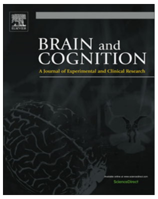
Brain and Cognition 97 (2015) 32–39

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c



Domain dependent associations between cognitive functioning and regular voluntary exercise behavior

Suzanne C. Swagerman [a,*], Eco J.C. de Geus [a,b], Marinka M.G. Koenis [c], Hilleke E. Hulshoff Pol [c], Dorret I. Boomsma [a], Kees-Jan Kan [a]

[a] *Department of Biological Psychology, VU University Amsterdam, van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands*
[b] *EMGO+ Institute of Health and Care Research, VU University Medical Center, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands*
[c] *Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands*

ARTICLE INFO

Article history:
Accepted 2 April 2015

Keywords:
Exercise
Physical activity
Neurocognition
Cognitive aging
Computerized Neurocognitive Battery

ABSTRACT

Regular exercise has often been suggested to have beneficial effects on cognition, but empirical findings are mixed because of heterogeneity in sample composition (age and sex); the cognitive domain being investigated; the definition and reliability of exercise behavior measures; and study design (e.g., observational versus experimental). Our aim was to scrutinize the domain specificity of exercise effects on cognition, while controlling for the other sources of heterogeneity.

In a population based sample consisting of 472 males and 668 females (aged 10–86 years old) we administered the Computerized Neurocognitive Battery (CNB), which provided accuracy and speed measures of abstraction and mental flexibility, attention, working memory, memory (verbal, face, and spatial), language and nonverbal reasoning, spatial ability, emotion identification, emotion- and age differentiation, sensorimotor speed, and motor speed. Using univariate and multivariate regression models, CNB scores were associated with participants' average energy expenditure per week (weekly METhours), which were derived from a questionnaire on voluntary regular leisure time exercise behavior.

Univariate models yielded generally positive associations between weekly METhours and cognitive accuracy and speed, but multivariate modeling demonstrated that direct relations were small and centered around zero. The largest and only significant effect size ($\beta = 0.11$, $p < 0.001$) was on the continuous performance test, which measures attention.

Our results suggest that in the base population, any chronic effects of voluntary regular leisure time exercise on cognition are limited. Only a relation between exercise and attention inspires confidence.

1. Introduction

Regular exercise has often been suggested to have beneficial effects on cognitive performance, but empirical findings do not always support this suggestion. As a result, the effectiveness of regular exercise behavior as a means to improve cognitive performance remains a subject of debate, not only among scientists, but also among policy makers. When published findings are summarized, associations between exercise behavior and cognitive performance appear positive on average, but vary considerably in strength (Fedewa & Ahn, 2011; Hindin & Zelinski, 2012; Ploughman, 2008; Singh, Uijtdewilligen, Twisk, van Mechelen, & Chinapaw, 2012; Taras, 2005; Trudeau & Shephard, 2008; Verburgh, Konigs, Scherder, & Oosterlaan, 2014). The literature provides four major sources of heterogeneity among study outcomes, the first concerning sample constitution (Singh et al., 2012). Study samples have differed greatly with respect to age, while the association strength between exercise behavior and cognitive performance is considered to differ between children, adolescents and adults (Hillman, Castelli, & Buck, 2005; Tomporowski, Davis, Miller, & Naglieri, 2008; but see Verburgh et al., 2014).

In childhood and adolescence exercise may influence the (rapid and specific) brain changes that take place during development, while in the elderly exercise may prevent (slow or general) deterioration of the brain during aging (Churchill et al., 2002; Fabel & Kempermann, 2008; Greenwood & Parasuraman, 2010; Hillman, Erickson, & Kramer, 2008; Kraft, 2012; Yuki et al., 2012).

---

* Corresponding author at: Department of Biological Psychology, VU University Amsterdam, van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. Fax: +31 (0)20 5988832.

E-mail addresses: s.c.swagerman@vu.nl (S.C. Swagerman), eco.de.geus@vu.nl (E.J.C. de Geus), m.m.g.koenis@umcutrecht.nl (M.M.G. Koenis), h.e.hulshoff@umcutrecht.nl (H.E. Hulshoff Pol), di.boomsma@vu.nl (D.I. Boomsma), k.j.kan@vu.nl (K.-J. Kan).

http://dx.doi.org/10.1016/j.bandc.2015.04.001
0278-2626/

Furthermore, rates of cognitive decline differ across sexes, which has been linked to the loss of estrogen (Kramer, Erickson, & Colcombe, 2006). Sex may be regarded as a source of heterogeneity in itself as the associations between exercise behavior and cognitive measures in samples consisting of a majority of women tend to be larger than in samples consisting of relatively many men (Colcombe & Kramer, 2003). A second major source of heterogeneity amongst study outcomes concerns the cognitive domain being measured. Recent studies (Colcombe & Kramer, 2003) suggest that cognitive functions are differently susceptible to exercise; executive functions may be more sensitive to exercise than, for example, long-term memory. Empirically however, little is known about how effects of exercise vary across cognitive domains, let alone about how these effects differ in their dependencies on age and sex. Many studies have focused on global cognitive measures, and outcomes thereof, such as academic achievement. This is unfortunate because they do not inform about the sensitivity of specific cognitive functions (Tomporowski et al., 2008). The present study is unique, in that we measured in a single, population representative sample cognitive performance across a wide range of well-defined, specific cognitive domains. The battery we used, the web-based Computerized Neurocognitive Battery (CNB), consists of 17 cognitive tests, and provides measures of accuracy as well as speed in the following cognitive domains: abstraction and mental flexibility, attention, working memory, memory (verbal, face, and spatial), language and nonverbal reasoning, spatial ability, emotion identification, emotion- and age differentiation, sensorimotor speed, and motor speed. Individual differences in these domains are substantially heritable and demonstrate genetic linkage (Almasy et al., 2008). Scores on the CNB are reliable and compare well to scores on traditional pen-and-paper tests in healthy samples as well as in clinical samples (e.g. schizophrenia patents, Gur, Ragland, Moberg, Bilker, et al., 2001; Gur, Ragland, Moberg, Turner, et al., 2001). While initially constructing the test battery, tests were selected from neuroimaging studies that showed selective activation of specific brain systems in the magnetic resonance imaging (MRI) scanner (Gur et al., 2010). Recently, the CNB tests adapted for administration in the MRI scanner replicated the brain areas that are activated by the CNB's cognitive domains. More specifically, the executive tests activated mainly frontal areas, memory tests involved anterior medial temporal regions, and a test measuring emotion identification activated temporo-limbic regions (Roalf et al., 2014).

A third source of heterogeneity amongst previous results, the definition and reliability of exercise behavior measures, has been discussed extensively in the literature. Studies have varied greatly in the conceptualization of exercise behavior, the broadest conceptualization being the inclusion of all forms of physical activity (i.e. every activity increasing energy expenditure above basal metabolic rate). However, self-reported physical activity corresponds poorly with actual physical activity (Prince et al., 2008). In addition, the idea that common, low intensity forms of physical activity will be sufficient to induce cognitive effects has been questioned; exercise likely needs to be carried out at a moderate to vigorous intensity to have effect on cognitive functioning (Colcombe & Kramer, 2003; Fedewa & Ahn, 2011; Hindin & Zelinski, 2012). It is recommended to focus on relatively vigorous activities, especially leisure time exercise activities: recall is relatively easy and quite accurate as these activities are self-initiated and often clearly defined in time. Indeed, voluntary regular leisure time exercise behavior demonstrated excellent test–retest reliability (de Moor, Boomsma, Stubbe, Willemsen, & de Geus, 2008; Stubbe, de Moor, Boomsma, & de Geus, 2007). In the present study, we will focus on this narrow but well-defined behavior, also because it is often the main target of health-promoting exercise interventions (Kahn et al., 2002).

A fourth source of heterogeneity concerns study design. This is an important source to recognize, because study designs are differently suited to estimate effects of physical activity. In experimental and clinical intervention studies the focus is usually on mean effects as a result of intervention, while the focus of observational studies lies on individual differences in voluntary behavior and on dose–response relationships. Furthermore, intervention studies – experimental studies included – have varied widely in their definition of intervention. In addition, not all intervention studies have been truly experimental; clinical intervention is often performed in non-random samples (Singh et al., 2012; Tomporowski et al., 2008). Another distinction concerns studies investigating the effects of acute physical exercise, and studies that investigate the effects of chronic physical exercise (Verburgh et al., 2014). In the first, the focus is on (short-term) cognitive enhancement right after a single bout of exercise, typically within less than an hour. In the latter, the focus is on (long-term) cognitive enhancement as the result of regular exercise over longer periods, typically weeks or months. Although there is ample evidence for beneficial effects of acute physical exercise (Verburgh et al., 2014), studies into the effects of chronic physical exercise are scarce, hence the call for more research.

The general objective of the present study is to investigate the chronic dose–response association between voluntary regular leisure time exercise behavior and cognitive performance across a wide range of cognitive domains, while controlling for other sources of heterogeneity. To this end, we first examine whether leisure time exercise associated with accuracy and speed scores, exploring whether and how these associations vary across domains. Next, we explore whether, how, and to what extent these associations vary when accounting for differences in age and sex. We end with a general discussion, in which the results of the present population-based observational study are compared with results from previous (high quality) intervention studies, which typically involve clinical-control designs.

2. Material and methods

2.1. Participants

The subject sample consisted of 472 males and 668 females from the Netherlands Twin Register (NTR) recruited from all over the Netherlands (Boomsma et al., 2006; van Beijsterveldt et al., 2013; Willemsen et al., 2013). The majority ($n = 1110$) was comprised of twin pairs and their family members (parents, children, siblings, and spouses) who volunteered in NTR projects. The rest ($n = 30$) was comprised of undergraduate students who piloted in these projects. The participants ranged in age from 10 to 86 years old (mean = 37.73, *SD* = 20.86, see Fig. 1).

2.2. Procedure

Studies and procedures were approved by the Medical Ethics Review Committee of the VU Medical Center Amsterdam and the Central Committee on Research Involving Human Subjects. The twins and their family members were approached by mail. In case of a positive response, a structured telephone call followed, which was informative about possible exclusion criteria (epilepsy, paralysis). The students were recruited at the university through flyers. They signed up themselves. Data collection took place either at home ($n = 536$) or in a laboratory (VU University Amsterdam, University Medical Center Utrecht, Amsterdam Medical Center, $n = 604$).

Cognitive performance was assessed on a 15 inch Macbook laptop, using the web-based Computerized Neurocognitive

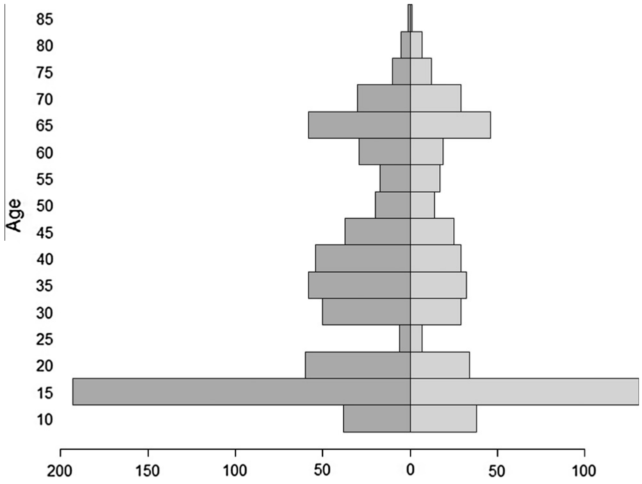


Fig. 1. Age distribution in females (dark gray) and males (light gray).

Battery (CNB, see below). The test administrator was placed behind the participant to be able to read the test instructions out loud and to provide feedback during practice trials. The administrator judged for each test if it was complete and valid (for example based on motivation or attention). On designated timepoints in between tests, the procedure, which lasted 1.30 h on average, could be paused. Students received study credits, others travel compensation and a gift voucher. All participants signed an informed consent form. For participants under 16 years parents gave additional written consent. All participants received feedback on their performance.

2.3. Materials

2.3.1. Cognitive performance

Cognitive performance (accuracy and speed) was assessed by the Dutch translation of the CNB as described by Gur et al. (2010, 2012). It comprises a total of 17 tests that assess cognitive performance on executive control, episodic memory, complex cognition, social cognition, and sensorimotor speed (Table 1). Accuracy was defined as either the percentage or the number of correct responses on a test, whereas speed was defined as minus the median response time $(R^{*}-1)$ in milliseconds for correct responses. Speed performance on the Finger-Tapping Test (TAP), however, was expressed as the number of taps one can produce within 60 seconds (alternating every 10 seconds between the left and right hand). TAP score thus indicates motor speed rather than response time. For all cognitive measures it held that higher scores reflected better cognitive performance.

2.3.2. Voluntary regular leisure time exercise behavior

Questions on exercise behavior were collected using a standardized interview (on the same day as the cognitive testing, $n = 894$) or a questionnaire (within 2 weeks of cognitive testing, $n = 246$) with identical questions. The first question was “Do you exercise regularly?”. When the answer was affirmative further information was gathered on the type of exercise (for example aerobics classes, soccer, or running) and on the involvement in this type of exercise (months a year, times a week, and average duration of the activity in minutes). Activities were excluded if they are not self-initiated or voluntary, like transportation (walking, biking), or physical education classes in school, as were general physical activities such as gardening. Voluntary exercise activities were only scored when participants had engaged in them for at least three months during the past year (Stubbe, Boomsma, & de Geus, 2005).

Next, we obtained the metabolic equivalent (MET) for each of the reported activities. Here, a MET = 1 corresponds to the rate of energy expenditure of an individual at rest (approximately one kcal/kg/h). Because children and adults differ in the energy cost of activities, MET scores of participants under age 18 were obtained using the Compendium of Physical Activities for Youth (Ridley, Ainsworth, & Olds, 2008), and of older participants using the Ainsworth’s compendium of physical activities (Ainsworth

Table 1
Cognitive domains and tests, order of administration, mean duration (in minutes), number of participants who completed the test (*N*), the mean score (and *SD*), and Cronbach’s alpha coefficients (α) of accuracy and speed.

Cognitive domain	Test name	Test label	Order	Duration	N	Accuracy			Speed		
						Mean	*SD*	α	Mean	*SD*	α
Executive control											
Abstraction/flexibility	Penn Conditional Exclusion Test[a]	CET	9	4.9	1125	1.9	0.8	[d]	2813.3	1392.6	[d]
Attention	Penn Continuous Performance Test[a]	CPT	3	5.3	1125	54.8	5.4	.86	487.7	49.1	.94
Working memory	Letter-N-Back Test[a]	LNB	6	9.2	1114	18.8	1.8	.77	537.7	118.0	.90
Memory											
Verbal Memory	Penn Word Memory Test[b]	CPW-i	5	3.1	1125	36.3	2.8	.62	1564.5	368.2	.92
	Penn Word Memory Test – delayed[b]	CPW-d	8	1.1	1124	35.0	3.3	.64	1541.7	376.6	.91
Face Memory	Penn Facial Memory Test	CPF-i	4	3.9	1123	31.4	3.5	.56	1992.7	544.2	.92
	Penn Facial Memory Test – delayed	CPF-d	7	1.5	1121	32.1	3.5	.57	1834.2	489.7	.89
Spatial Memory	Visual Object Learning Test[a]	VOLT-i	13	2.7	1117	16.0	2.3	.48	1973.8	554.6	.87
	Visual Object Learning Test – delayed[a]	VOLT-d	17	0.5	1115	15.4	2.4	.48	1811.5	519.7	.86
Complex cognition											
Nonverbal reasoning	Penn Matrix Reasoning Test	MAT	12	7.8	1129	13.9	5.2	−.30	10806.0	6959.8	.83
Language reasoning	Penn Verbal Reasoning Test[a,b]	VRT	14	1.8	1123	69.2	20.6	.53	8465.8	3332.5	.74
Spatial ability	Variable Penn Line Orientation Test[a]	LOT	16	5.5	1119	12.9	3.7	.79	10506.8	3861.8	.97
Social cognition											
Emotion Identification	Penn Emotion Identification Test	EI	2	2.3	1132	32.1	3.5	.62	2273.4	685.7	.92
Emotion Differentiation	Measured Emotion Differentiation Test	EDT	10	3.4	1131	28.0	3.5	.69	3721.0	1369.1	.94
Age Differentiation	Age Differentiation Test	ADT	15	3.0	1122	26.8	3.9	.74	3238.4	1493.5	.94
Sensorimotor											
Sensorimotor speed	Motor Praxis Test	MP	1	1.8	1130	20.0	0.4	.93	793.2	221.3	.95
Motor speed	Penn Computerized Finger-Tapping Test[a]	TAP	11	3.5	[c]	[c]	[c]	[c]	110.6	15.1	.96

[a] Short test version.
[b] Different items for children.
[c] No accuracy score available for TAP.
[d] Not amenable for calculating.

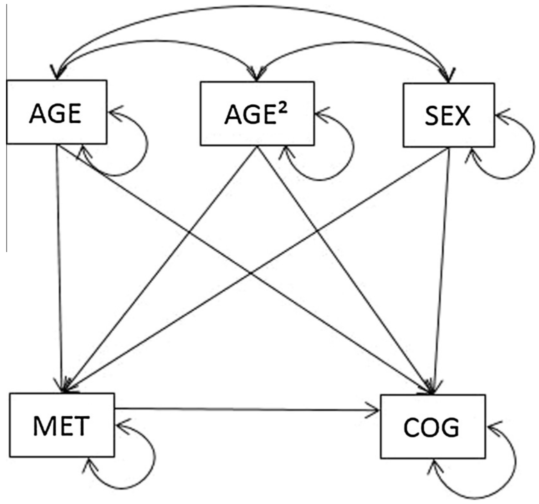


Fig. 2. Graphical representation of the multivariate model. The relation between cognitive test performance (COG) and weekly METhours (MET) depends on sex and age.

et al., 1993). Finally, we computed each individual's weekly METhours by multiplying each activity's MET by the hours per week spent on each activity and by summing these up over the exercise activities. Non-exercisers received a weekly METhour score of 0. Previous studies have shown that this variable has a high 6-month test–retest reliability of 0.82 (de Moor et al., 2008).

2.4. Statistical analyses

Descriptive statistics, which were calculated using SPSS 21.0 statistical package for Windows (IBM Corp., 2011), included means and standard deviations of the measured variables and of test duration (Table 1). Cronbach's alpha coefficients of internal consistency, which are commonly interpreted as indicators of reliability, were also calculated. To explore whether voluntary regular leisure time exercise behavior associated with cognitive performance, we ran for each cognitive variable a univariate regression model, in which cognitive performance was regressed on weekly METhours. Next, in order to statistically control for the effects of age and sex we fitted a multivariate path model (see Fig. 2) in which cognitive performance was regressed on weekly METhours, sex and (linear and quadratic terms of) age, while weekly METhours was regressed on sex and age. The quadratic age term was defined as the square of grand mean centered values and was included because inspection of the raw data suggested nonlinear relationships between cognitive performance and age (Fig. 3A and B provide examples). Sex and age were allowed to intercorrelate (although their intercorrelations were expected not to differ from zero).

Cases were excluded from statistical analyses whenever participants were considered to experience too much difficulty ($n = 1$) or when test performance was judged invalid by the experimenter (~0.8%), for example when computer or mouse issues had occurred, when participants demonstrated a lack of motivation, or when participants reported (noncognitive) impairments such as rheumatoid arthritis. Data of children in elementary school (under age 13) were removed from analyses ($n = 4$).

Because the majority of the cognitive variables were non-normally distributed (skewed) and scores were family clustered, the standard errors of the parameter estimates required correction. Correction was accomplished by analyzing the data in R using package lavaan (Rosseel, 2012) which included the option to use a robust sandwich estimation procedure and family number as cluster variable. This procedure allows for the analysis of nonnormally distributed, continuous outcome variables.

Because analyses were carried out for 33 cognitive, possibly related, measures, the Matrix Spectral Decomposition program (Li & Ji, 2005) was used to estimate the number of independent dimensions in the data, which was 23. This yielded a preferred significance level of $\alpha = 0.05/23 \approx 0.002$.

3. Results

3.1. Descriptive statistics

The descriptive statistics are provided in Table 1. Fig. 3C illustrated the complete distribution of weekly METhours across age and sex. The mean weekly METhours in the total sample was 15.6, males scoring higher (20.2) than females (12.3, $\beta = 7.97$, $p < 0.001$), as did young participants compared to older participants ($r = -0.23$, $p < 0.001$).

3.2. Modeling results

Table 2 and Fig. 4 summarize the modeling results. The univariate model yielded standardized regression coefficients that can be interpreted as bivariate correlations between weekly METhours and cognitive performance. With respect to accuracy these ranged from −0.02 (VRT) to 0.14 (MAT) and with respect to speed from

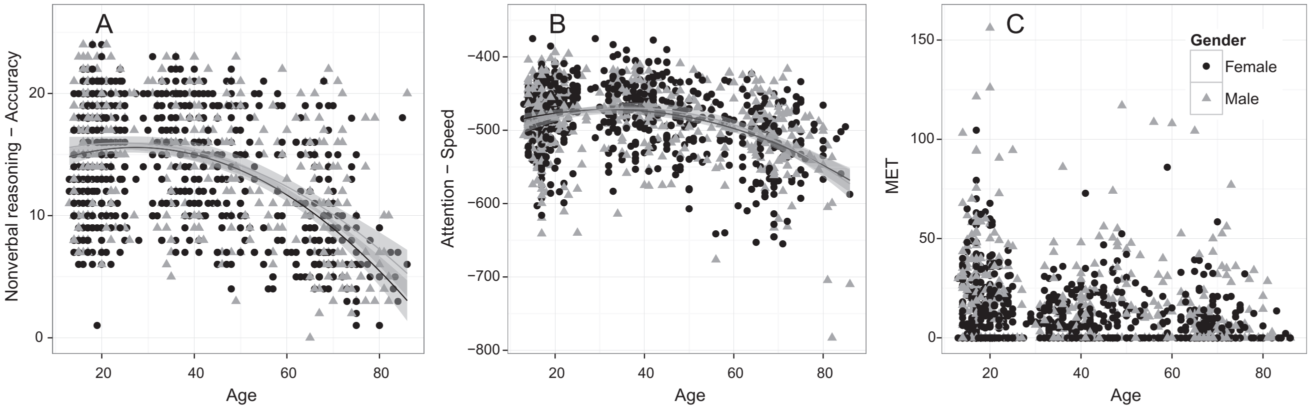


Fig. 3. (A and B) Examples of the nonlinear relationships between age and cognitive performance (accuracy and speed), including 95% intervals around the quadratic regression lines. (C) Illustration of non-normal distribution of weekly METhours against the age of female (black ●) and male (gray ▲) participants.

Table 2
Results from the univariate and multivariate analyses.

Cognitive domain	Test label	Univariate analyses Accuracy on Weekly MET-h		Multivariate analyses Accuracy on Weekly MET-h		Accuracy on Age		Accuracy on Age^2		Accuracy on Sex	
		β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*
Executive control											
Abstraction/flexibility	CET	.13 (.03)	.00	.04 (.03)	.10	−.30 (.03)	.00	−.11 (.04)	.00	.05 (.03)	.12
Attention	CPT	.11 (.03)	.00	.11 (.03)	.00	.13 (.04)	.00	−.45 (.04)	.00	.00 (.03)	.97
Working memory	LNB	.07 (.03)	.02	.03 (.03)	.40	−.16 (.03)	.00	−.24 (.05)	.00	−.03 (.03)	.35
Memory											
Verbal Memory	CPW-i	.04 (.03)	.18	.03 (.03)	.35	−.11 (.04)	.01	−.09 (.04)	.03	−.10 (.03)	.00
	CPW-d	.10 (.03)	.00	.05 (.03)	.05	−.19 (.04)	.00	−.21 (.04)	.00	−.05 (.03)	.07
Face Memory	CPF-i	.02 (.03)	.44	.00 (.03)	.89	.02 (.04)	.71	−.32 (.04)	.00	−.02 (.03)	.62
	CPF-d	.01 (.03)	.79	−.03 (.03)	.38	−.14 (.04)	.00	−.26 (.04)	.00	−.09 (.03)	.01
Spatial Memory	VOLT-i	.08 (.03)	.01	.01 (.03)	.80	−.24 (.04)	.00	−.09 (.04)	.01	.04 (.03)	.17
	VOLT-d	.09 (.03)	.00	.02 (.03)	.42	−.20 (.04)	.00	−.07 (.04)	.05	.07 (.03)	.02
Complex cognition											
Nonverbal reasoning	MAT	.14 (.03)	.00	.04 (.03)	.16	−.29 (.04)	.00	−.26 (.04)	.00	.05 (.03)	.06
Language reasoning	VRT	−.02 (.03)	.50	−.03 (.03)	.26	.14 (.04)	.00	−.30 (.04)	.00	.12 (.03)	.00
Spatial ability	LOT	.11 (.03)	.00	.03 (.03)	.34	−.14 (.04)	.00	−.17 (.04)	.00	.18 (.03)	.00
Social cognition											
Emotion Identification	EI	.13 (.03)	.00	.06 (.03)	.04	−.34 (.04)	.00	−.16 (.04)	.00	−.07 (.03)	.01
Emotion Differentiation	EDT	.05 (.03)	.07	.01 (.03)	.59	−.15 (.04)	.00	−.21 (.04)	.00	−.07 (.03)	.03
Age Differentiation	ADT	.04 (.03)	.24	.01 (.03)	.75	−.11 (.04)	.00	−.21 (.04)	.00	−.07 (.03)	.02
Sensorimotor											
Sensorimotor speed	MP	.10 (.03)	.00	.05 (.02)	.00	−.08 (.03)	.00	−.23 (.07)	.00	.04 (.03)	.21
Motor speed	TAP	–	–	–	–	–	–	–	–	–	–
		Univariate analyses Speed on Weekly MET-h		Multivariate analyses Speed on Weekly MET-h		Speed on Age		Speed on Age^2		Speed on Sex	
		β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*
Executive control											
Abstraction/flexibility	CET	.06 (.03)	.03	−.03 (.02)	.16	−.41 (.03)	.00	−.14 (.04)	.00	−.05 (.03)	.09
Attention	CPT	.03 (.03)	.29	−.01 (.03)	.75	−.11 (.04)	.00	−.31(.04)	.00	−.01 (.03)	.65
Working memory	LNB	.05 (.03)	.13	−.02 (.03)	.62	−.26 (.04)	.00	−.02 (.04)	.62	.03 (.03)	.33
Memory											
Verbal Memory	CPW-i	.12 (.03)	.00	.01 (.02)	.57	−.37 (.03)	.00	−.35 (.05)	.00	−.02 (.02)	.33
	CPW-d	.10 (.03)	.00	−.01 (.02)	.73	−.42 (.03)	.00	−.27 (.04)	.00	−.03 (.02)	.15
Face Memory	CPF-i	.06 (.03)	.06	.01 (.03)	.75	−.12 (.03)	.00	−.26 (.05)	.00	.00 (.03)	.90
	CPF-d	.06 (.03)	.03	−.01 (.03)	.68	−.26 (.03)	.00	−.21 (.05)	.00	.01 (.03)	.86
Spatial Memory	VOLT-i	.09 (.03)	.00	−.03 (.02)	.22	−.48 (.03)	.00	−.15 (.04)	.00	.00 (.03)	.98
	VOLT-d	.12 (.03)	.00	−.01 (.02)	.58	−.53 (.03)	.00	−.12 (.04)	.00	−.01 (.03)	.83
Complex cognition											
Nonverbal reasoning	MAT	−.01 (.03)	.71	−.04 (.03)	.14	−.23 (.04)	.00	.12 (.06)	.05	−.06 (.03)	.04
Language reasoning	VRT	−.01 (.03)	.62	−.03 (.03)	.35	−.05 (.04)	.18	−.02 (.04)	.57	−.01 (.03)	.85
Spatial ability	LOT	.10 (.03)	.00	−.04 (.02)	.11	−.37 (.03)	.00	−.30 (.04)	.00	.13 (.03)	.00
Social cognition											
Emotion Identification	EI	.07 (.04)	.06	−.04 (.03)	.20	−.46 (.03)	.00	−.27 (.04)	.00	−.07 (.03)	.01
Emotion Differentiation	EDT	.07 (.04)	.04	−.05 (.03)	.10	−.49 (.03)	.00	−.18 (.03)	.00	.00 (.03)	.95
Age Differentiation	ADT	.08 (.04)	.07	−.05 (.03)	.13	−.53 (.03)	.00	−.13 (.03)	.00	−.03 (.02)	.23
Sensorimotor											
Sensorimotor speed	MP	.16 (.03)	.00	.02 (.02)	.28	−.52 (.03)	.00	−.25 (.04)	.00	.01 (.02)	.78
Motor speed	TAP	.18 (.03)	.00	.06 (.03)	.02	−.19 (.03)	.00	−.34 (.03)	.00	.24 (.03)	.00

β = standardized regression coefficient, *SE* = standard error, *p* = *p*-value. See Table 1 for full name of the various cognitive tests.

−0.01 (reasoning tests MAT and VRT) to 0.18 (TAP), hence from negatively small to positively small. Medians were also small (0.09 for accuracy and 0.07 for speed), yet at $\alpha = 0.002$ about half of the coefficients were significant. However, the multivariate model, which yielded standardized path coefficients that can be interpreted as partial correlations, demonstrated that direct relationships were small and centered close to 0. Coefficients for accuracy ranged from −0.03 (CPF-d, VRT) to 0.11 (CPT, median = 0.03). And for speed coefficients ranged from −0.05 (EDT, ADT) to 0.06 (TAP, median = −0.02). Only the coefficient between weekly METhours and accuracy on the attention test (CPT) was significant ($\beta = 0.11$, $SE = 0.03$, $p < 0.001$).

4. Discussion

The aim of this paper was to scrutinize the domain dependency of the association between exercise behavior and cognition, while controlling for other major sources of heterogeneity. To this end, we explored in a population based sample, and across a wide range of cognitive domains, the age and sex independent associations using reliable and narrowly defined measures of voluntary regular leisure time exercise behavior.

Univariate analyses confirmed the existence of multiple associations between regular exercise behavior and cognitive performance. At face value, these findings may seem to support the

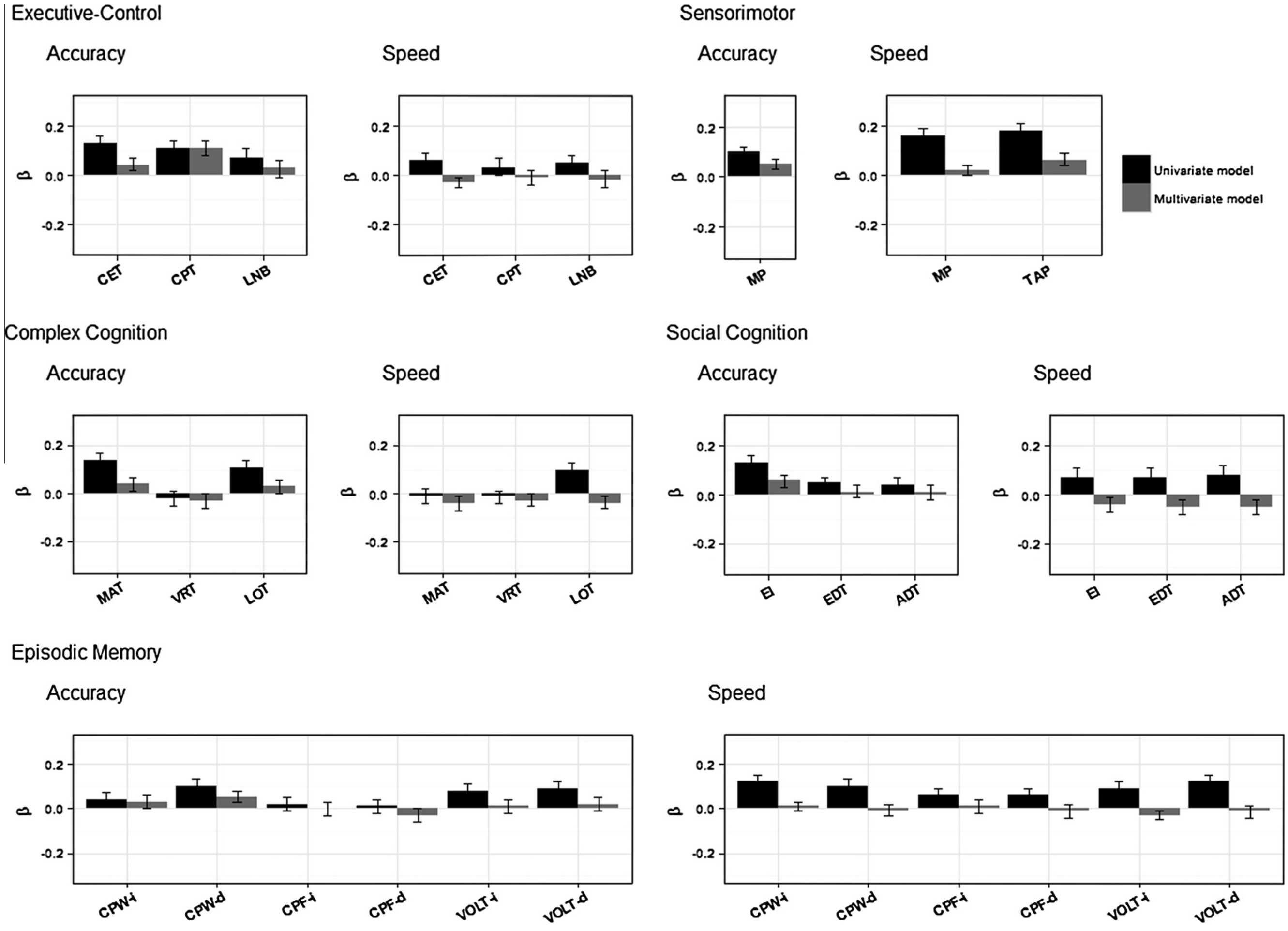


Fig. 4. Standardized associations (β) between physical activity (in weekly METhours) and cognitive accuracy and speed across the cognitive domains, including 95% confidence intervals. The effect sizes (ES) of weekly METhours in the univariate model (black bars) are generally positive, but vary across cognitive domains. After taking into account confounding effects of sex and age in the multivariate model, effect sizes of direct associations (gray bars) are small and centered around 0. See Table 1 for cognitive domain and full name of the cognitive tests (CET, CPT, LNB, CPW-i, CPW-d, CPF-i, CPF-d, VOLT-i, VOLT-d, MAT, VRT, LOT, EI, EDT, ADT, MP, TAP).

idea of beneficial effects of regular leisure time exercise on cognitive accuracy and speed, however, this interpretation requires some caution. First, in line with results from reviews, the majority of the associations between exercise behavior and cognitive measures were positive, but associations varied in strength; null effects, including ones in the negative direction, were also found. This pattern thus reiterates the heterogeneous findings in the literature and implies that not all cognitive functions may benefit equally from voluntary exercise.

Second, our analyses clearly demonstrate the presence of confounding effects. Sex differences were established in both exercise behavior and cognitive performance and these varied in sign and strength across cognitive domains. Exercise behavior and cognitive performance decreased with age, also replicating previous findings (de Moor, Beem, Stubbe, Boomsma, & de Geus, 2006). The linear and quadratic associations with age were found to vary considerably across cognitive domains. Effects of age, sex, and exercise on cognitive performance can thus be confounded, while the magnitude of the confounding effect is dependent on the specific cognitive domain.

After regressing out sex and age, and while using a liberal significance level of $\alpha = 0.05$, only four out of 33 relationships between weekly METhours and cognitive performance would reach the level of significance. This is close to the number of expected false positives. Publication bias in previously reported results is thus a serious issue. When proper correction for multiple testing is applied, the association between weekly METhours and cognitive performance may not survive statistical scrutiny. With exception of the Continuous Performance Test (CPT), none of the standardized regression coefficients in our study was above 0.1 (or below -0.1), therefore any effects must be considered small.

Combining previous and present results, we conclude that only the association between chronic, regular exercise behavior and attentional performance inspires some confidence. Accuracy on the CPT, a widely used neuropsychological test that measures a person's sustained and selective attention, showed the strongest association with voluntary exercise behavior. Multiple clinical studies that explored exercise as a possible treatment option for children with attention deficit/hyperactivity disorder (ADHD) have provided support for a beneficial exercise effect on the ability to focus on relevant stimuli and ignore competing stimuli (Berwid & Halperin, 2012; Pontifex, Saliba, Raine, Picchietti, & Hillman, 2013; Wigal, Emmerson, Gehricke, & Galassetti, 2013). High intensity physical activity in ADHD children may improve their continuous performance test score, for example, irrespective of the effect of the often prescribed drug methylphenidate (Medina et al., 2010). Such clinical studies demonstrate the importance of acknowledging that our results concerning voluntary exercise should not be taken as precluding beneficial effects of exercise on cognition in specific settings. As mentioned in the introduction, study design has been found to be a major source of heterogeneity among previous results as reported in the existent literature (Singh et al., 2012). Experimental studies in which effects of exercise on cognition can be attributed to intervention or treatment have shown larger associations than observational studies in which exercise-related differences between participants may be drowned out by the many other sources of individual differences in cognitive ability, including genetic factors. In part, this may reflect non-specific effects of the participation in an exercise regime; Barnes et al. (2013) found that in a sample of participants with nonclinical cognitive complaints, each of four groups (control and intervention conditions of mental and physical activity) showed increased global cognitive function. Relatively large associations obtained in experimental studies may also be due to the fact that interventions were performed in vulnerable populations, where exercise may truly have relatively large effects. Elderly with cognitive complains or stroke have shown to benefit substantially from exercise (Barnes et al., 2013; Marzolini, Oh, McIlroy, & Brooks, 2012). Here exercise may protect against brain atrophy, increase brain connectivity, or protect against white matter damage caused by heavy alcohol consumption (Karoly et al., 2013).

Despite inconsistent findings in humans, the effectiveness of exercise has been shown more consistently in animal studies, which have suggested insight into the mechanisms involved in the beneficiary effects of exercise (Lista & Sorrentino, 2010). How these processes translate to human cognition has mainly been discussed in the light of cognitive ageing: various plausible pathways have been hypothesized to explain the effects of exercise on cognitive functioning and aging processes. Exercise effects may act through a diverse set of (supra)molecular mechanisms such as angiogenesis due to increased blood flow, neurogenesis and synaptogenesis (both consistently shown in the hippocampus, involved in learning and memory). These mechanisms are controlled by processes that have also been directly associated with exercise: through for example brain derived neurotrophic factor (BNDF), growth factors, neurotransmitters (including glutamate, serotonin, noradrenaline, dopamine and acetylcholine), hormones and second messenger systems (Fabel & Kempermann, 2008; Lista & Sorrentino, 2010; van Praag, 2008). In addition, neuroimaging studies in humans have shown that exercise may induce structural changes in the hippocampus and the frontal and parietal cortex (Erickson et al., 2009, 2010), as well as functional changes (Colcombe et al., 2004; Voss et al., 2010).

We end by stressing that the question of immediate importance to policy makers should not be the question whether there are associations between exercise behavior and cognitive functioning, but rather whether and how changes in exercise behavior relate to changes in cognitive functioning. Furthering the knowledge about the sources of heterogeneity in the results may be viewed as a first step. In view of the present findings, we suggest that further exploration of the association between changes in voluntary regular leisure time exercise behavior and changes in cognitive functioning is needed. Such explorations should be carried out in homogeneous samples using valid and reliable measures of exercise behavior or other forms of physical activity and neurocognitive functioning. Apart from the use of valid, reliable instruments to measure physical activity, we advance the use of strong research designs – e.g., experimental, longitudinal, or genetically informative (e.g., de Moor et al., 2008) designs – because these are essential to address the crucial question whether physical activity truly is a causal means to improve cognition.

Funding

This work was supported by the Netherlands Consortium for Healthy Ageing (NCHA), the Netherlands Organization for Scientific Research (NWO, 433-09-220), European Research Council (ERC 230374) and the Neuroscience Campus Amsterdam (NCA).

Acknowledgments

We thank all the participants.

References

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Jr., Montoye, H. J., Sallis, J. F., et al. (1993). Compendium of physical activities: Classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise, 25*, 71–80.

Almasy, L., Gur, R. C., Haack, K., Cole, S. A., Calkins, M. E., Peralta, J. M., et al. (2008). A genome screen for quantitative trait loci influencing schizophrenia and neurocognitive phenotypes. *American Journal of Psychiatry, 165*, 1185–1192.

Barnes, D. E., Santos-Modesitt, W., Poelke, G., Kramer, A. F., Castro, C., Middleton, L. E., et al. (2013). The Mental Activity and eXercise (MAX) Trial: A randomized controlled trial to enhance cognitive function in older adults. *JAMA Journal of Internal Medicine, 173*, 797–804.

Berwid, O. G., & Halperin, J. M. (2012). Emerging support for a role of exercise in attention-deficit/hyperactivity disorder intervention planning. *Current Psychiatry Reports, 14*, 543–551.

Boomsma, D. I., de Geus, E. J., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., et al. (2006). Netherlands Twin Register: From twins to twin families. *Twin Research and Human Genetics, 9*, 849–857.

Churchill, J. D., Galvez, R., Colcombe, S., Swain, R. A., Kramer, A. F., & Greenough, W. T. (2002). Exercise, experience and the aging brain. *Neurobiology of Aging, 23*, 941–955.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14*, 125–130.

Colcombe, S. J., Kramer, A. F., Erickson, K. I., Scalf, P., McAuley, E., Cohen, N. J., et al. (2004). Cardiovascular fitness, cortical plasticity, and aging. *Proceedings of the National academy of Sciences of the United States of America, 101*, 3316–3321.

de Moor, M. H. M., Beem, A. L., Stubbe, J. H., Boomsma, D. I., & de Geus, E. J. C. (2006). Regular exercise, anxiety, depression and personality: A population-based study. *Preventive Medicine, 42*, 273–279.

de Moor, M. H. M., Boomsma, D. I., Stubbe, J. H., Willemsen, G., & de Geus, E. J. C. (2008). Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Archives of General Psychiatry, 65*, 897–905.

Erickson, K. I., Prakash, R. S., Voss, M. W., Chaddock, L., Hu, L., Morris, K. S., et al. (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus, 19*, 1030–1039.

Erickson, K. I., Raji, C. A., Lopez, O. L., Becker, J. T., Rosano, C., Newman, A. B., et al. (2010). Physical activity predicts gray matter volume in late adulthood The Cardiovascular Health Study. *Neurology, 75*, 1415–1422.

Fabel, K., & Kempermann, G. (2008). Physical activity and the regulation of neurogenesis in the adult and aging brain. *Neuromolecular Medicine, 10*(2), 59–66.

Fedewa, A. L., & Ahn, S. (2011). The effects of physical activity and physical fitness on children's achievement and cognitive outcomes: A meta-analysis. *Research Quarterly for Exercise and Sport, 82*, 521–535.

Greenwood, P. M., & Parasuraman, R. (2010). Neuronal and cognitive plasticity: A neurocognitive framework for ameliorating cognitive aging. *Frontiers in Aging Neuroscience, 2*, 150.

Gur, R. C., Ragland, J. D., Moberg, P. J., Bilker, W. B., Kohler, C., Siegel, S. J., et al. (2001). Computerized neurocognitive scanning: II. The profile of schizophrenia. *Neuropsychopharmacology, 25*, 777–788.

Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., et al. (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology, 25*, 766–776.

Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., et al. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology, 26*, 251–265.

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., et al. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods, 187*, 254–262.

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

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

Hindin, S. B., & Zelinski, E. M. (2012). Extended practice and aerobic exercise interventions benefit untrained cognitive outcomes in older adults: A meta-analysis. *Journal of the American Geriatrics Society, 60*, 136–141.

IBM Corp. (2011). *IBM SPSS Statistics for Windows, Version 21.0*. Armonk, NY: IBM Corp.

Kahn, E. B., Ramsey, L. T., Brownson, R. C., Heath, G. W., Howze, E. H., Powell, K. E., et al. (2002). The effectiveness of interventions to increase physical activity – A systematic review. *American Journal of Preventive Medicine, 22*, 73–108.

Karoly, H. C., Stevens, C. J., Thayer, R. E., Magnan, R. E., Bryan, A. D., & Hutchison, K. E. (2013). Aerobic exercise moderates the effect of heavy alcohol consumption on white matter damage. *Alcohol-Clinical and Experimental Research, 37*, 1508–1515.

Kraft, E. (2012). Cognitive function, physical activity, and aging: Possible biological links and implications for multimodal interventions. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 19*, 248–263.

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

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity, 95*, 221–227.

Lista, I., & Sorrentino, G. (2010). Biological mechanisms of physical activity in preventing cognitive decline. *Cellular and Molecular Neurobiology, 30*, 493–503.

Marzolini, S., Oh, P., McIlroy, W., & Brooks, D. (2012). The effects of an aerobic and resistance exercise training program on cognition following stroke. *Neurorehabilitation and Neural Repair, 27*, 392–402.

Medina, J. A., Netto, T. L., Muszkat, M., Medina, A. C., Botter, D., Orbetelli, R., et al. (2010). Exercise impact on sustained attention of ADHD children, methylphenidate effects. *Attention Deficit and Hyperactivity Disorders, 2*, 49–58.

Ploughman, M. (2008). Exercise is brain food: The effects of physical activity on cognitive function. *Developmental Neurorehabilitation, 11*, 236–240.

Pontifex, M. B., Saliba, B. J., Raine, L. B., Picchietti, D. L., & Hillman, C. H. (2013). Exercise improves behavioral, neurocognitive, and scholastic performance in children with attention-deficit/hyperactivity disorder. *Journal of Pediatrics, 162*, 543–551.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 5*, 56.

Ridley, K., Ainsworth, B. E., & Olds, T. S. (2008). Development of a compendium of energy expenditures for youth. *International Journal of Behavioral Nutrition and Physical Activity, 5*, 45.

Roalf, D. R., Ruparel, K., Gur, R. E., Bilker, W., Gerraty, R., Elliott, M. A., et al. (2014). Neuroimaging predictors of cognitive performance across a standardized neurocognitive battery. *Neuropsychology, 28*(2), 161–176.

Rosseel, Y. (2012). Lavaan: An R Package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36.

Singh, A., Uijtdewilligen, L., Twisk, J. W., van Mechelen, W., & Chinapaw, M. J. (2012). Physical activity and performance at school: A systematic review of the literature including a methodological quality assessment. *Archives of Pediatric & Adolescent Medicine, 166*, 49–55.

Stubbe, J. H., Boomsma, D. I., & de Geus, E. J. C. (2005). Sports participation during adolescence: A shift from environmental to genetic factors. *Medicine and Science in Sports and Exercise, 37*, 563–570.

Stubbe, J. H., de Moor, M. H. M., Boomsma, D. I., & de Geus, E. J. C. (2007). The association between exercise participation and well-being: A co-twin study. *Preventive Medicine, 44*, 148–152.

Taras, H. (2005). Physical activity and student performance at school. *Journal of School Health, 75*, 214–218.

Tomporowski, P. D., Davis, C. L., Miller, P. H., & Naglieri, J. A. (2008). Exercise and children's intelligence, cognition, and academic achievement. *Educational Psychology Review, 20*, 111–131.

Trudeau, F., & Shephard, R. J. (2008). Physical education, school physical activity, school sports and academic performance. *International Journal of Behavioral Nutrition and Physical Activity, 25*, 5–10.

van Beijsterveldt, C. E. M., Groen-Blokhuis, M., Hottenga, J. J., Franic, S., Hudziak, J. J., Lamb, D., et al. (2013). The Young Netherlands Twin Register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Research and Human Genetics, 16*, 252–267.

van Praag, H. (2008). Neurogenesis and exercise: Past and future directions. *NeuroMolecular Medicine, 10*, 128–140.

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

Voss, M. W., Prakash, R. S., Erickson, K. I., Basak, C., Chaddock, L., Kim, J. S., et al. (2010). Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. *Frontiers of Aging Neuroscience*, 2.

Wigal, S. B., Emmerson, N., Gehricke, J. G., & Galassetti, P. (2013). Exercise: Applications to childhood ADHD. *Journal of Attention Disorders, 17*, 279–290.

Willemsen, G., Vink, J. M., Abdellaoui, A., den, B. A., van Beek, J. H., Draisma, H. H., et al. (2013). The adult Netherlands Twin Register: Twenty-five years of survey and biological data collection. *Twin Research and Human Genetics, 16*, 271–281.

Yuki, A., Lee, S., Kim, H., Kozakai, R., Ando, F., & Shimokata, H. (2012). Relationship between physical activity and brain atrophy progression. *Medicine and Science in Sports and Exercise, 44*, 2362–2368.