Supplemental Information

This appendix has been provided by the authors to give readers' additional information about the study.

PA PROGRAM DESCRIPTION

Goals of the FITKids Afterschool Program

The primary goal of the FITKids afterschool program was to increase cardiovascular fitness through participation in developmentally appropriate PA. Given the stage of development, secondary goals focused on experientially increasing motor skill competence (eg, dribbling a basketball) and social responsibility (eg, fair play, cooperation) within a PA setting.

Background

FITKids (http://clinicaltrials.gov/ct2/show/NCT01334359?term=FITKids&rank=1) was offered for 2 hours every day afterschool for ~90% of the school days on the annual academic calendar. Based on empirical research associated with the Coordinated Approach to Child Health,[31] the FITKids program was adapted from Community Access to Child Health (CATCH) to meet the specific contextual needs of local schoolchildren by aligning with the state learning standards. The first clinical trial of CATCH, sponsored by the National Institutes of Health, was in 1987 and now more than 25 years later has been implemented into over 9000 communities. This program was selected as the basis for FITKids given its effects on PA behaviors[32] and known sustainability.[33] Further, CATCH effectively reduces overweight among school-aged children.[34,35] FITKids is purposefully designed to exceed the national recommendation of 60 minutes of moderate to vigorous PA (MVPA) each day through engagement in integrated fitness activities and low organizational games.

Daily Lessons

Within a daily lesson, children experienced a consistent organizational structure of "instant activities," snack/educational component, and game play. Specifically, when children arrived at the program site, they immediately engage in instant activities, or a series of individual and partner tasks that are centered on particular health-related fitness component (eg, cardiovascular endurance, muscular strength). The program also consisted of the consumption of a healthy snack and the introduction of a themed educational component related to health promotion (ie, goal setting, self-management, nutrition). The activities were mostly aerobic in nature, but simultaneously provide opportunities to refine motor skills (eg, throwing, catching). On the weekends, the children were encouraged to continue their participation in PA with their family unit, and PA worksheets were used during school holidays to log continued engagement.

Instant Activities and Activities in an Instant

Using the CATCH Kids Club afterschool program and CATCH K-5 supplemental physical education curriculum materials (Flaghouse, Inc, Hasbrouck Heights, NJ), activity cards containing instructions were selected to provide age-appropriate physical activities as the children arrive at the program site. Daily, 4 to 6 activity cards targeting aerobic activities, muscular strength and endurance, or a movement concept such as rhythms or cooperative games were selected and offered for children in accordance with the FITKids curriculum. After completing an activity station, children were asked to move to another station, whereby they individually experienced a new station and new partner or working group. Other instant activities from local physical education programs were also integrated into the program so there was alignment with content experienced in formal physical education during the school day. Lessons early in the school year focused on instructing the children how to safely participate in the activities by using appropriate techniques. Once children become acquainted with how to correctly perform the instant activities, they then focused on achieving their personal goals such as a certain number of repetitions or amount of time spent in the target heart zone. The various fitness stations, performed as instant activities, lasted ~30 minutes each lesson.

Because children wore a pedometer, HR monitor, or activity monitor during each lesson (described later), the FITKids leaders knew if a child was not meeting his or her PA goal for the day. "Activities in an instant" are spontaneous, brief activities that were designed to increase a child's HR through engagement in brief activities that are of a moderate to vigorous intensity and are intermittent in nature (eg, jumping jacks, agilities, etc). To obtain 70 minutes of MVPA each lesson, activities in an instant were used as a form of active rest, as a way to reset attention, or to get children moving after a water break.

Snack and Educational Time

Each year, a FITKids curriculum of weekly themes was applied to the school calendar of the local school district. The educational component was either about a health-related fitness component (eg, strength), nutrition, or a benefit of PA. The educational component was brief and often presented as a health tip (eg, "eat five fruits and vegetables a day") followed by a learning task (eg, a worksheet that identified common fruits and vegetables). Formally, the children were taught the "Go," "Slow," and Whoa" strategy for selecting food for consumption based on how the food was prepared.

Weekends and School Holidays

On long weekends and extended school holidays (eg, spring break), children were provided with PA logs and sample activities that they could do at home. The children were sometimes given a pedometer to track the number of steps that they took over the break. These strategies were simply ways of extending and supporting PA in the home environment during prolonged absences from the program.

PA Levels

PA is defined as bodily movement that requires energy expenditure above normal physiological demands. Each day children participated in an average of 70 minutes of MVPA as measured by E600 Polar HR monitors (Polar Electro, Finland, and Accusplit Eagle 170 pedometers, San Jose, CA). In accordance with the American College of Sports Medicine guidelines, children were asked to participate between 55% and 80% of their maximal HR, which was measured during pretesting. When not wearing a HR monitor, children were asked to predict the number of steps that they might take during a game and then track their progress toward their prediction. At first, the predictions among the children were highly inaccurate. However, in time, their ability to predict the number of steps became more refined. Intensity, energy expenditure, and minutes within the target heart zone were regularly measured to individually and collectively track all program participants (Supplemental Figures 5 and 6). Using the Polar activity monitor, a watch that displays the number of minutes spent in MVPA as a bar moving across the bottom of the interface, allowed the child to learn to self-regulate his or her level of participation. Children knew by looking at their wrist whether they had accomplished their goals of 70+ minutes in MVPA.

Environment

The environment was positive, nonelimination, and minimally competitive. During play in low-organizational games (eg, tag, active rock, paper, scissors challenge, etc), the children were not eliminated from the activity when "tagged" by their peers; instead they were given a different role within the game, thus keeping all children active. With the exception of snack/educational time and water breaks, rest was active (eg, walking, dynamic stretching, etc).

FITKids Leaders and Training

The instructional leaders of FITKids were certified physical education teachers, with college students majoring in kinesiology serving as their assistants. The FITKids leaders participated in CATCH training, which was experienced as 1-day (booster) or 3-day (initial) workshops regarding how to use a variety of interactive teaching strategies. Further, the training addressed the developmental needs of children and how best to create a positive, minimally competitive environment. Annually, the FITKids staff participated in a 1-day training program that included strategies to successfully implement key components of the CATCH program, FITKids procedures, introduction of educational components, safety, and first aid and cardiopulmonary resuscitation (CPR) certification.

COGNITIVE TASKS AND NEUROELECTRIC ASSESSMENT

Flanker Task

To assess aspects of inhibitory control, neuroelectric and behavioral indices of performance were measured during a modified version of the Eriksen flanker task.[22] This task is conceptually simplistic in that it requires the discrimination of a centrally presented target stimulus amid lateral flanking stimuli. In this task, participants were required to make a left hand thumb press on a Neuroscan STIM system response pad (Compumedics, Charlotte, NC) when the target stimulus pointed left, and a right hand thumb press when the target stimulus pointed right. Thus, participants were instructed to respond as accurately as possible to the direction of a centrally presented target goldfish amid either congruous (the target faces the same direction) or incongruous (the target faces the opposite direction) flanking goldfish stimuli (Supplemental Figure 7).[23] After a block of 40 practice trials, participants completed 2 blocks of 75 trials presented with equiprobable congruency and directionality. The stimuli were 3-cm tall yellow fish, which were presented focally for 200 milliseconds on a blue background with a fixed intertrial interval of 1700 milliseconds.

Switch Task

To assess aspects of working memory and cognitive flexibility, neuroelectric and behavioral indices of performance were measured during a color-shape switch task.[24] This task requires participants to learn a set of response mappings arbitrarily assigned to a set of colors (blue and green) and shapes (circle and square), then use a rule-set cue (the direction of the characters arms) to flexibly shift visuospatial attention toward the correct feature set and execute the correct response mapping (Supplemental Figure 8). In this task, participants completed 2 blocks of 60 homogeneous trials (1 block of color only; 1 block of shape only) in which they attended to a centrally presented character. Participants were instructed to make a left hand thumb press on a Neuroscan STIM system response pad (Compumedics) when the character was blue (in the color condition) or a circle (in the shape condition) and a right hand thumb press when the character was green (in the color condition) or a square (in the shape condition). During the heterogeneous condition, participants performed both the color and shape tasks together with the specific task on each trial indicated by the direction of the character's arms (arms up: respond based on the shape of the character; arms down: respond based on the color of the character). After a block of 40 practice trials, participants completed 3 blocks of 50 heterogeneous trials with equiprobable task occurrence and response directionality. The stimuli were 5.5-cm tall and 9-cm wide characters presented focally for 3000 milliseconds on a black background (or until a response was given) with a fixed intertrial interval of 3500 milliseconds.

ERP Recording

A Neuroscan Synamps 2 amplifier (Compumedics) was used to acquire EEG activity in response to the cognitive tasks described earlier from 64 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 based on the International 10-10 system[36] by using a Neuroscan Quikcap (Compumedics). Recordings were referenced to averaged mastoids (M1, M2), with AFz serving as the ground electrode, and impedance less than 10 kOmega. Additional electrodes were placed above and below the left orbit and on the outer canthus of each eye to monitor electro-oculographic activity with a bipolar recording. Continuous 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. Continuous data were corrected offline for electro-oculographic artifacts by using a spatial filter.[37] Stimulus-locked epochs were created for correct trials from −100 to 1500 milliseconds around the stimulus, baseline corrected by using the −100 to 0 milliseconds prestimulus period, and filtered by using a zero phase shift low-pass filter at 30 Hz (24 dB/octave). Artifact in the EEG signal was identified if an amplitude excursion of $\pm 75\ \mu V$ occurred or if the overall variance of an epoch exceeded >1.5 SDs of the mean of all accepted epochs. Artifact-free trials that were accompanied by correct responses were averaged. The P3 component was evaluated as the mean amplitude within a 50-millisecond interval surrounding the largest positive going peak within a 300 to 700 milliseconds latency window.[38,39] Amplitude was measured as the difference between the mean prestimulus baseline and mean peak-interval amplitude; peak latency was defined as the time point corresponding to the maximum peak amplitude.

Aerobic Fitness Assessment

Cardiorespiratory fitness was assessed by using a test of $V_{O_{2peak}}$, which describes the physiological limit to the rate at which an individual can deliver/consume oxygen and is considered the criterion measure of cardiorespiratory fitness.[20] Oxygen consumption was measured by using a computerized indirect calorimetry system (ParvoMedics True Max 2400) while participants ran/walked on a motor-driven treadmill at a constant speed with incremental increases of 2.5% grade every 2 minutes until the participant was no longer able to maintain the exercise intensity.[20] A Polar HR monitor (Polar WearLink+ 31; Polar Electro) was used to measure HR throughout the test, and ratings of perceived exertion were assessed every 2 minutes by using the children's OMNI scale.[40] $V_{O_{2peak}}$ was based upon maximal effort as evidenced by (1) a peak HR > 185 beats per minute[20] and a HR plateau[41]; (2) RER > 1.0[42]; (3) a score on the children's OMNI ratings of perceived exertion scale >7[40]; and/or (4) a plateau in oxygen consumption corresponding to an increase of less than 2 mL/kg per minute despite an increase in workload.

STATISTICAL ANALYSIS

All statistical analyses were conducted with $\alpha = .05$ by using the Greenhouse-Geisser statistic with subsidiary univariate ANOVAs and Tukey's HSD tests for posthoc comparisons. Analysis of each task was conducted separately for each variable of interest in PASW Statistics, 19.0 by using multiple imputation for missing data points. Flanker task performance (response accuracy, RT) was assessed by using a 2 (group: intervention, wait-list) × 2 (time: pretest, posttest) × 2 (congruency: congruent, incongruent) multivariate repeated measures ANOVA. The P3 component in response to the flanker task was assessed separately for amplitude and latency by using a 2 (group: intervention, wait-list) × 2 (time: pretest, posttest) × 2 (congruency: congruent, incongruent) × 7 (electrode site: Fz, FCz, Cz, CPz, Pz, POz, Oz) multivariate repeated measures ANOVA. Switch task performance (response accuracy, RT) was assessed by using a 2 (group: intervention, wait-list) × 2 (time: pretest, posttest) × 2 (switch: homogeneous, heterogeneous) multivariate repeated measures ANOVA. The P3 component in response to the switch task was assessed separately for amplitude and latency by using a 2 (group: intervention, wait-list) × 2 (time: pretest, posttest) × 2 (switch: homogeneous, heterogeneous) × 7 (electrode site: Fz, FCz, Cz, CPz, Pz, POz, Oz) multivariate repeated measures ANOVA.

RESULTS

All significant effects are characterized in Supplemental Table 3 (for aerobic fitness and BMI), Supplemental Table 4 (for response accuracy and RT), and Supplemental Table 5 (for P3 amplitude and latency). Preliminary analyses were performed to test whether gender, race, or recruitment wave was related to any anthropomorphic, behavioral, or neuroelectric variables. Findings revealed no significant main effects or interactions (Ps $\geq$ .06), thus all further analyses were collapsed across gender, race, and recruitment wave.

Aerobic Fitness and Weight Status

Analyses revealed a main effect of Time with increased aerobic fitness assessed by using $\dot{V}O_{2peak}$ at posttest (39.7 mL/kg per minute, 95% CI: 38.8 to 40.7) relative to pretest (38.3 mL/kg per minute, 95% CI: 37.4 to 39.3; $P < .001$, $d = 0.45$). This effect was superseded by an interaction of Treatment $\times$ Time. Decomposition of the Treatment $\times$ Time interaction revealed increases in aerobic fitness for both the intervention group (2.1 mL/kg per minute, 95% CI: 1.2 to 2.9, $P < .001$, $d = 0.65$) and the wait-list control group (0.7 mL/kg per minute, 95% CI: 0.1 to 1.4, $P = .02$, $d = 0.32$). However, the intervention group demonstrated greater improvement in aerobic fitness than the wait-list control group (1.3 mL/kg per minute, 95% CI: 0.3 to 2.4, $P = .014$, $d = 0.34$). No differences in aerobic fitness between the intervention and wait-list groups were observed at pretest or posttest ($P \geq .07$).

Analysis of aerobic fitness percentile based on normative data provided by Shvartz and Reibold[21] revealed a main effect of Time with increased aerobic fitness percentile at posttest (23.0 percentile, 95% CI: 19.8 to 26.1) relative to pretest (19.7 percentile, 95% CI: 16.9 to 22.6, $P = .005$, $d = 0.32$). This effect was superseded by an interaction of Treatment $\times$ Time. Decomposition of the Treatment $\times$ Time interaction revealed increased aerobic fitness percentile only among intervention participants (5.6 percentile, 95% CI: 2.0 to 9.3, $P = .003$, $d = 0.43$) and not among wait-list participants (0.8 percentile, 95% CI: -1.8 to 3.4, $P = 0.53$, $d = 0.08$). Further, the intervention group demonstrated greater improvement in aerobic fitness percentile than the wait-list control group (4.8 mL/kg per minute, 95% CI: 0.4 to 9.2, $P = .034$, $d = 0.29$). No differences in aerobic fitness percentile between the intervention and wait-list groups were observed at pretest or posttest ($P \geq .13$).

Analysis of BMI revealed increased BMI at posttest (19.5, 95% CI: 18.9 to 20.2) relative to pretest (19.0, 95% CI: 18.4 to 19.5). This effect was superseded by an interaction of Treatment $\times$ Time. Decomposition of the Treatment $\times$ Time interaction revealed increases in BMI only among the wait-list control group (0.9, 95% CI: 0.6 to 1.2, $P < .001$, $d = 0.81$) and not among intervention participants (0.2, 95% CI: 0.0 to 0.5, $P = .06$, $d = 0.27$). Further, the wait-list group demonstrated a greater increase than the intervention group (0.7, 95% CI: 0.3 to 1.1, $P = .001$, $d = 0.47$). No differences in BMI between the intervention and wait-list groups were observed at pretest or posttest ($P \geq .44$).

Attentional Inhibition

Behavioral Findings

Analysis of the flanker task revealed increased response accuracy at posttest (82.5%, 95% CI: 81.2 to 83.9) relative to pretest (77.2%, 95% CI: 75.7 to 78.6, $P < .001$, $d = 1.1$). This effect was superseded by an interaction of Treatment $\times$ Time. Decomposition of the Treatment $\times$ Time interaction revealed increases in response accuracy for both the intervention group (7.0%, 95% CI: 4.6 to 9.4, $P < .001$, $d = 0.8$) and the wait-list control group (3.8%, 95% CI: 1.6 to 5.9, $P = .001$, $d = 0.45$). However, the intervention group demonstrated greater improvement than the wait-list control group (3.2%, 95% CI: 0.0 to 6.5, $P = .047$, $d = 0.27$). Differences in response accuracy between the intervention and the wait-list control group were only observed at pretest (-3.6%, 95% CI: -6.6 to -0.6, $P = .018$, $d = 0.32$). A main effect of congruency was also observed for response accuracy with superior performance on congruent (83.1%, 95% CI: 81.9 to 84.3) relative to incongruent trials (76.6%, 95% CI: 75.3 to 77.9, $P < .001$, $d = 1.55$). Analysis of RT in response to the flanker task revealed shorter latency at posttest (488.9 milliseconds, 95% CI: 476.4 to 501.3) relative to pretest (527.5 milliseconds, 95% CI: 512.0 to 543.1, $P < .001$, $d = 0.82$). Similarly, shorter RT was observed for congruent (492.9 milliseconds, 95% CI: 480.9 to 504.8) relative to incongruent trials (523.6 milliseconds, 95% CI: 510.9 to 536.3, $P < .001$, $d = 0.74$). No interactions with group assignment were observed for RT ($P \geq .18$).

Neuroelectric Findings

Analysis of P3 amplitude in response to the flanker task revealed increased P3 amplitude for incongruent (7.3 μV, 95% CI: 6.8 to 7.9) relative to congruent trials (7.0 μV, 95% CI: 6.5 to 7.5, $P = .009$, $d = 0.2$). This main effect of congruency was superseded by an interaction of Treatment $\times$ Time $\times$ Congruency. Decomposition of the Treatment $\times$ Time $\times$ Congruency interaction was conducted by examining Treatment $\times$ Time within the congruent and incongruent trials separately. Within the congruent trials, no main effects or interactions were observed ($P \geq .06$). Within the incongruent trials, an interaction of Treatment $\times$ Time was observed ($F_{1,219} = 5.2$, $P = .023$, $\eta^2 = 0.02$). Decomposition of the Treatment $\times$ Time interaction revealed that only the intervention group demonstrated increased P3 amplitude at posttest relative to pretest in response to incongruent trials (1.4 μV, 95% CI: 0.3 to 2.6, $P = .012$, $d = 0.34$), with greater change in P3 amplitude on incongruent trials observed from pretest to posttest relative to the wait-list group (1.9 μV, 95% CI: 0.3 to 3.5, $P = .023$, $d = 0.31$). No differences in P3 amplitude in response to the incongruent trials of the flanker task were observed between the intervention and wait-list groups at pretest or posttest ($P \geq .15$). A main effect of Electrode Site was also observed, which was superseded by an interaction of Congruency × Electrode Site. Decomposition of the Congruency × Electrode Site interaction revealed increased P3 amplitude for incongruent relative to congruent trials only at the FCz, Cz, and CPz electrode sites ($\geq$0.5 μV, 95% CI $\geq$0.1 to $\leq$1.3, Ps $\leq$ .015, ds $\geq$ 0.19).

Analysis of P3 latency in response to the flanker task revealed faster P3 latency for congruent (520.2 milliseconds, 95% CI: 512.4 to 528.1) relative to incongruent (529.7 milliseconds, 95% CI: 522.9 to 536.5, $P = .002$, $d = 0.38$) trials. This effect was superseded by an interaction of Treatment × Time × Congruency. Decomposition of the Treatment × Time × Congruency interaction was conducted by examining Treatment × Time within the congruent and incongruent trials separately. Within the congruent trials, no significant main effects or interactions were observed (Ps $\geq$ .66). Within the incongruent trials, an interaction of Treatment × Time was observed ($F_{1,219} = 6.3$, $P = .013$, $\eta^2 = 0.03$). Decomposition of the Treatment × Time interaction revealed that only the intervention group demonstrated faster P3 latency at posttest relative to pretest in response to incongruent trials (20.1 milliseconds, 95% CI: 2.6 to 37.6, $P = .025$, $d = 0.31$), with a greater change in P3 latency from pretest to posttest relative to the wait-list group (32.0 milliseconds, 95% CI: 6.9 to 57.2, $P = .013$, $d = 0.34$). At pretest, faster P3 latency in response to the incongruent trials of the flanker task was also exhibited by the wait-list control group (521.5 milliseconds, 95% CI: 508.0 to 535.0) relative to the intervention group (541.9 milliseconds, 95% CI: 529.9 to 554.0, $P = .027$, $d = 0.3$). No differences in P3 latency in response to the incongruent trials of the flanker task were observed between the intervention and wait-list groups at posttest ($P > .2$). A main effect of Electrode Site was also observed, which was superseded by an interaction of Time × Electrode Site. Decomposition of the Time × Electrode Site interaction revealed faster P3 latency for posttest relative to pretest only at the FCz electrode site (28.3 milliseconds, 95% CI: 8.6 to 48.1, $P = .005$, $d = 0.56$).

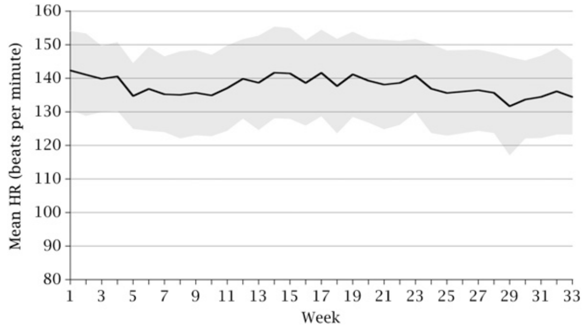
Cognitive Flexibility

Behavioral Findings

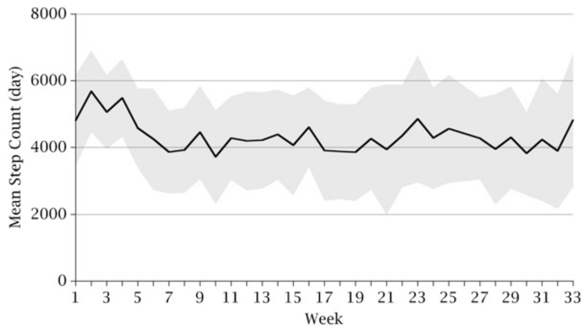
Analysis of the switch task revealed increased response accuracy at posttest (82.9%, 95% CI: 81.8 to 84.0) relative to pretest (77.6%, 95% CI: 76.4 to 78.8, $P < .001$, $d = 1.4$) and greater response accuracy for homogeneous trials (88.9%, 95% CI: 88.2 to 89.6) relative to heterogeneous trials (71.6%, 95% CI: 70.2 to 73.1, $P < .001$, $d = 4.7$). These effects were superseded by interactions of Time × Cost and Treatment × Time × Cost. Decomposition of the Treatment × Time × Cost interaction was conducted by examining Treatment × Time within the homogenous and heterogeneous trials separately. Within the homogeneous trials, only a main effect of Time was observed with increased response accuracy at posttest (90.0%, 95% CI: 89.2 to 90.8) relative to pretest (87.7%, 95% CI: 86.8 to 88.7, $F_{1,219} = 20.6$, $P < .001$, $d = 0.79$). Within the heterogeneous trials, increased response accuracy was observed at posttest (75.8%, 95% CI: 74.1 to 77.4) relative to pretest (67.5%, 95% CI: 65.7 to 69.2, $F_{1,219} = 79.8$, $P < .001$, $d = 1.43$). This effect was superseded by an interaction of Treatment × Time ($F_{1,219} = 6.5$, $P = .011$, $\eta^2 = 0.03$). Decomposition of the Treatment × Time interaction revealed increases in response accuracy for both the intervention group (10.7%, 95% CI: 8.2 to 13.3, $P < .001$, $d = 1.14$) and the wait-list control group (6.0%, 95% CI: 3.3 to 8.6, $P < .001$, $d = 0.59$). However, the intervention group demonstrated greater improvement than the wait-list control group (4.8%, 95% CI: 1.1 to 8.4, $P = .011$, $d = 0.35$). No differences in response accuracy between the intervention and wait-list groups were observed at pretest or posttest ($P \geq .09$). Analysis of RT in response to the switch task revealed shorter RT at posttest (1089.2 milliseconds, 95% CI: 1065.6 to 1112.7) relative to pretest (1117.5 milliseconds, 95% CI: 1092.9 to 1142.0, $P = .04$, $d = 0.2$). Shorter RT was also observed for homogeneous trials (780.9 milliseconds, 95% CI: 762.1 to 799.7) relative to heterogeneous trials (1425.8 milliseconds, 95% CI: 1398.3 to 1453.2, $P < .001$, $d = 4.65$).

Neuroelectric Findings

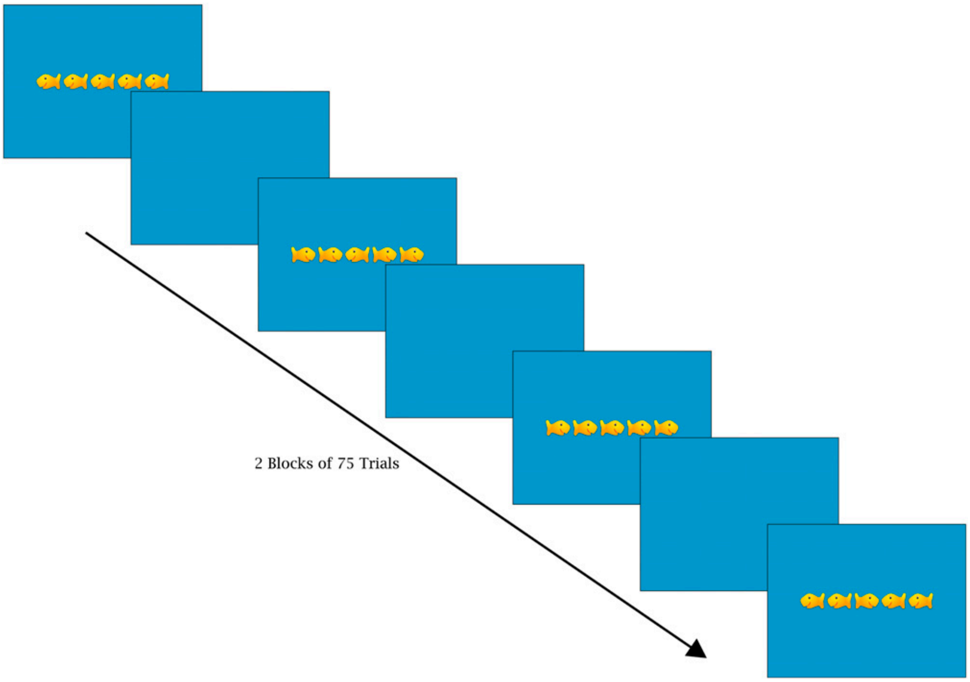
Analysis of P3 amplitude in response to the switch task revealed increased P3 amplitude at posttest (7.5 μV, 95% CI: 7.0 to 8.0) relative to pretest (6.8 μV, 95% CI: 6.3 to 7.2, $P = .004$, $d = 0.47$) and increased P3 amplitude for homogeneous (7.9 μV, 95% CI: 7.4 to 8.3) relative to heterogeneous trials (6.4 μV, 95% CI: 6.0 to 6.9, $P < .001$, $d = 0.93$). These effects were superseded by an interaction of Treatment × Time × Cost. Decomposition of the Treatment × Time × Cost interaction was conducted by examining Treatment × Time within the homogenous and heterogeneous trials separately. Within the homogeneous trials, no significant main effects or interactions were observed ($P \geq .06$). Within the heterogeneous trials, increased P3 amplitude was observed at posttest (6.9 μV, 95% CI: 6.3 to 7.4) relative to pretest (6.0 μV, 95% CI: 5.4 to 6.6, $F_{1,219} = 6.7$, $P = .01$, $d = 0.45$). This effect was superseded by an interaction of Treatment × Time ($F_{1,219} = 4.1$, $P = .04$, $\eta^2 = 0.02$). Decomposition of the Treatment × Time interaction revealed that only the intervention group demonstrated increased P3 amplitude at posttest relative to pretest in response to heterogeneous trials (1.5 μV, 95% CI: 0.6 to 2.5, $P = .002$, $d = 0.43$), with a greater change in P3 amplitude from pretest to posttest relative to the wait-list group (1.4 μV, 95% CI: 0.0 to 2.7, $P = .04$, $d = 0.27$). No differences in P3 amplitude in response to the heterogeneous trials of the switch task were observed between the intervention and wait-list groups at pretest or posttest ($P \geq .2$). A main effect of Electrode Site was also observed, which was superseded by an interaction of Cost × Electrode Site. Decomposition of the Cost × Electrode Site interaction revealed increased P3 amplitude for homogeneous relative to heterogeneous trials at the Cz, CPz, Pz, POz, and Oz electrode sites ($\geq$1.7 μV, 95% CI $\geq$0.7 to $\leq$2.8, $Ps \leq .001$, $ds \geq 0.52$). Analysis of P3 latency in response to the switch task revealed main effects of Cost and Electrode Site, which were superseded by an interaction of Cost × Electrode Site. Decomposition of the Cost × Electrode Site interaction revealed faster P3 latency for homogeneous relative to heterogeneous trials at the Cz, CPz, Pz, and POz electrode sites ($\geq$17.2 milliseconds, 95% CI $\geq$3.1 to $\leq$65.0, $Ps \leq .017$, $ds \geq 0.23$) and faster P3 latency for heterogeneous relative to homogeneous trials at the Fz electrode site (31.5 milliseconds, 95% CI: 10.9 to 52.2, $P = .003$, $d = 0.37$).



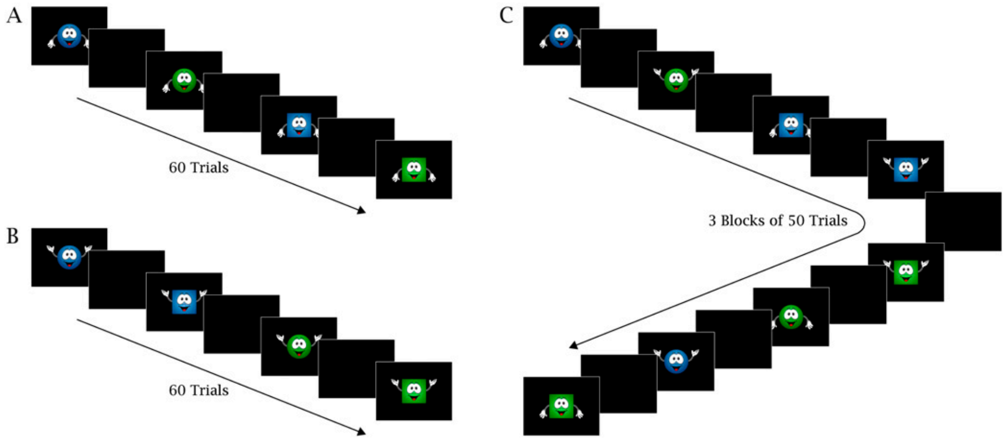
SUPPLEMENTAL FIGURE 5
Mean (SE) HR assessed during the afterschool program over the course of the intervention.



SUPPLEMENTAL FIGURE 6
Mean (SE) number of steps taken assessed by using a pedometer during the afterschool program over the course of the intervention.

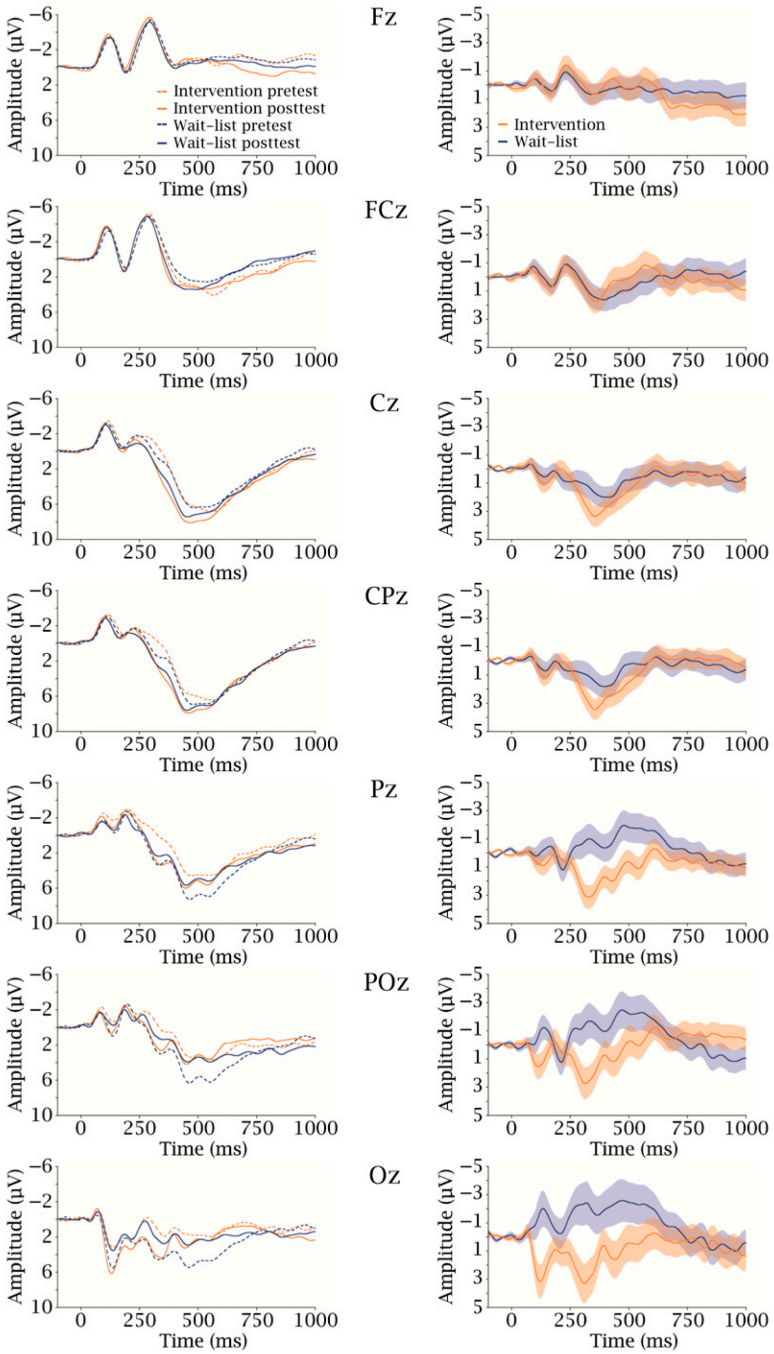


SUPPLEMENTAL FIGURE 7
Illustration of the modified Eriksen flanker task with fish.



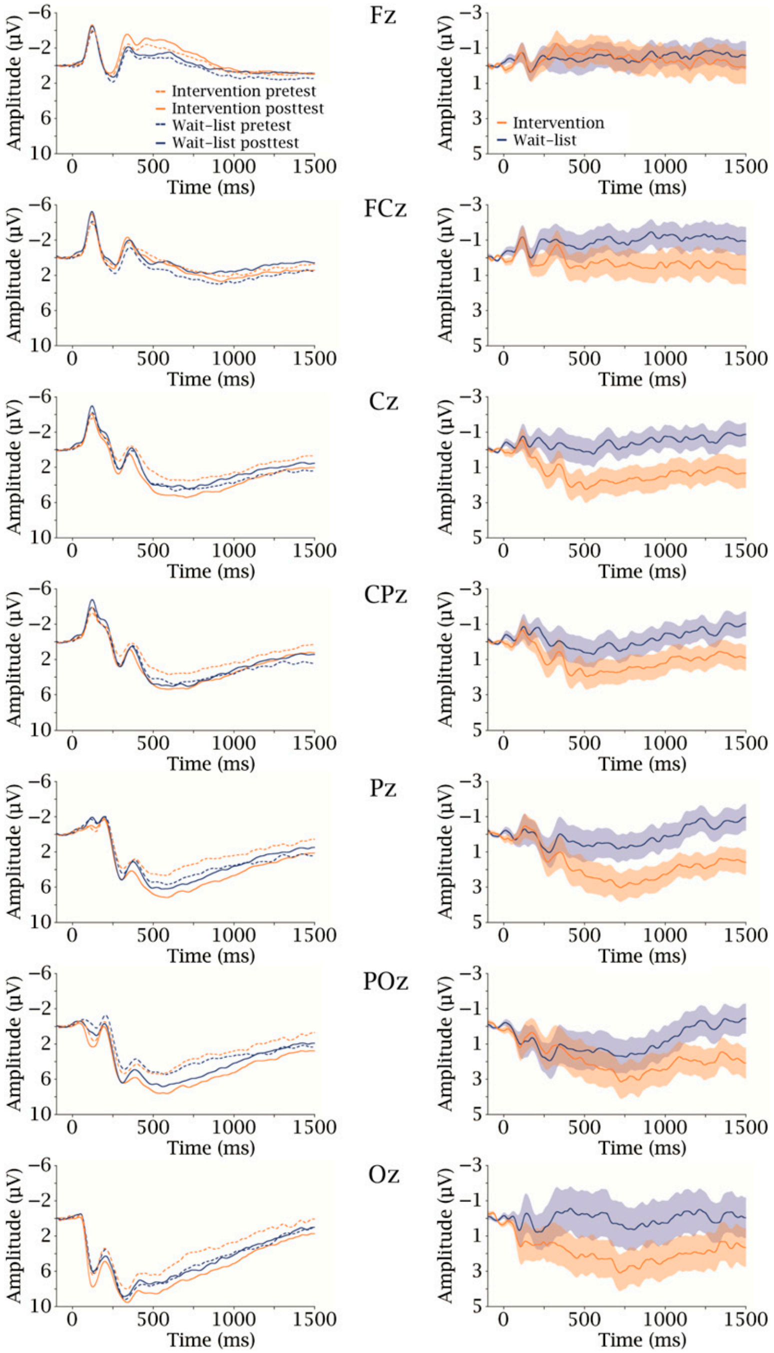
SUPPLEMENTAL FIGURE 8

Illustration of the color-shape switch task. Participants completed 2 blocks of the homogeneous condition (A, one block responding to the color; B, one block responding to the shape), then completed 3 blocks of the heterogeneous condition (C, arms down indicates respond to color, arms up indicates respond to shape).



SUPPLEMENTAL FIGURE 9

Grand average event-related potential waveforms (left) and difference waves (±SE) of the change from pre- to posttest (right) to the incongruent trials of the flanker task as a function of group and electrode site.



SUPPLEMENTAL FIGURE 10

Grand average event-related potential waveforms (left) and difference waves (±SE) of the change from pre- to posttest (right) to the heterogeneous trials of the switch task as a function of group and electrode site.

SUPPLEMENTAL TABLE 3 Statistical Summary Table for Anthropomorphic Effects

Measure and Effect	df	*F*	*P*	η^2
Aerobic fitness (Vo_{2peak} mL/kg/min)				
Time	1, 219	27.9	<.001	0.11
Treatment × Time	1, 219	6.2	.014	0.03
Aerobic fitness percentile				
Time	1, 219	8.2	.005	0.04
Treatment × Time	1, 219	4.6	.034	0.02
BMI				
Time	1, 219	33.7	<.001	0.13
Treatment × Time	1, 219	12.3	.001	0.05

Only significant ($P < .05$) effects are reported.

SUPPLEMENTAL TABLE 4 Statistical Summary Table for Task Performance Effects

Measure and Effect	df	*F*	*P*	η^2
Flanker task response accuracy				
Time	1, 219	43.9	<.001	0.17
Congruency	1, 219	330.4	<.001	0.60
Treatment × Time	1, 219	3.9	.05	0.02
Flanker task RT				
Time	1, 219	29.9	<.001	0.12
Congruency	1, 219	388.4	<.001	0.64
Switch task response accuracy				
Time	1, 219	80.1	<.001	0.27
Cost	1, 219	743.2	<.001	0.77
Time × Cost	1, 219	42.7	<.001	0.17
Treatment × Time × Cost	1, 219	7.3	.007	0.03
Switch task RT				
Time	1, 219	4.3	.04	0.02
Cost	1, 219	2568.8	<.001	0.92

Only significant ($P < .05$) effects are reported.

SUPPLEMENTAL TABLE 5 Statistical Summary Table for Neuroelectric Effects

Measure and Effect	df	*F*	*P*	η^2
Flanker task P3 amplitude				
Congruency	1, 219	6.9	.009	0.03
Electrode site	6, 219	56.3	<.001	0.20
Congruency × Electrode Site	6, 219	7.6	<.001	0.03
Treatment × Time × Congruency	1, 219	4.5	.036	0.02
Flanker task P3 latency				
Congruency	1, 219	10.0	.002	0.04
Electrode site	6, 219	6.3	.001	0.03
Time × Electrode Site	6, 219	2.5	.045	0.01
Treatment × Time × Congruency	1, 219	6.4	.012	0.03
Switch task P3 amplitude				
Time	1, 219	8.4	.004	0.04
Cost	1, 219	38.3	<.001	0.15
Electrode site	6, 219	210.5	<.001	0.49
Cost × Electrode Site	6, 219	6.8	<.001	0.03
Treatment × Time × Cost	1, 219	6.0	.015	0.03
Switch task P3 latency				
Cost	1, 219	9.3	.003	0.04
Electrode site	6, 219	121.8	<.001	0.36
Cost × Electrode Site	6, 219	11.6	<.001	0.05

Only significant ($P < .05$) effects are reported.