Sergeant, J. A. (2012). Developmental trajectories of neural mechanisms supporting conflict and error processing in middle

childhood. *Developmental Neuropsychology*, *37*, 358–378. https://doi.org/10.1080/87565641.2011.653062

van Praag, H., Christie, B. R., Sejnowski, T. J., & Gage, F. H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Neurobiology*, *96*, 13427–13431. https://doi.org/10.1073/pnas.96.23.13427

Williams, S. N., & Undieh, A. S. (2009). Dopamine D1-like receptor activation induces brain-derived neurotrophic factor protein expression. *Neuro-Report*, *20*, 606–610. https://doi.org/10.1097/WNR.0b013e32832a0a98

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Rolling Meadows, IL: Riverside Publishing.

Yeung, N., Botvinick, M. W., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959. https://doi.org/10.1037/0033-295X.111.4.931