Aerobic exercise modulates transfer and brain signal complexity following cognitive training

Chun-Hao Wang\textsuperscript{a}, David Moreau\textsuperscript{b}, Cheng-Ta Yang\textsuperscript{c,e}, Yun-Yen Tsai\textsuperscript{a}, Jui-Tang Lin\textsuperscript{a}, Wei-Kuang Liang\textsuperscript{d,*}, Chia-Liang Tsai\textsuperscript{a,\textdegree}\textdegree

\textsuperscript{a} Institute of Physical Education, Health & Leisure Studies, National Cheng Kung University, No. 1, University Road, Tainan City, Taiwan
\textsuperscript{b} School of Psychology and Centre for Brain Research, University of Auckland, Auckland, New Zealand
\textsuperscript{c} Department of Psychology, National Cheng Kung University, Social Sciences Building, No. 1, University Road, East District, Tainan City 701, Taiwan
\textsuperscript{d} Institute of Cognitive Neuroscience, National Central University, Jhongli 320, Taiwan
\textsuperscript{e} Institute of Allied Health Sciences, National Cheng Kung University, No.1, University Road, Tainan City, Tainan

A R T I C L E   I N F O

Keywords:
Physical exercise
Dual n-back training
Working memory
Transfer effect
Brain signal complexity

A B S T R A C T

Although recent evidence has demonstrated the potent effect of physical exercise to increase the efficacy of cognitive training, the neural mechanisms underlying this causal relationship remain unclear. Here, we used multiscale entropy (MSE) of electroencephalography (EEG)—a measure of brain signal complexity—to address this issue. Young males were randomly assigned to either a 20-day dual n-back training following aerobic exercise or the same training regimen following a reading. A feature binding working memory task with concurrent EEG recording was used to test for transfer effects. Although results revealed weak-to-moderate evidence for exercise-induced facilitation on cognitive training, the combination of cognitive training with exercise resulted in greater transfer gains on conditions involving greater attentional demanding, together with greater increases in cognitive modulation on MSE, compared with the reading condition. Overall, our findings suggest that the addition of antecedent physical exercise to brain training regimen could enable wider, more robust improvements.

1. Introduction

Cognitive enhancement has inspired considerable interest among the scientific community and the public at large in the past decade (Hillman, Erickson, & Kramer, 2008; Klingberg, 2010; Moreau, 2014; Moreau, Wang, Tseng, & Juan, 2015). One of the key reasons for this growing interest is the potential to improve higher-order cognitive functions and facilitate everyday activities, with important ramifications to academic and professional settings (Ramidis et al., 2014; Ward et al., 2017).

A number of promising behavioral approaches to cognitive enhancement have emerged over the years. These include computerized working memory training (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Klingberg, 2010), videogame training (Anguera et al., 2013), brain stimulation (Juan & Muggleton, 2012; Juan, Tseng, & Hsu, 2017), neural feedback (Wang & Hsieh, 2013), mindfulness (Jaiswal, Tsai, Juan, Liang, & Muggleton, 2018; Zeidan, Johnson, Diamond, David, & Goolkasian, 2010), and chronic (Ludyga, Gerber, Kamijo, Brand, & Pühse, 2018; Moreau, Kirk, & Walddie, 2017; Moreau, Morrison, & Conway, 2015; Rogge et al., 2017) or acute (Chang et al., 2017; Kao, Westfall, Soneson, Gurd, & Hillman, 2017; Ludyga, Pühse, Lucchi, Marti, & Gerber, 2019; Tsai et al., 2018a) participation in physical exercise. Of these, working memory training has gained popularity in recent years because of its potential utility as a cost-effective method to improve cognition across the lifespan (Guye, De Simoni, & von Bastian, 2017; Jaeggi, Buschkuehl, Jonides, & Shah, 2011). Working memory is considered a key component for a number of higher-order cognitive abilities, such as problem-solving, reasoning, executive control and decision-making (Diamond, 2013; Engle, 2002). A number of research groups have shown interest in studying the generalizability of working memory training, with equivocal results (Jaeggi et al., 2011; Kundu, Sutterer, Emrich, & Postle, 2013; Melby-Lervåg & Hulme, 2013).

Received 26 November 2018; Received in revised form 21 February 2019; Accepted 22 March 2019
Available online 31 March 2019
0301-0511/ © 2019 Elsevier B.V. All rights reserved.
Although transfer still remains contentious (Au et al., 2015; Melby-Lervåg & Hulme, 2013; Moreau & Conway, 2014; Redick et al., 2013; Soberi, Antfolk, Karlsson, Salo, & Laine, 2017), oftentimes because training is restricted to domain-specific abilities (Green & Bavelier, 2008; Moreau & Conway, 2014), the potential benefits and applications are still very appealing, and in this regard warrant further investigations (Sternberg, 2008).

Several research groups have become interested in the extent to which working memory training can be supplemented by additional interventions, to maximize its benefits. Indeed, there is growing evidence that multimodal training (e.g., combining cognitive training with other approaches) can produce greater cognitive gains than unimodal training (Bamidis et al., 2014; Heisz et al., 2017; Oswald, Gunzelmann, Rupprecht, & Hagen, 2006; Ward et al., 2017). One promising approach to maximize the effect of cognitive training is through a combination with physical exercise, which has been widely demonstrated to have a positive influence on working memory performance across lifespan (Erickson et al., 2011; Kamijo et al., 2011; Ludgya et al., 2018; Moreau et al., 2017; Stroth, Hille, Spitzer, & Reinhardt, 2009; Wang & Tsai, 2016). For example, neuroelectric evidence, albeit limited, suggests that physical exercise has the capacity to improve working memory performance along with benefits for task preparation evidenced by the early component of contingent negative variation (CNV) (Kamijo et al., 2011; Ludgya et al., 2018). Indeed, animal studies have revealed the potential mechanisms underlying such neuroplasticity by showing that physical exercise—more specifically, aerobic regimens—can upregulate expression of long-term potentiation, neurotrophins and growth factors (for example, brain-derived neurotrophic factor, BDNF) (Cotman, Berchtold, & Christie, 2007; O’Callaghan, Olhe, & Kelly, 2007; Vaynman, Ying, & Gomez-Pinilla, 2004). This cascade of effects typically leads to enhanced hippocampal neurogenesis (Raichlen & Alexander, 2017; Van Praag, 2008), in turn resulting in increased learning-related plasticity (Cassilhas, Tullik, & de Mello, 2016; Höttin & Röder, 2013). In agreement with this idea, after a 4-month intervention, Ward et al. (2017) observed improved learning on a battery of tasks (visuo-spatial reasoning, mental planning, and working memory updating) when supplementing 20 computerized cognitive training sessions with 28 physical exercise training sessions, in comparison with the former only. This suggests that physical exercise may have a facilitating effect on cognitive training interventions.

Furthermore, and given the positive effects of physical exercise on brain plasticity (Cotman & Berchtold, 2002), it also seems plausible that physical exercise can facilitate transfer of the training gains. Three recent studies have examined transfer effects in experiments combining cognitive and physical training in healthy adults (Daugherty et al., 2017; Heisz et al., 2017; Ward et al., 2017). For example, Heisz et al. (2017), comparing exercise training (high intensity interval training, HIIT) alone against a combination of HIIT exercise and cognitive training among healthy adults, found that individuals sensitive to exercise training (i.e., greater exercise-induced increases in fitness and serum BDNF) in the combined group exhibited greater performance improvements than exercise training alone. Similarly, Daugherty et al. (2017) demonstrated that a combination of physical exercise and cognitive training resulted in greater improvements on visuo-spatial reasoning than an active control condition based on an adaptive visual attention task. Together, these findings are consistent with the claim that a combination of physical exercise and cognitive demands has the capacity to maximize training benefits on cognition (Bamidis et al., 2014; Moreau & Conway, 2014; Moreau, Morrison, et al., 2015; Oswald et al., 2006; Wang, Yang, Moreau, & Muggleton, 2017). However, despite greater training effects observed after cognitive training and physical exercise (i.e., consisted of high intensity cardiovascular and resistance training) combined, relative to cognitive training alone, Ward et al. (2017) found no positive transfer to any of the cognitive ability measures they included.

The discrepancies across studies may in part be due to differences in experimental designs and training regimens. For example, comparing multimodal studies with exercise control groups may obscure some of the effects on training gains and transferability (Daugherty et al., 2017; Heisz et al., 2017), with the evidence showing no positive transfer of training gains when compared with cognitive training controls (Ward et al., 2017). It thus remains unclear whether a combination of exercise and cognitive training is superior to cognitive training alone, in terms of transfer. Here, we sought to replicate and extend upon findings from earlier multimodal intervention studies, in a population of healthy young adults. Instead of using an exercise training control group, however, we compared performance differences between two cognitive training groups, but with different additional interventions (exercise vs. reading). Unlike previous work in which the exercise and cognitive training were completely separate (Daugherty et al., 2017; Ward et al., 2017), participants in the exercise condition engaged with the cognitive training regimen immediately after acute bouts of moderate intensity aerobic exercise. This was motivated by the meta-analyses demonstrating positive changes in cognitive performance following acute exercise (Chang, Labban, Gapin, & Etnier, 2012; Ludgya, Gerber, Brand, Holsboer-Trachsler, & Pühs, 2016). In addition, empirical evidence shows significant acute exercise-induced increases in levels of plasticity-related biomarkers (Skriver et al., 2014), which presumably could improve the effectiveness of cognitive training and transfer. To our knowledge, there has never been a study examining the efficacy of cognitive training preceded by physical exercise. In the control condition, participants were assigned academic readings before receiving the cognitive training regimen, to control for placebo effects or other potential confounding factors (Moreau, Kirk, & Walde, 2016).

We chose an adaptive dual n-back task, in which participants were instructed to indicate whether the color or shape feature of a stimulus was the same as the one presented n-items before, as the cognitive training regimen for the following reasons: (1) n-back tasks have shown transfer effects in several studies (Jaeggi et al., 2008; Soberi et al., 2017); (2) one recent study, using latent variable analysis, has shown that the n-back task is a valid indicator of working memory, and correlates strongly with other working memory tasks (Schmiedek, Lövdén, & Lindenberger, 2014), but see (Redick & Lindsey, 2013); and (3) a single bout of moderate aerobic exercise has been associated with improved n-back performance (Weng, Pierce, Darling, & Voss, 2015) and enhanced modulations of brain activations (Li et al., 2014) during higher n-back levels (i.e., 2-back condition). For all these reasons, we postulated a facilitating effect of aerobic exercise for the dual n-back training.

Regarding transfer, we chose a shape-color feature-binding task to examine whether n-back training can lead to improvements in an untrained working memory task, structurally different from the training task (i.e., near transfer)—evidence for working memory improvements at the construct level. Because the transfer effect of multimodal training might be task-dependent, with findings showing greater benefits on working memory tasks with higher interference (Heisz et al., 2017), our shape-color feature binding paradigm included three working memory conditions with different attentional demands (shape or color feature only vs. shape-color binding feature). This particular design allowed examining whether any transfer effect at the construct level was general or selective.

Beyond previous behavioral results (Daugherty et al., 2017; Heisz et al., 2017; Ward et al., 2017), there is scant evidence on the changes in functional brain activity that accompany improvements in performance associated with multimodal interventions that include physical exercise. Given that the adaptability of neural systems to a changing environment has been shown in a number of studies employing physical exercise (Cotman & Berchtold, 2002; Knaepen, Goekint, Heyman, & Meeusen, 2010) or cognitive training interventions (Klingberg, 2010; Olesen, Westerberg, & Klingberg, 2003), this study aimed to examine whether the training-induced changes in cognitive transfer is linked to differential neural adaptability. To better quantify the neural
adaptability, we calculated the multiscale entropy analysis (MSE) (Costa, Goldberger, & Peng, 2002; Costa, Goldberger, & Peng, 2005; Garrett, McIntosh, & Grady, 2011; Liang et al., 2014) of the functional brain signals acquired during the transfer task. MSE, in short, is a fractal scaling measure of physiological complexity that calculates sample entropy at multiple time scales (i.e., coarse-grained time series) (Costa et al., 2002; Costa et al., 2005), which has been suggested to reflect the adaptability of biological systems (Costa et al., 2005). A smaller (or fine-grained) time scale indicates that the non-overlapping window is shorter, whereas a larger (or coarse-grained) time scale suggests that the non-overlapping window is longer. Specifically, greater sample entropy at smaller scales is thought to reflect more local network processing, whereas larger scales is considered to represent more distributed network processing (Grundy, Anderson, & Bialystok, 2017; Vakorin, Lippé, & McIntosh, 2011). Moreover, because sample entropy indexes the occurrence of repetitive patterns of a given time-series (Richman & Moorman, 2000), higher MSE values thus imply that the signal is less predictable and information rich, whereas lower ones indicate that the time series is more regular and less complex (Costa et al., 2005; Deco, Jirsa, & McIntosh, 2011; Yang et al., 2013).

In recent years, MSE has been successfully applied to measure brain signal complexity in electroencephalographic (EEG) signals for probing the brain mechanisms of cognition and behavior (Catarino, Churches, Baron-Cohen, Andrade, & Ring, 2011; Grundy et al., 2017; Heisz, Shedden, & McIntosh, 2012; Liang et al., 2014; Wang et al., 2014). For example, Liang et al. (2014), using a stop-signal task, found that the MSE values of EEG in successful stop trials was higher relative to that in failed stop trials, suggesting a greater ability to adapt to environmental changes during successful stopping (i.e., an irregularly-intervening sudden “stop” signal). Moreover, Grundy et al. (2017) observed greater MSE during task switching in bilinguals relative to that in monolinguals. Because greater MSE has been linked to brain’s ability to switch between functional states more readily (Beharelle, Kovacevic, McIntosh, & Levine, 2012; Grundy et al., 2017), their finding thus suggests that second-language experience modulates the adaptability of neural network during switching between languages. In terms of learning behavior, Heisz et al. (2012) found that brain signal complexity measured by MSE was higher for familiar faces and increased as the familiarity of faces became higher, suggesting the more meaningful information becomes available about a given task after learning, the higher brain complexity is. With regard to the relationship between physical exercise and cognitive functioning, Wang et al. (2014) found that elderly adults with higher level of physical activity exhibited greater MSE values along with higher response accuracy when performing a visuo-spatial working memory task, suggesting a positive association between physical activity level and neural adaptability supporting working memory processes. Given that MSE is sensitive to both cognitive learning and exercise effects, it could represent a promising approach to reveal the training-induced changes in neural adaptability.

We made a number of predictions concerning the present study. First, and given the positive effects of moderate acute exercise on n-back performance in prior work (Li et al., 2014; Weng et al., 2015), we hypothesized that the antecedent acute exercise may induce greater training improvements on the n-back task as compared to the control condition. In addition, based on findings showing that interventions involving both cognitive and physical demands can induce greater cognitive benefits (Heisz et al., 2017; Moreau, Morrison, et al., 2015), we hypothesized that a greater transfer effect would be observed for individuals engaging in the intervention that combined cognitive and physical demands, and that such effect would be greater for the condition involving higher attentional demands (i.e., shape-color binding condition). Finally, given that both cognitive learning and physical exercise have been demonstrated to have positive relationship with neural adaptability (Heisz et al., 2012; Wang et al., 2014), we hypothesized that individuals in the experimental condition would exhibit greater changes in MSE values than those whose regimen only includes cognitive demands.

2. Materials and methods

2.1. Participants

A total of 38 male participants were recruited via online public advertisement. Prerequisites included the following: (a) age over 20 years old, (b) body mass index (BMI) < 27 kg/m² (defined as mid obesity by the Ministry of Health and Welfare, Taiwan) (c) overall energy expenditure achieved a minimal of at least 600 metabolic equivalents per week (MET-min/week). Of the 38 participants enrolled, eight noncompliant subjects were dropped from this study because they were unable to follow the required training protocol. As a result, a total of 30 participants (exercise: n = 15; control: n = 15) completed all procedures. The sample size for each group was similar to those of prior behavioral or neurophysiological studies investigating training and transfer effect in healthy adults (Jaeggi et al., 2008; Kundu et al., 2013). All the participants were right-handed and had normal or corrected-to-normal vision. No individuals reported a history of neurological problems or cardiovascular diseases, nor were any taking medication known to affect cognitive function. Informed consent was obtained prior to the study, which was approved by the Human Research Ethics Committee (HREC) at National Cheng Kung University. In accordance with regulations of the HREC, the data acquired in the present study cannot be shared with researcher without the written re-consent of the participants.

2.2. Procedures

All participants took part in the experimental procedure that consisted of four phases carried out on separate occasions. In the first phase, the research assistants explained the experimental procedure, and asked participants to complete a consent form, as well as the International Physical Activity Questionnaire (IPAQ) and a demographic questionnaire. Participants’ heights and weights were measured to calculate BMIs. Participants were then instructed to practice and perform the training task to determine baseline n-back levels. Finally, they were fitted with a Polar heart rate (HR) monitor (RX800XC, Finland) and took the continuous graded maximal exercise test to measure baseline VO2max. After finishing the first phase, participants

Exercise group

Aerobic exercise Break Cognitive training

HR recovery 60 – 70% HR reserve Dual n-back WM task

Control group

Reading books Break Cognitive training

Dual n-back WM task

Fig. 1. Training protocol for the exercise and the control groups.
were randomly assigned to either the exercise and computer-based cognitive training (exercise group) or the reading and computer-based cognitive training (control group) training regimen (Fig. 1).

The EEG pretest session was conducted in the second phase, on a different day from the first phase to avoid potential acute exercise effects (Tsai et al., 2018a). Each participant was asked to sit comfortably in a dimly lit and soundproof room in front of a screen positioned at eye level, at a distance of approximately 100 cm. An electrocap and electrooculographic (EOG) were attached to the participant’s scalp and face before performing the transfer task. After 12 practice trials for each condition, participants started the task, while EEG signal was simultaneously recorded.

Both groups started training in the third phase, which was scheduled on a different day from the second phase. The training task (i.e., the dual n-back task) was adopted from Jaeggi et al. (2008). The cognitive training regimen lasted 5 weeks, with a 4-day/week schedule. This translated into 20 training sessions, each ranging from 30 to 40 min. Prior to performing the training task, the two groups received different interventions. Participants in the exercise group performed a 30-min running bout on a treadmill. The session consisted of a 5-min warm up, 20 min of aerobic exercise at moderate intensity (i.e., 60–70% heart rate reserve), and a 5-min cool down (Chang et al., 2017). Once participants’ heart rate had returned to within 10% of pre-exercise levels (approximately 15–20 min after a single bout of acute exercise) (Tsai et al., 2018a), they then performed the training task. One meta-analysis study has shown that such a short delay period results in greater cognitive benefits (Chang et al., 2012). Participants in the control group were assigned academic readings for the same duration and frequency as the exercise group, to account for general or non-specific treatment effects (Moreau et al., 2016).

In the fourth phase (the post-training session), participants were required to complete a VO2max test and perform the transfer task along with EEG recording. Both tests were completed within five days of training completion, and each test was administrated on different days to avoid acute exercise effects. All the procedures were identical to the pretest session.

2.3. Measures

2.3.1. Cardiorespiratory fitness test

We measured participants’ maximal cardiorespiratory endurance by using a modified Bruce Protocol treadmill test on a Medtrac ST55 Control Treadmill (Quinton Instrument Company, USA). This protocol involved running on a treadmill, with both the speed and slope increasing every 3 minutes (Kalyani, Ebadi, Mehri, & Jamshidi, 2008), until volitional exhaustion occurred or other criteria were met, as explained in detail hereafter. After getting familiar with the exercise equipment, each participant was fitted with head gear and a mouthpiece to collect expired gases using semi-computerized open-circuit spirometry with the logic pathway on a Vmax system (Vmax Spectra Series Model 29, VIASYS Respiratory Care Inc., USA), to extract the following respiratory parameters: oxygen uptake (VO2), minute ventilation (Ve), carbon dioxide output (VCO2), and respiratory exchange ratio (RER, VO2/VCO2), with a sampling interval of 20 s to determine the maximal oxygen uptake during the graded exercise test (GXT). Throughout the GXT, we monitored heart rate and rhythm via electrocardiography (Yorba Linda, VIASYS Respiratory Care Inc., California) and a Polar HR monitor. Each test session included a 3-min warm up, a GXT on a motor-driven treadmill and a 3-min cool down. During the VO2max test, participants were verbally encouraged to continue exercising until exhausted, and the test was terminated based on the following four criteria: (1) indication of maximal exhaustion; (2) peak HR reaching more than 90% of the theoretic age-predicted maximum (220 – age); (3) a plateau in oxygen consumption corresponding to an increase of less than 150 mL in VO2 values despite the increase in exercise workload; or (4) an RER greater than 1.15 (Medicine, 2000; Wang et al., 2015, 2019).

2.3.2. Training task

The present study employed an adaptive dual n-back task described by Jaeggi et al. (2008), in which eight types of shape with eight hues were presented sequentially on a computer screen, at a rate of 3000 ms [i.e., stimulus presentation (500 ms); inter-trial interval (ITI) (2500 ms)] (see Fig. 2). The training task was freely available software (http://brainworkshop.sourceforge.net/) (Jaeggi et al., 2008). In this study, participants had to process two streams of stimuli (color and shape) simultaneously. All stimuli varied across eight hues and seventeen shapes and were presented at the center of a 3 x 3 grid on the monitor. A response was required whenever one of the presented stimuli matched the one presented n positions back (i.e., 1-back, 2-back and so on) in the sequence. There were eight color targets and seventeen shape targets per block, and the shape-color combinations were determined randomly. Participants were asked to make responses manually by pressing the “P” key with their left index finger for the color targets (color matched), and on the “J” key with their right index finger for shape targets (color matched). No responses were needed for non-targets (non-matched stimulus). We did not teach or recommend any particular strategies to improve performance.

Level of training difficulty (i.e., n-back level) was used to track participants’ performance. The task was adaptive, with level changing as a function of individual task performance. After each block, performance for each participant was recorded and analyzed automatically, and n-back levels were adapted to individual performance: an accuracy score (20 + n trials within a block) ≥80% correct led to an increase to the immediately superior n-back level. Three consecutive scores of < 50% correct led to a decrease of the immediately inferior n-back level. For all other cases (i.e., 50–79% correct), n-back levels remained unchanged (Kundu et al., 2013). There were 20 + n trials per block, and 20 blocks within each training session.

2.3.3. Transfer task

We used a shape-color feature binding working memory task with three different conditions (color vs. shape vs. shape-color binding) to assess transfer. The three conditions were blocked, with interleaved blocks, resulting in three sequences that either started with a color, shape, or a binding block, counterbalanced across all participants. Each block began with 12 practice trials, before the 100 testing trials. In this task, participants were first presented with a fixation cross (500 ms), followed by the stimulus at encoding phase (200 ms), a retention interval (1500 ms), the retrieval phase (until response or 2000 ms had elapsed after the onset of the display), and ITI (1500 ms) (Fig. 3a).

The stimuli were identical to those in Yang, Tseng, Huang, and Yeh (2012) (see Experiment 7). A set of 64 colored shapes was used, and the stimulus set was based on all possible combinations of eight shapes and eight colors (Fig. 3b). Each shape subtended a visual of approximately 25°.
1.06° × 1.06°. Each trial consisted of an encoding phase and a retrieval phase. In the encoding phase, four colored shapes were randomly selected from the stimulus set, consistent with the limits of working memory capacity (Cowan, 2001), and were placed randomly at the four corners of a square grid that subtended an area of 6.96° × 6.96°. In the retrieval phase, a single stimulus was presented at the center of the screen. This position was used as a neutral location because the displayed stimuli in the encoding phase were never presented there (Wheeler & Treisman, 2002).

In the color condition (color-only), the stimuli in the encoding phase were four colored squares (drawn from a pool of eight colors), and a single colored square in the retrieval phase. Participants were asked to judge whether the color in the retrieval phase had been present during the encoding phase. In the shape condition (shape-only), the stimulus display consisted of four black shapes (drawn from a pool of eight shapes), and a single black shape in the retrieval phase, with participants judging whether this shape was present initially. In the shape-color binding condition, four colored shapes were displayed and never repeated at encoding phase in each trial, and a single colored shape was presented in the retrieval phase. On target trials, this involved an original shape-color combination randomly selected from the original display at the encoding phase. By contrast, on lure trials, color and shape were re-combined from two different selected items at random from the encoding display. Participants were required to judge whether the combination of color and shape was present initially.

Participants were told to make judgements accurately by pressing the “N” key on the keyboard for target-present trials and the “M” key for target-absent trials, whereas speed was not emphasized (Allen, Baddeley, & Hitch, 2006; Yeh, Yang, & Chiu, 2005). For each condition, 50% of the trials were target-present trials, whereas the remaining 50% consisted of target-absent trials (i.e., lure). An auditory feedback (beep) denoted an incorrect response.

2.4. EEG recording

The electroencephalographic recording procedure was performed in a manner similar to previous studies (Wang, Yang, et al., 2017). EEG activity was recorded using a Nu-Amps EEG amplifier and the Scan 4.5 package (Neuroscan Inc., El Paso, TX, USA) with 32 electrodes mounted in an elastic cap (Quik-Cap; Compumedics, Neuroscan Inc.) designed for the International 10–20 System. The left and right mastoids were used as online references and a ground electrode was placed on the mid-forehead on the Quik-Cap. In addition, two sets of bipolar electrodes were placed on the upper and lower sides of the left eye, and on the canthi of both eyes in order to monitor vertical (VEOG) and horizontal (HEOG) eye-movements. Electrode impedances were kept below 10 kΩ. Electroencephalography data were acquired with an analog—digital rate of 1000 Hz per channel, filtered with a 5th order Butterworth bandpass filter (0.05–70 Hz), a 60-Hz notch filter, and written continuously to hard disk for subsequent offline analysis.

2.5. Data reduction

2.5.1. Behavioral data

E-prime 2.0 was used to record behavioral performance of the transfer task in terms of hit and false alarm rates, which were used to compute $d'$, a valid index for assessing working memory performance (Hsu, Tseng, Liang, Cheng, & Juan, 2014; Tseng, Chang, Chang, Liang, & Juan, 2016). The estimation of $d'$ signal followed the detection theory (Macmillan & Creelman, 1991). The rationale for this estimate is that the greater difference between hit and false alarm rate, the better sensitivity. The formula for $d'$ calculation is $z$ (hit rate) − $z$ (false alarm rate), which indicates the $z$ score difference between the averages of the target-present and the target-absent distributions. By subtracting their $z$-scores, we can obtain a valid discrimination index which is not confounded by participants’ response bias.
2.5.2. EEG preprocessing

Prior to the complexity analysis of EEG, ocular artifacts were corrected from the EEG signal with a regression-based algorithm implemented in the Neuroscan Edit 4.5 software (Neuroscan, Inc.) (Semlitsch, Anderer, Schuster, & Presslich, 1986). The offline ocular-corrected EEG data were then locked to the stimulus onset, and were segmented into epochs: 1500 ms before the encoding stimulus onset (S1), and lasting until 1500 ms after the retrieval stimulus (S2). Trials containing artifacts with amplitudes exceeding ±150 μV were discarded. After the artifact rejection, the average numbers of segments were 77.64 ± 8.54 for the exercise group (color: 83.80 ± 9.05, shape: 78.27 ± 10.10, and binding: 70.87 ± 9.88) and 84.24 ± 5.14 for the control group (color: 87.13 ± 3.93, shape: 86.73 ± 5.23, and binding: 78.87 ± 7.98) in the pretest session, while were 84.56 ± 6.79 for the exercise group (color: 87.26 ± 8.44, shape: 84.87 ± 8.08, and binding: 81.53 ± 7.38) and 82.82 ± 8.18 for the control group (color: 85.40 ± 7.62, shape: 82.53 ± 9.28, and binding: 80.53 ± 10.33) in the posttest session. A 5th order Butterworth low-pass filter with cutoff frequency at 70 Hz was applied to the epoched EEG data (Liang et al., 2014).

2.5.3. Multiscale entropy analysis of EEG

This study investigated neural adaptability by calculating brain signal complexity at different time scales via multiscale entropy analysis (Costa et al., 2002, 2005; Goldberger et al., 2000). The MSE analysis was calculated in two steps and was performed from time scale 1 to 20 in the following time windows: (1) 0 to 1500 ms relative to encoding stimulus onset (S1) and (2) −100 to 1000 ms for the retrieval stage (S2) in each of the condition (i.e., color, shape, and binding conditions) of the transfer task. This choice of time intervals of retrieval phase for the MSE analysis was to obtain as many data points as possible and to avoid any influence from response execution.

For the MSE analysis, first, the algorithm progressively down-sampled the EEG post stimulus time series \(x_1, x_2, \ldots, x_n\) for each trial in each condition. This down-sampling procedure was defined as a coarse-grained procedure along various time scales in the MSE analysis. For timescale \(r\), the coarse-grained time series \(Y^{(r)} = (y(1), y(2), \ldots, y(N^{(r)}))\) was obtained by averaging data points within non-overlapping windows of length \(r\). Therefore, each element of a coarse-grained time series, \(y(j)\), is calculated according to:

\[
y(j) = \frac{1}{r} \sum_{i=(j-1)r+1}^{jr} x_i, \quad \text{where } 1 \leq j \leq N^{(r)}, \quad N^{(r)} = \frac{N}{r}.
\]

Second, the algorithm computed the sample entropy for each coarse-grained time series \(Y^{(r)}\). Note that all the superscripts \((r)\) are omitted in the following to simplify the notation. There are two specified parameters for calculating the sample entropy: pattern length \(m\) and tolerance level \(r\) for similarity comparison. Given the coarse-grained time series \(Y\), sample entropy was calculated as follows: first, construct \(N - m + 1\) vectors

\[
Y_m(i): Y_m(i) = (y(i + k)), \quad 0 \leq k \leq m - 1
\]

and the distance between two vectors is defined as the absolute maximum difference between the corresponding scalar components

\[
d(Y_m(i), Y_m(j)) = \max(|y(i + k) - y(j + k)|), \quad 0 \leq k \leq m - 1.
\]

Given \(r\), \(n^{(r)}\) is defined as the number of vectors \(Y_m(i)\) falling within vector distance \(r\)’s of \(Y_m(i)\) without allowing self-matches, where \(s\) is the standard deviation of the original time series. Similarly, \(n^{(r+1)}\) is defined as the number of vectors \(Y_{m+1}(i)\) falling within vector distance \(r\)’s of \(Y_{m+1}(i)\). Finally, sample entropy was defined by the negative natural logarithm of the conditional probability that a time series of length \(N\), having repeated itself within a tolerance \(r\)’s (similarity factor) for \(m\) points pattern, will also repeat itself for \(m + 1\) points pattern

\[
S_e(m, r, N) = \ln \frac{\sum_{i=1}^{N-m} n^{r+1}_i}{\sum_{i=1}^{N-m} n^{r}_i}.
\]

Although no guidelines exist to optimize parameters of \(m\) and \(r\) values when calculating sample entropy values in EEG studies, some theoretical and clinical applications have suggested setting \(m = 1\) or 2 and \(r = 0.1–0.25\) to provide a high validity for sample entropy in EEG signal of length \(N\) ranging from 100 to 5000 data points (Escudero, Abásolo, Hornero, Espino, & López, 2006; Takahashi et al., 2009). Because we only obtained short EEG epochs (i.e., 1500 and 1100 data points for encoding and retrieval stages, respectively), here the pattern length, \(m\), was set to 1 (i.e., one consecutive data point was used for pattern matching), whereas the similarity criterion, \(r\), was set to 0.25, meaning that data points were considered to be indistinguishable if the absolute amplitude difference between them was ≤25% of the time series standard deviation. Because previous research has suggested that data lengths of \(10^m\) to \(20^m\) (pattern length) should be sufficient to estimate sample entropy (Richman & Moorman, 2000), estimation of sample entropy in the current coarse-grained EEG data (before the coarse-graining procedure, 1500 time points in encoding stage and 1100 ms time points in retrieval stage) may be sufficient for \(m = 1\) with time scales 1–20. The time scale indicates the length of each non-overlapping time bin within which the original data were averaged; for example, time scale 20 refers to averaging within each 20 ms window when the original sampling rate was 1000 Hz. EEG data processing was performed using SPMS and custom MATLAB (Math Works) scripts. The algorithm for MSE analysis can be found at http://www.pysnetresearch.org/tools.html.

Our focus was on midline electrodes (i.e., Fz, FCz, Cz, CPz, Pz, Oz) for the following two reasons. First, MSE values at these electrode have been demonstrated to have the capacity to reveal individual differences in domain-general cognitions (Grundy et al., 2017; Wang et al., 2014). Second, empirical studies have shown that event-related potential (ERP) or EEG signals from midline electrodes over fronto-parietal areas can capture the critical processes associated with working memory processes (Chang, Huang, Chen, & Hung, 2013; Kamijo, O’Leary, Pontifex, Themanson, & Hillman, 2010; Kamijo et al., 2011; Wang, Tseng, Liu, & Tsai, 2017).

We tested the effects of Session (pretest, posttest), Condition (color, shape, binding), Group (exercise, control), and Electrode (Fz, FCz, Cz, CPz, Pz, Oz) separately and used \(p < .05\) (uncorrected) or \(q^* < .05\) with false discovery rate (FDR) correction for multiple comparisons (Benjamini & Yekutieli, 2001). However, for any post hoc analysis, we only reported effects with \(q^* < .05\) FDR corrected.

3. Results

3.1. Demographic data

Participant demographic data are shown in Table 1. Two sample t-tests revealed that demographic variables including age, \(t(28) = −.62, p = .536\), height, \(t(28) = 1.02, p = .317\), weight, \(t(28) = .14, p = .889\), BMI, \(t(28) = −.49, p = .631\), and IPAQ, \(t(18.73) = −.15, p = .140\), did not differ between groups. Further, for the VO2max value, a mixed-
design ANOVA with Group (exercise vs. control) as a between-subject factor and Session (pretest vs. posttest) as a within-subject factor showed no significant main effects of Group \([F(1, 28) = 1.48, p = .234]\), Session \([F(1, 28) = .65, p = .427]\), or Group × Session interaction \([F(1, 28) = 1.11, p = .301]\).

3.2. Behavioral performance

Behavioral data were analyzed using two complementary statistical frameworks, with both frequentist (in SPSS 18.0) and Bayesian (R-proj:SCR_001905; R Core Team, 2018) hypothesis testing. Although we mainly focus on the frequentist results, the use of Bayesian statistics can further quantify the strength of evidence, which helps inform the interpretation of our behavioral results. All priors used in our analyses were the default scales (Morey & Rouder, 2015), in line with previous exercise training (Moreau et al., 2017) or cognitive training studies (Stojanoski, Lyons, Pearce, & Owen, 2018; Tsai et al., 2018b). Typically, a Bayes Factor (BF) greater than three is considered moderate evidence in favor of the hypothesis tested (i.e., the null or the alternative), whereas a BF between one and three suggests the data are inconclusive (Dienes, 2014). All analyses were set at 10^4 iterations, with diagnostic checks for convergence. One chain per analysis was used for all analyses reported in the paper, with a thinning interval of 1 (i.e., no iteration was discarded).

3.2.1. Training effects

A mixed-design ANOVA and a Bayesian mixed-design ANOVA on n-back level scores with Session (Baseline to D20) as a within-subject factor and Group (exercise, control) as a between-subject factor, showed a significant main effect of Session \([F(3.71, 103.81) = 86.04, p < .001, \eta^2_p = 0.75, \text{Greenhouse–Geisser-corrected}; \text{BF}_{10} = 4.73, \pm 0.1\%]\), suggesting a strong training effect for both groups. This indicates that training was effective, with steady task improvements across the training sessions. However, there was no Group effect \([F(1, 28) = 2.23, p = .146; \text{BF}_{10} = 0.62, \pm 0.54\%]\) and only weak-to-moderate evidence for a Group × Session interaction \([F(20, 56) = 1.67, p = .035, \eta^2_p = 0.06; \text{BF}_{10} = 2.33, \pm 1.75\%]\). Post hoc t-tests with Bonferroni correction and Bayesian pairwise comparisons did not reveal any group difference across sessions. This suggests that type of regimen did not induce fundamentally different trends of learning in terms of training performance, at the level of n-back task (Fig. 4). However, the focus of the present study was to explore transfer effects, which we detail hereafter.

3.2.2. Transfer effects

Participants’ d’ scores were submitted to a 2 (Session: pretest, posttest) × 3 (Condition: color, shape, binding) × 2 (Group: exercise, control) mixed-design ANOVA and Bayesian mixed-design ANOVA. We found overwhelming evidence for a main effect of Condition \([F(1, 61, 40.54) = 16.66, p < .001, \eta^2_p = 0.37, \text{Greenhouse–Geisser-corrected}; \text{BF}_{10} = 397140, \pm 0.42\%]\), with higher d’ values for the color condition \((2.66 \pm 0.97)\), intermediate d’ for the shape condition \((2.17 \pm 0.62)\), and lower d’ for the shape-color binding condition \((1.79 \pm 0.64)\). In addition, we observed a main effect of Session \([F(1, 28) = 7.61, p = .010, \eta^2_p = 0.21; \text{BF}_{10} = 3.27, \pm 0.67\%]\) and weak evidence for a Group × Session interaction \([F(2, 56) = 5.61, p = .025, \eta^2_p = 0.17; \text{BF}_{10} = 1.81, \pm 2.1\%]\). Post hoc t-tests and Bayesian pairwise comparisons for the Session difference within Group showed d’ score overall gains from pretest to posttest for the exercise group (pretest: 1.88 ± 0.77 vs. posttest: 2.47 ± 0.68; \(p = .007\); \text{BF}_{10} = 7.39, ± 3.6\%e-6\%) but not for the control group (pretest: 2.22 ± 0.44 vs. posttest: 2.26 ± 0.77; \(p = .73\); \text{BF}_{10} = 0.28, ± 2.9e-5\%).

In addition, we subtracted working memory performance in the pretest from the posttest session for each participant to compute training gain scores. Independent t-tests and Bayesian pairwise comparisons on gain scores for each condition showed strong evidence of group differences for the shape and binding conditions \((t(28) = −3.07, p = .005; \text{BF}_{10} = 8.89, ± 6.7e−7\%\) and \(t(28) = −3.89, p = .001; \text{BF}_{10} = 50.2, 3.9e−7\%\), respectively), but not for the color condition \((t(28) = −0.29, p = .775; \text{BF}_{10} = 0.35, ± 6.4e−7\%)\) (Fig. 2). Because the shape and binding conditions were arguably more difficult, with lower d’ overall scores at pretest (respectively \(M = 2.1, SD = 0.64\) and \(M = 1.6, SD = 0.71\), compared with \(M = 2.4, SD = 1.09\) for the color condition), this result might suggest that group differences in transfer only emerge given a certain level of task difficulty (Fig. 5).

3.3. EEG multiscale entropy analysis

Because there have been no studies using Bayesian hypothesis testing to examine the effect of exercise or cognitive training in the context of MSE analysis of EEG signals, we lacked appropriate references for setting plausible priors on EEG data. We thus conducted frequentist analyses only on our EEG data.

---

Fig. 4. Adaptive dual n-back training (mean n) across sessions, split by group (Exercise vs. Control). Small dots show individual data points, large dots show the mean for each group at a given time point, and error bars represent standard error of the means.
3.3.1. Encoding & maintenance stage

For the Condition effect, no differences in MSE across color, shape, and binding conditions were detected across groups and electrodes for all time scales in the pretest session (all $q^s > .05$ FDR corrected or all $p > .05$ uncorrected), suggesting that, prior to training, the three conditions were not associated with differences in brain signal complexity during encoding and maintenance phases in working memory. In the posttest session, although repeated measures ANOVA did not survive after FDR correction for multiple comparisons, a trend of Condition was observed for the exercise group (all $p < .05$ uncorrected) but not for the control group (all $q^s > .05$ FDR corrected or all $p > .05$ uncorrected) (Fig. 6a). We detected trends around Fz (scale 1 – 16, 18, 19) and FCz (scale 10 – 14, 16). Specifically, post hoc analyses with FDR correction for multiple comparisons revealed higher sample entropy values at Fz (scale 1 – 20), FCz (scale 1 – 20), and Pz (scale 1 – 10) in the binding condition relative to color condition (post hoc comparisons, all $q^s < .05$ FDR corrected), whereas no differences were observed in other condition comparisons (binding vs. shape or shape vs. color) (all $q^s > .05$ FDR corrected) (Fig. 7a). This pattern contrasted with that of the control group, for which we found no differences in MSE values at any time scales, task conditions, electrode sites, or test sessions (all $q^s > .05$ FDR corrected or all $p > .05$ uncorrected) (Fig. 7b). No differences in MSE values between groups were found for any time scales, task conditions, electrode sites, or test sessions (all $q^s > .05$ FDR corrected or all $p > .05$ uncorrected). No condition, group, or session effects were observed at Oz (all $q^s > .05$ FDR corrected or all $p > .05$ uncorrected).

3.3.2. Retrieval stage

For the Condition effect, no differences in MSE between color, shape, and binding conditions were found across groups and electrodes for all time scales in the pretest session (all $q^s > .05$ FDR corrected or all $p > .05$ uncorrected), suggesting that the type of condition did not significantly modulate brain signal complexity in both groups prior to the intervention. In the posttest session, although the repeated measures ANOVA did not survive FDR correction when controlling for multiple tests across conditions, electrodes, and time scales, we observed a main effect of Condition at a less conservative significance level...
for the exercise group (all ps < .05 uncorrected) but not for the control group (all qs* > .05 FDR corrected or all ps > .05 uncorrected) (Fig. 6b). Such trend was detected around Fz (scale 1–17, 19), FCz (scale 2–17, 19), Cz (scale 10, 13–14, 17), CPz (scale 9, 11–12, 14) and Pz (scale 1–12). More specifically, the post hoc analysis with FDR correction for multiple comparisons revealed that the exercise group exhibited higher sample entropy values at Fz (scale 1–17), FCz (scale 1–19), Cz (scale 1–14, 16), CPz (scale 1–14, 19), and Pz (scale 1–13) (all qs* < .05 FDR corrected) in the binding condition, while at Fz (scale 4–20), FCz (scale 3–20), Cz (scale 7–20, 16), CPz (scale 8–20), and Pz (scale 9, 11–17, 19–20) (all qs* < .05 FDR corrected) in the shape condition, relative to the color condition (Fig. 8a). In contrast, this effect was found only at Fz (scale 3–14) in the binding condition for the control group (all qs* < .05 FDR corrected) (Fig. 8b). No differences in MSE values between groups were found for any time scales, task conditions, electrode sites, or test sessions (all qs* > .05 FDR corrected or all ps > .05 uncorrected). No condition, group, or session effects were observed at Oz (all qs* > .05 FDR corrected or all ps > .05 uncorrected).

4. Discussion

Our study was the first to investigate the efficacy of antecedent aerobic exercise on adaptive working memory training using multiscale entropy analysis of EEG signals. Based on previous findings, we hypothesized that the employ of preceding exercise would lead to greater training improvements and would facilitate transfer at both behavioral and electrophysiological levels. We designed an intervention combining acute bouts of aerobic exercise and dual n-back training, and a control condition combining reading with the same dual n-back regimen.

Our results revealed that both groups substantially improved dual n-back performance across the intervention, with further evidence...
showing a marginally higher increase in the exercise group. In terms of transfer, the exercise group showed greater improvements post intervention, with stronger gains on the conditions requiring greater attentional demands (i.e., shape and binding conditions) in comparison to controls. Specifically, we observed training-induced changes in task-related modulations on MSE values in the exercise group only, with greater MSE in the binding condition than in the color condition during the encoding and maintenance phases, and was more pronounced in the shape and binding conditions relative to the color condition during the retrieval phase. These findings suggest that participants on the multi-modal regimen inclusive of physical and mental demands increased their working memory capacity, leading to transfer to a different task.

Our findings are in line with the claim that cognitive training supplementing with physical exercise improves transfer, and is particularly encouraging given that working memory capacity is a reliable predictor of professional and academic success (Deary, Strand, Smith, & Fernandes, 2007). This study further extends these views in several ways, which we discuss hereafter.

4.1. The effect of acute bouts of exercise on working memory training

The present study adopted a domain-general computerized working memory training paradigm, the dual n-back task, which has been used repeatedly in children and young adults (Jaeggi et al., 2008, 2011; Kundu et al., 2013). Our findings are in line with studies demonstrating substantial improvements in n-back levels associated with repeated training, regardless of the specific type of intervention.

This study went beyond prior work by examining the impact of aerobic exercise on training effects. Specifically, we employed the dual n-back training paired with antecedent aerobic exercise to explore
whether the beneficial effects acute exercise has on n-back tasks (Li et al., 2014; Weng et al., 2015) can be extended to a greater extent on a longitudinal basis. We found that aerobic exercise provided a marginal boost to n-back training performance. This may be consistent with work by Ward et al. (2017), which showed that 20 physical exercise sessions did not significantly improve learning on a dual n-back paradigm.

Nevertheless, these findings do not mean that physical exercise has no effect on learning during dual n-back training. Despite the lack of a visible effect at the behavioral level, single bouts of exercise may induce changes at the macro-neural level (Li et al., 2014), which might eventually facilitate performance. Measures using neuroimaging approaches might be more sensitive to this type of acute exercise-induced differences in cognitive training. In addition, it is also plausible that other type of acute exercise, such as high-intensity interval training (HIIT) could lead to greater training improvements, as suggested by recent evidence (Kao et al., 2017; Kujach et al., 2017; Moreau et al., 2017). Moreover, while we employed an adaptive cognitive training, the intensity of aerobic training was controlled to be consistent over the training sessions. Thus, it is likely that a progressive aerobic training would have resulted in greater enhancement in both aerobic fitness level and cognitive learning, and this might lead to the group-differences in the n-back difficulty. Although it is unfortunate that this study is unable to test these possibilities, our results are certainly noteworthy and encouraging of further investigation.

4.2. The effect of combined working memory and exercise training on transfer

One of the central questions of the present study was whether aerobic exercise could elicit greater transfer of dual n-back training to an untrained working memory task. Evidence for reliable transfer effects has been elusive thus far (Au et al., 2015; Jaeggi et al., 2008; Jaeggi, Buschkeuhl, Shah, &Jonides, 2014; Melby-Lervåg &Hulme, 2013; Redick et al., 2013), and even near-transfer effects have not been found consistently (Melby-Lervåg, Redick, &Hulme, 2016; Soveri et al., 2017). Recently, near-transfer effects have been refined by distinguishing task-specific from task-general near transfer. In the present case, the former refers to transfer to any untrained n-back tasks, while the latter implies transfer to a different working memory task (Soveri et al., 2017). This distinction is important, since it has been argued that task-specific near transfer often emerges as a consequence of the acquisition of task-specific strategies rather than improved working memory per se (Soveri et al., 2017). Thus, in order to explore whether aerobic exercise has the capacity to increase the effectiveness of n-back training at the construct level, this study focused on task-general near transfer.

Specifically, we adopted a shape-color feature binding task to examine whether training effects could transfer to holding multiple features temporarily in an integrated manner, an important component of working memory and of other higher cognitive functions such as problem solving and decision making (Wheeler &Treisman, 2002). Analyses of behavioral measures revealed that the recall of shape-color binding was less accurate than the recall of shape, and that both were less accurate than the recall of color, across groups and testing sessions. These findings are in line with previous research (Brown &Brockmole, 2010; Ueno, Allen, Baddeley, Hitch, &Saito, 2011), suggesting that the feature binding task was reliably administered in this study. Performance differences across conditions may be due to color being a more salient and more easily processed feature than shape (Allen et al., 2006), in contrast with the integration of different features, typically more effortful and demanding (Ueno et al., 2011). These findings can be explained within Baddeley (2000)’s framework of working memory, which claims that feature binding requires central executive resources in the episodic buffer component; but see (Allen et al., 2006) for an opposite view.

Although it has been demonstrated that multimodal interventions combining cognitive and exercise training could produce greater training gains than exercise training alone (Daugherty et al., 2017; Heisz et al., 2017), this effect has not been as reliable when compared to cognitive training alone (Ward et al., 2017). The present results provide further evidence that dual n-back training paired with repeated acute bouts of aerobic exercise can induce domain-general changes in cognitive function. Importantly, the group receiving dual n-back training along with a non-exercise control condition (i.e., reading) did not show this effect. We should point out, however, that task-general near transfer following n-back training has been shown to be very small (i.e., approximately \(d = 0.2\)) (Soveri et al., 2017), and as such its practical implication might be limited (Melby-Lervåg et al., 2016). In line with this idea, we did not observe clear transfer for the control group. One possible explanation may be the possibility that the strategies honed during dual n-back training could not be transferred outside of that specific context. This is plausible given that the stimuli were presented sequentially in the n-back training regimen, but simultaneously in the feature binding task. If confirmed, this suggests that transfer from n-back training is task-specific (Soveri et al., 2017).

More specifically, we observed that the exercise group showed greater gains in the shape and binding conditions involving greater task difficulty than controls. One recent study by Heisz et al. (2017) observed a synergistic effect when combining cognitive and exercise training in a working memory task with a high degree of interference, but not for the condition tapping general recognition, suggesting that transfer might be task-dependent, and especially pronounced for higher levels of difficulty. Consistent with this idea, the magnitude of transfer is presumed to be associated with cognitive overlap between training and transfer tasks (Dahlin, Neely, Larson, Bäckman, &Nyberg, 2008). Because the n-back task involves a broad range of processes, spanning from updating, to active maintenance, capacity and interference control (Jaeggi, Buschkeuhl, Perrig, &Meier, 2010; Schmiedek et al., 2014), all with shape and color features, it is possible that exercise predominantly facilitates some of these components, yielding the strongest effects when task difficulty is high.

4.3. Training-induced changes in modulation of brain signal complexity in the transfer task

In order to further understand the mechanisms that could explain transfer, we investigated the training-induced changes in neural adaptability when performing the transfer task via MSE analysis of EEG (Costa et al., 2005; Liang et al., 2014). Broadly construed, MSE evaluates the amount of information contained in an electrophysiological signal produced by an area of cerebral cortex across fine-to coarse-grained timescales (Costa et al., 2002, 2005; Liang et al., 2014), with higher value of MSE having been associated with learning (Deco et al., 2011; Lafontaine et al., 2016) and greater knowledge representation (Heisz et al., 2012), and task demands (Catarino et al., 2011). These findings present evidence to suggest that MSE may reflect the repertoire of responses and neural adaptability in response to a particular stimulus represented in the brain, and may change transiently within a neural system based on the task demands (Heisz, Vakorin, Ross, Levine, &McIntosh, 2014). That is, the more information becomes available about a given task, the higher brain complexity typically is (Deco et al., 2011). We should note, however, that in the present study we did not observe any task-modulated changes in MSE value in the pretest session, suggesting that participants might not be able to acquire and maintain greater amount of information in conditions involving higher working memory load (i.e., memory for shape or shape-color binding), and thus failing to induce any differences in brain signal complexity across task conditions.

Interestingly, after 20 sessions of training, while the main effect of condition did not reach significance after FDR correction for multiple comparisons across electrodes and time scales, we found a trend during working memory encoding and maintenance stage in the exercise...
group. This was also observed within each electrode using FDR correction for multiple comparisons, suggesting that the binding condition may be associated with higher sample entropy values in EEG than the color condition at all time scales over frontal areas, and at finer time scales over the parietal area. This finding suggests that the intervention combining cognitive training and aerobic exercise helped increase neural adaptability for complex information processing, thus increase the sensitivity of brain response to task demands. Moreover, it may also indicate a training-induced enrichment in functional network that can accommodate greater amounts of integrated visual feature configurations, which in turn leads to increases in brain signal complexity. Specifically, previous work has suggested that MSE at finer scales represents local information, whereas MSE at coarser scales is an indicator of more distributed information (Heisz et al., 2012; Vakorin et al., 2011). Therefore, our findings may suggest enhancements in neural adaptability across local and distributed networks in the frontal regions, together with increased local neural processing in the parietal regions when encoding and maintaining a combination of shape-color features. This is consistent with the changes in activity that have been observed in these regions during encoding and maintenance in working memory (Bor, Duncan, Wiseman, & Owen, 2003; Klingberg, 2010; Smith & Jonides, 1999). By contrast, the absence of such cognitive modulation on MSE in controls suggests that n-back training alone might not be able to improve the neural adaptability associated with domain-general encoding.

Analyses of brain signal complexity during the retrieval phase of working memory provides additional insight into the effect of the multimodal regimen on central executive component of working memory. After training, the exercise group exhibited a trend for the effect of cognitive demands on MSE values across time scales at most electrodes over fronto-parietal areas. A further post hoc analysis with FDR correction for such effect within each electrode revealed greater MSE values in the binding and shape conditions relative to the color condition. MSE of EEG signal has been demonstrated to have great sensitivity to uncover subtle differences associated with experimental learning effect (Heisz et al., 2012) and to reflect individual differences in cognitive capacity (Grundy et al., 2017; Heisz, Gould, & McIntosh, 2015; Wang et al., 2014). Accordingly, our results indicate that a combination of cognitive and exercise training improves neural adaptability of the fronto-parietal network that supports executive processes in working memory. Further, given that neural complexity is considered an index of individuals’ adaptability to alternate between multiple functional states (Deco et al., 2011; Grundy et al., 2017), our findings may point to an improved switching ability between brain states in the exercise group, with greater changes in brain states seen in the condition requiring more executive processes (i.e., binding condition). This observation appears consistent with previous literature showing a selective improvement of combined cognitive and exercise training for a memory task requiring high-interference but not for a task involving general recognition (Heisz et al., 2017). Given that higher brain MSE is considered a marker for greater information processing related to improved behavioral performance (Carpentier, Moreno, & McIntosh, 2016), our MSE findings might help to explain why the exercise group showed greater cognitive gains than controls in the shape and binding conditions, but not in the color condition. That is, greater transfer effects may be visible on task conditions with a certain level of difficulty, and thus higher changes in brain signal complexity. Clearly, the above discussed findings should be taken with caution and requires replications in future studies given that they did not reach significance with more conservative correction for multiple comparisons.

In contrast, for the control condition, the MSE differentiation seen at finer time scales over frontal area when comparing binding and color conditions might indicate increased capacity in information processing at more local levels. This is plausible given that executive demands are crucial in the dual n-back task to access and manipulate visual features stored in working memory. This finding may also imply that MSE is sensitive enough to detect subtle differences associated with training effects, even in the absence of changes in behavioral performance.

Finally, we did not observe any effects at the occipital site, which could be explained by the fact that the feature binding task does not primarily rely on basic visual function, but rather on working memory capacity and executive process dominated by the frontoparietal network, which may suggest the specificity of the transfer of training regimes.

4.4. Limitations and implications for future research

We should also point out potential limitations and raise questions that warrant further investigations. First, given the possible influences of some demographic variables such as intelligence quotient (IQ) (Fry & Hale, 2000) and socioeconomic status (SES) (Hackman & Farah, 2009) on cognitive performance, future investigations should control these factors as confounding variables when appropriate, which may help resolve discrepancies across studies. Second, since a limited sample size reduces the likelihood of detecting a true effect (Button et al., 2013), the effects failing to survive FDR correction might be due to a lack of statistical power. Further investigation, with larger sample sizes, is thus warranted to address this issue. It is also worth noting that the parameters (i.e., duration, intensity, type, etc.) of exercise intervention were not experimentally manipulated in this study, and thus we cannot claim that cognitive training supplemented by moderate aerobic exercise is ideal – greater cognitive learning and its generalizability could possibly be elicited with longer durations, higher intensity, or even other types of exercise (i.e., HIIT or motorically complex exercise) (Kao et al., 2017; Kujach et al., 2017; Moreau, Morrison, et al., 2015, 2017; Wang et al., 2013). Moreover, using a single working memory task to assess one type of transfer effect (i.e., task-general near transfer effect) limits the generalizability of this study to other dimensions of transfer (i.e., task-specific near transfer or far transfer). Using multiple measurements helps minimizing noise, or random sources of error, while maximizing the influence of a true effect (Moreau et al., 2016). Therefore, it will be necessary to replicate and extend upon current findings by using multiple measurements of transfer effects. Finally, this study only assessed the transfer effect immediately following the intervention. Future research may extend the current findings by examining the retention effect following the cessation of the intervention.

Despite these limitations, results from this study provide further support for the effectiveness of multimodal interventions combining both cognitive training and exercise to elicit potent cognitive benefits in healthy populations (Daugherty et al., 2017; Heisz et al., 2017; Ward et al., 2017). These findings could have important ramifications to everyday life, especially in specific populations, for example those with low working memory capacity (Gathercole, Alloway, Willis, & Adams, 2006; Wang, Tseng, et al., 2017). In this sense, our findings also have practical implications for educational policies, as they suggest that exercise may be a viable approach to maximize learning efficiency in the classroom.

Author contributions statement

CHW, CTY, and WKL conceived and designed the experiments. CHW, JTL, and YYT performed the experiments. CHW, DM, CTY, and WKL analyzed the data. CHW, CLT, CTY and WKL contributed materials/experimental tools. CHW wrote the first draft of the manuscript. DM, CTY, WKL and CLT edited/commented on the manuscript. All authors reviewed the manuscript.

Conflict of interests

The authors declare no conflict of interest.
Acknowledgements

This work was sponsored by the Ministry of Science and Technology, R. O. C. (Grant numbers: MOST 104-2410-H-006-059; MOST 106-2410-H-008-038-MY3). The authors are very grateful to Chan Kuan for his assistance with Figs. 1 and 2 illustrations. In addition, we would like to thank Prof. Chi-Hung Juan for kindly providing EEG recording equipments and technical supports.

References


function in children in a randomized, placebo-controlled trial. eLife, 6


