



## PAPER

# The sexual dimorphic association of cardiorespiratory fitness to working memory in children

Eric S. Drollette,<sup>1</sup> Mark R. Scudder,<sup>1</sup> Lauren B. Raine,<sup>1</sup> R. Davis Moore,<sup>1</sup> Matthew B. Pontifex,<sup>2</sup> Kirk I. Erickson<sup>3,4</sup> and Charles H. Hillman<sup>1</sup>

1. Department of Kinesiology and Community Health, University of Illinois at Urbana-Champaign, USA

2. Department of Kinesiology, Michigan State University, USA

3. Department of Psychology, University of Pittsburgh, USA

4. The Center for the Neural Basis of Cognition, University of Pittsburgh, USA

## Abstract

The present investigation examined the sexual dimorphic patterns of cardiorespiratory fitness to working memory in preadolescent children (age range: 7.7–10.9). Data were collected in three separate studies (Study 1:  $n = 97$ , 42 females; Study 2:  $n = 95$ , 45 females; Study 3:  $n = 84$ , 37 females). All participants completed a cardiorespiratory fitness assessment in addition to a specific measure of working memory (i.e. the operation span task, the  $n$ -back task, or the Sternberg task). Results from all three samples revealed that higher cardiorespiratory fitness levels were associated with better working memory performance only for males with no such relation observed for females. In addition, the sexually dimorphic pattern was selective for the most challenging working memory conditions in each task. Together, these findings reveal new evidence that cardiorespiratory fitness is selectively related to better working memory performance for male children. This investigation provides additional insight into how interventions aimed at improving fitness may influence cognitive development differentially among preadolescent children.

## Research highlights

- The relation of fitness to working memory is selective for prepubertal male children with no such relation observed for females.
- For males, the fitness–working memory relationship is generalized across several constructs of working memory, yet selective for conditions that placed greater demand on working memory.
- Only at higher fitness levels do males demonstrate greater working memory performance compared to females, revealing that sexual dimorphic patterns of cardiorespiratory health differentially influence working memory performance.
- These sex-related differences provide a link for future research aimed at elucidating the benefits of fitness on mechanisms of brain health and cognition.

## Introduction

Physical inactivity has been described as a global health pandemic (Kohl, Craig, Lambert, Inove, Alkandari *et al.*, 2012) and has been identified as the fourth leading risk factor for non-communicable diseases (World Health Organization [WHO], 2009). Perhaps more alarming is evidence indicating that this trend is increasing among school-age children (Hallal, Andersen, Bull, Guthold, Haskell *et al.*, 2012), with more youth being diagnosed with type-II diabetes and obesity as inactivity rates continue to rise (Eisenmann, 2003). Such trends are likely exacerbated by reductions in opportunities for physical activity during the school day (e.g. more than 44% of school districts in the US have reported reductions; Andersen, Crespo, Bartlett, Cheskin & Pratt, 1998; Centers for Education Policy, 2007; Institute of Medicine of the National Academies [IOM], 2013; Sisson, Church,

Address for correspondence: Charles H. Hillman, Department of Kinesiology & Community Health, 317 Louise Freer Hall, 906 South Goodwin Avenue, University of Illinois, Urbana, IL 61801, USA; e-mail: chhillma@illinois.edu

Martin, Tudor-Locke, Smith *et al.*, 2009). Thus, it is not surprising that various organizations have begun to focus on school-based interventions for children with an emphasis not only on appropriate energy balance for obesity prevention, but also on the cognitive and brain health benefits associated with improvements in cardiorespiratory fitness (see Hillman, Erickson & Kramer, 2008, for review; IOM, 2013; National Association for Sport and Physical Education, 2008).

To date, the majority of research is consonant in demonstrating a positive relationship between fitness, academic achievement (Castelli, Hillman, Buck & Erwin, 2007; Chomitz, Slining, McGowan, Mitchell, Dawson *et al.*, 2009; Donnelly, Greene, Gibson, Smith, Washburn *et al.*, 2009), and aspects of cognitive control that support scholastic performance (Diamond, Barnett, Thomas & Munro, 2007) including working memory (Kamijo, Pontifex, O'Leary, Scudder, Wu *et al.*, 2011; Scudder, Lambourne, Drollette, Herrmann, Washburn *et al.*, 2014). Cognitive control refers to top-down, goal-directed operations that assist with selection, scheduling, maintaining, and coordinating processes that underlie perception, memory, and action (Norman & Shallice, 1986; Rogers & Monsell, 1995). Working memory is a key component of cognitive control that represents a hierarchical system involving the short-term, transitory storage and manipulation of information in the service of motivated behavior (Baddeley & Hitch, 1974; Cowan, 1995, 1999; Kane & Engle, 2002). Prior evidence elucidating the relationship between working memory and fitness has been derived from cross-sectional research. For example, Scudder *et al.* (2014) found that a field test of cardiorespiratory fitness had a beneficial relation with working memory performance using a spatial *n*-back task in a cohort of 397 preadolescent children. Interestingly, the relation was strengthened during task conditions requiring greater amounts of working memory (i.e. 2-back trials).

This and other investigations (Kamijo *et al.*, 2011; Scudder *et al.*, 2014) have provided emerging evidence on the relation of higher fitness with enhanced working memory performance among children while minimizing potential bias by matching groups based on individual differences (e.g. age, socioeconomic status, pubertal timing, IQ, sex); however, few studies have systematically explored fitness-related differences in children's working memory among these key demographic characteristics. Specifically, a separate line of research has demonstrated a unique contribution of sex to various cognitive abilities including working memory. Across the lifespan, research findings demonstrate a male advantage on select spatial and working memory-related tasks (Astur, Tropp, Sava, Constable & Markus, 2004; Kaufman, 2007; Levine, Huttenlocher, Taylor & Langrock, 1999; Masters &

Sanders, 1993; Moore & Johnson, 2008; Parsons, Larson, Kratz, Thiebaut, Bluestein *et al.*, 2004; Peters, Laeng, Latham, Jackson, Zaiyouna *et al.*, 1995; Quinn & Liben, 2008; Vecchi & Girelli, 1998), including: the mental rotation task (Linn & Petersen, 1985; Shepard & Metzler, 1971; Vandenberg & Kuse, 1978), the spatial digit span task (Geiger & Litwiller, 2005), the spatial transformation task (Levine *et al.*, 1999), working memory span tasks (Kaufman, 2007), and object location memory tasks (Postma, Jager, Kessels, Koppeschaar & Van Honk, 2004).

These working memory differences have been attributed to a variety of developmental variations between males and females including differential rate of brain development and function. Specifically, research has demonstrated sexually dimorphic patterns in the latency and duration of brain maturation (Blakemore, Burnett & Dahl, 2010; Colom, Stein, Rajagopalan, Martínez, Hermel *et al.*, 2013; Giedd, Vaituzis, Hamburger, Lange, Rajapakse *et al.*, 1996), with such differences observed in cortical and sub-cortical structures associated with cognitive control processes (Bramen, Hranilovich, Dahl, Forbes, Chen *et al.*, 2011; Christakou, Halari, Smith, Ifkovits, Brammer *et al.*, 2009; Davies, Segalowitz & Gavin, 2004; De Bellis, Keshavan, Beers, Hall, Frustaci *et al.*, 2001; Rubia, Hyde, Halari, Giampietro & Smith, 2010) including regions associated with working memory (Colom *et al.*, 2013; D'Esposito, Detre, Alsop, Shin, Atlas *et al.*, 1995). Such regions include the prefrontal cortex (D'Esposito, Postle & Rypma, 2000; Diamond, 2002; Owen, Evans & Petrides, 1996; Petrides, 1989), the anterior cingulate cortex (ACC; Osaka, Osaka, Kondo, Morishita, Fukuyama *et al.*, 2003; Smith & Jonides, 1999), and the hippocampus (Axmacher, Mormann, Fernández, Cohen, Elger *et al.*, 2007; Axmacher, Henseler, Jensen, Weinreich, Elger *et al.*, 2010; Chein, Moore & Conway, 2011; Colom *et al.*, 2013; Faraco, Unsworth, Langley, Terry, Li *et al.*, 2011; Hannula, Tranel & Cohen, 2006; Oztekin, McElree, Staresina & Davachi, 2009; van Vugt, Schulze-Bonhage, Litt, Brandt & Kahana, 2010; Warren, Duff, Tranel & Cohen, 2011). For example, developmental neuroimaging studies demonstrate that during cognitive control tasks (i.e. switching, inhibition, and working memory) neural activation (i.e. fMRI) increases with age to a greater degree in frontal regions for females and temporal-parietal regions for males (Bell, Willson, Wilman, Dave & Silverstone, 2006; Christakou *et al.*, 2009; De Bellis *et al.*, 2001; Goldstein, Jerram, Poldrack, Anagnoson, Breiter *et al.*, 2005; Rubia *et al.*, 2010; Thomsen, Hugdahl, Ersland, Barndon, Lundervold *et al.*, 2000; Weiss, Siedentopf, Hofer, Deisenhammer, Hoptman *et al.*, 2003). Furthermore, total volume of the female hippocampus during the early stages of development gradually increases through adolescence and results in greater cell density compared to

the male hippocampus (Filipek, Richelme, Kennedy & Caviness, 1994; Giedd *et al.*, 1996; Murphy, DeCarli, McIntosh, Daly, Mentis *et al.*, 1996). By adulthood the hippocampal formation appears lateralized with the right formation being larger (Colom *et al.*, 2013) and more active (fMRI; Frings, Wagner, Unterrainer, Spreer, Halsband *et al.*, 2006) than the left in males compared to females. Given that neural activation associated with working memory depends on such inter-related cortical and sub-cortical regions, it is interesting that these hemispheric lateralization effects between males and females have also been observed while participants performed working memory tasks (Hugdahl, Thomsen & Erslund, 2006; Speck, Ernst, Braun, Koch, Miller *et al.*, 2000). Taken together, these sexually dimorphic differences in brain development suggest a neural substrate for sex differences in working memory performance.

Further, these regions associated with working memory are also amenable to changes in cardiorespiratory fitness (Chaddock, Erickson, Prakash, Kim, Voss *et al.*, 2010; Colcombe, Erickson, Raz, Webb, Cohen *et al.*, 2003; Colcombe, Kramer, Erickson, Scalf, McAuley *et al.*, 2004; Colcombe, Erickson, Scalf, Kim, Prakash *et al.*, 2006; Erickson, Prakash, Voss, Chaddock, Hu *et al.*, 2009; Erickson, Voss, Prakash, Basak, Szabo *et al.*, 2011; Kramer & Erickson, 2007). For example, research has demonstrated larger bi-lateral hippocampal volume (along with superior relational memory performance) in higher-fit compared to lower-fit children (Chaddock *et al.*, 2010), and in older adults who participated in an exercise intervention compared to control participants (Erickson *et al.*, 2011). Such volumetric fluctuations in cortical and sub-cortical regions are typical especially among developing populations, and possibly represent neuronal proliferation, synaptic pruning, and myelination (Eriksson, Perfilieva, Björk-Eriksson, Alborn, Nordborg *et al.*, 1998; Kornack & Rakic, 1999); processes which facilitate the growth, development, and efficiency of cognitive processing. Therefore, it is interesting that fitness is not only associated with improved working memory performance, but also with greater volume in brain structures supporting working memory.

Taken together, it is surprising that previous research has left the contribution of fitness on working memory performance between males and females unexplored given these unique sexually dimorphic differences in behavior outcomes and underlying neural mechanisms associated with working memory. To our knowledge, no investigations have explored the combination of these research threads; however, previous research has provided some evidence for sexually dimorphic effects of fitness on scholastic performance (Grissom, 2005; Kwak, Kremers, Bergman, Ruiz, Rizzo *et al.*, 2009; Sigfúsdóttir,

Kristjánsson & Allegrante, 2007) with higher fitness associated with better academic achievement only in males (Kwak *et al.*, 2009). Thus, given the unique relation of working memory with various academic outcomes (Gathercole & Pickering, 2000; Geary, Hoard, Byrd-Craven & DeSoto, 2004), the aim of the present study was to extend these findings and examine whether fitness differentially modulates sex differences in working memory. In accordance with prior work (Kamijo *et al.*, 2011; Scudder *et al.*, 2014) we hypothesized a relationship between higher fitness and better working memory performance, with selectively greater associations for task conditions that placed increased demand on working memory. Further, based on the academic achievement literature (Kwak *et al.*, 2009; Sigfúsdóttir *et al.*, 2007), it was predicted that such associations would manifest more strongly for males relative to females.

## Study 1

The operation span task (OSPAN; Turner & Engle, 1989) is a working memory span (WMS) task (Conway, Kane, Bunting, Hambrick, Wilhelm *et al.*, 2005) which taps constructs including maintenance of memory representations and controlled attention under conditions of distraction or interference (see Engle & Kane, 2004, and Conway *et al.*, 2005, for reviews). Such WMS tasks have demonstrated important individual differences related to sex. For example, research has indicated that adult males outperform females on various WMS tasks including the OSPAN (Kaufman, 2007). To our knowledge, the OSPAN has not been used in prior studies examining the relation between fitness and working memory performance in children and thus provides new insight into the differential relation of fitness and working memory between males and females.

## Methods

### Participants

One hundred and seventeen (56 females) preadolescent children were recruited from the East-Central Illinois area. Participants were excluded for incomplete demographic data ( $n = 8$ ; e.g. socioeconomic status, puberty timing, grade, etc.), task performance below 50% ( $n = 6$ ), and failure to achieve maximal aerobic capacity criterion during the cardiorespiratory fitness assessment ( $n = 6$ ; see Cardiorespiratory Fitness Assessment methods). Thus, analyses were conducted on the remaining 97 (42 females) preadolescent children (age range: 7.9–10.9 years). All

participants provided written assent and their legal guardians provided written informed consent in accordance with the Institutional Review Board (IRB) of the University of Illinois at Urbana-Champaign. Prior to testing, legal guardians completed a battery of health history and demographic questionnaires on behalf of the participant. Based on these questionnaires all participants included in the study had normal to corrected normal vision, indicated English as their primary language, reported no physical disabilities that could be exacerbated by exercise participation (Physical Activity Readiness Questionnaire [PAR-Q]; Thomas, Reading & Shephard, 1992), reported absence of any doctor-diagnosed neurological or attentional disorders, and were prepubescent (i.e. a score  $\leq 2$  on a 5-point scale) at the time of testing based on the results of the Tanner Staging System (Tanner, 1962; Taylor, Whincup, Hindmarsh, Lampe, Odoki *et al.*, 2001). In addition, socioeconomic status (SES) was assessed and calculated using a trichotomous index based on: (1) highest level of education obtained by the mother and father, (2) number of parents who worked full time, and (3) participation in a free or reduced-price lunch program at school (Birnbau, Lytle, Murray, Story, Perry *et al.*, 2002). Body mass index (BMI) was calculated as weight divided by height (i.e. kg/m<sup>2</sup>). Lastly, an estimate of intelligence quotient (IQ) was collected based on an age-normed standardized IQ exam<sup>1</sup> administered by a trained experimenter.

#### OSPAN task

Stimuli were presented focally on a computer screen at a distance of 1 meter using Neuroscan Stim<sup>2</sup> software (Compumedics, Charlotte, NC). A single word was presented followed by an arithmetic problem; both of which were required to be read aloud, and together constituted one trial. After reading the arithmetic

problem aloud, participants were instructed to respond as accurately as possible with a left thumb press (using a response pad) if the solution was incorrect, or a right thumb press if the solution was correct. Once the to-be-remembered word was read aloud and the math operation was solved, the next trial was presented. Following the final trial of each set a prompt appeared instructing the participant to recall and write down remembered words (in order of presentation). This recall phase signified the end of a set, and each set varied between one and four trials.

All participants completed four blocks that each contained four sets, one of each set size (1, 2, 3, and 4 trials presented in random order), for a total of 40 word-operation trials or 16 sets. To ensure that math computation successfully prevented mental rehearsal, a 50% accuracy inclusion criterion was set. Although prior research recommends that a higher math accuracy cut-off be used for inclusion criterion (Conway *et al.*, 2005), such a criterion applied to the present study only eliminated females from the sample, thus biasing the results. Therefore, a chance criterion (i.e. 50%) was adopted to control potential bias related to sex differences in arithmetic ability.<sup>2</sup> For practice, all participants received one block of set sizes 2, 3, and 4 (nine trials total). Words were presented for a duration of 1000 ms with an inter-stimulus interval (ISI) between words and arithmetic problems set at 1100 ms. Arithmetic problems were presented for up to 10 seconds or until a response was made. Word recall OSPAN accuracy was calculated as total trials in a correctly recalled set (e.g. 2 correctly recalled words in a set of 2 = score of 2; 3 correctly recalled words in a set of 4 = score of 0), and recalled in the correct order, divided by total trials possible (i.e. 40). (see Conway *et al.*, 2005, for calculation method).

#### Cardiorespiratory fitness assessment

Maximal aerobic capacity (VO<sub>2max</sub>; expressed in ml/kg/min) was measured on a motor-driven treadmill following a modified Balke protocol (American College of Sports Medicine [ACSM], 2010). Prior to testing, all participants had their height and weight measured and were fitted with a Polar heart rate (HR) monitor (Polar WearLink<sup>®</sup>+ 31, Polar Electro, Finland) to measure HR throughout the test. Following a warm-up period, the treadmill was set to a constant speed during the remainder of the test, while grade increments of 2.5% occurred every 2 minutes until volitional exhaustion. Oxygen consumption was measured using a computerized indirect calorimetry system (ParvoMedics True Max 2400) with averages for oxygen uptake (VO<sub>2</sub>) and respiratory exchange ratio (RER) assessed every

<sup>1</sup> Intelligent quotient (IQ) was assessed using either the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990), the Kaufman Brief Intelligence Test II (KBIT-II; Kaufman & Kaufman, 2004), or the Woodcock-Johnson III (WJ-III; Woodcock, McGrew & Mather, 2001).

<sup>2</sup> Given the low criterion for math accuracy, planned correlation analysis was performed between measures of math accuracy and OSPAN recall scores to determine the potential of a performance trade-off. If a trade-off occurred between processing (i.e. solving math solutions) and storage (i.e. maintaining word sets) we would expect participants to perform poorly on math in order to facilitate word recall performance, thus resulting in a negative correlation. However, results across all OSPAN measures revealed a positive correlation,  $r(97) \geq .47$ ,  $p \leq .01$ , suggesting that increases in math scores are related to increases in OSPAN recall scores. As such, the demands of the processing and storage components of the OSPAN appear adequate to prevent potential trade-off strategies among low math performers.



20 seconds. Concurrently, ratings of perceived exertion (RPE) were measured every 2 minutes using the children's OMNI scale (Utter, Robertson, Nieman & Kang, 2002).  $VO_{2max}$  was defined as the highest oxygen consumption corresponding to a plateau in oxygen uptake (i.e. an increase of less than 2 ml/kg/min despite an increase in exercise workload) or a minimum of 2 of the following criteria: (1) a peak heart rate  $\geq 185$  bpm (ACSM, 2010) and a heart rate plateau (Freedson & Goodman, 1993); (2)  $RER \geq 1.0$  (Bar-Or, 1983); and/or (3) ratings on the children's OMNI scale of perceived exertion  $\geq 8$  (Utter *et al.*, 2002). Participants were excluded if they did not achieve a plateau or any combination of two of the remaining criteria.

### Procedure

All participants completed preliminary screening, along with cognitive and cardiorespiratory fitness assessments, in a single session in the laboratory. In preparation for testing, participants were instructed to avoid moderate to vigorous physical activity and to limit foods and beverages containing caffeine (12 hours prior to testing) while otherwise maintaining a normal daily routine. After completing informed assent/consent, a brief IQ exam was administered by a trained laboratory staff member while parents completed health history and demographics questionnaires, as well as other prescreening measures described previously. Cognitive testing took place in a separate room while participants were seated in a quiet testing chamber. All participants were instructed on appropriate completion of the working memory task and given practice trials prior to testing. Following cognitive testing, participants completed a graded exercise test on a motor-driven treadmill to assess cardiorespiratory fitness level. Fitness assessment was always performed at the end of either session to avoid any confounding effects related to acute physical activity on cognitive task performance (Drollette, Scudder, Raine, Moore, Saliba *et al.*, 2014; Drollette, Shishido, Pontifex & Hillman, 2012; Hillman, Pontifex, Raine, Castelli, Hall *et al.*, 2009; Pontifex, Saliba, Raine, Picchietti & Hillman, 2013).

### Statistical analysis

Data were analyzed using SPSS (SPSS v. 22, Chicago, IL) with the family-wise alpha threshold for all tests set at  $p = .05$ . Demographic, fitness, and behavioral differences between males and females were evaluated using independent *t*-tests. Bivariate correlations were conducted utilizing Pearson product-moment correlation coefficients. Collective (i.e. all participants) and separate

correlation analysis (i.e. groups separated by sex) were conducted between working memory measures, age, grade, SES, IQ, BMI, and cardiorespiratory fitness (i.e.  $VO_{2max}$ ). Further, primary linear regression analyses were performed with demographic factors that were significantly correlated with working memory performance entered as covariates, sex (coded: 0 = females, 1 = males), fitness, and a two-way interaction of sex  $\times$  fitness. If the interaction product of sex  $\times$  fitness reached significance (i.e. suggesting sexual dimorphic patterns of cardiorespiratory fitness on working memory) further decomposition was performed using hierarchical linear regression analyses separately for male and female groups. This was accomplished by regressing working memory measures with statistically significant demographic correlates in Step 1 and fitness in Step 2. The change in the *R*-square value between Step 1 and Step 2 was used to determine whether the independent contribution of fitness explained a significant portion of variance in working memory. Based on the *a priori* hypotheses, only regression analyses revealing an effect of fitness are reported.

Further, given the small sample size (when parsing groups by sex) an additional bootstrapping procedure (Efron & Tibshirani, 1993; Shrout & Bolger, 2002) was utilized to increase confidence in the reliability of the regression results. This was accomplished by re-sampling with replacement 500 samples derived from the original sample. Thus, 95% confidence intervals of unstandardized coefficients derived from bootstrap analysis are included in the regression results. In addition, simple slope analysis (Aiken & West, 1991; Jaccard & Turrisi, 2003) was performed to further decompose the interaction of sex  $\times$  fitness. The purpose of this analysis was to examine the main effect of sex at higher and lower levels of cardiorespiratory fitness. This was accomplished by: (1) standardizing (i.e. *z*-score) dependent and predictor variables, (2) re-centering fitness at one standard deviation above and below the mean, and (3) performing the primary linear regression analysis with the re-centered fitness values.

### Results

Means and standard error of the mean (SEM) are reported for demographic, fitness, and OSPAN behavioral variables in Table 1. As expected, fitness levels were higher for males ( $42.8 \pm 1.12$  ml/kg/min) compared to females ( $39.4 \pm 1.03$  ml/kg/min),  $t(95) = 2.19$ ,  $p \leq .05$  (see Sallis, 1993, for a review on cardiorespiratory fitness and sex differences). However, males (23.5–60.8 ml/kg/min) and females (23.1–59.8 ml/kg/min) had a similar range of fitness levels. Lastly, males ( $91.6 \pm 1.05\%$ )

**Table 1** Study 1: Mean  $\pm$  1 SEM values for demographic, fitness, and OSPAN behavioral measures

Measure	Combined	Male	Female
<i>n</i>	97	55	42
Age (years)	9.4 $\pm$ 0.09	9.5 $\pm$ 0.11	9.3 $\pm$ 0.14
Grade	3.6 $\pm$ 0.11	3.6 $\pm$ 0.16	3.6 $\pm$ 0.15
SES	2.2 $\pm$ 0.08	2.1 $\pm$ 0.11	2.3 $\pm$ 0.12
IQ	112.8 $\pm$ 1.22	111.6 $\pm$ 1.60	114.5 $\pm$ 1.90
BMI (kg/m <sup>2</sup> )	18.5 $\pm$ 0.39	18.3 $\pm$ 0.52	18.7 $\pm$ 0.61
BMI Percentile (%)	61.4 $\pm$ 3.32	60.3 $\pm$ 4.57	62.8 $\pm$ 4.86
VO <sub>2max</sub> (ml/kg/min)*	41.3 $\pm$ 0.79	42.8 $\pm$ 1.12	39.4 $\pm$ 1.03
VO <sub>2max</sub> Percentile (%)	29.8 $\pm$ 2.79	28.6 $\pm$ 3.75	31.4 $\pm$ 4.19
OSPAN Accuracy	0.28 $\pm$ 0.02	0.31 $\pm$ 0.03	0.25 $\pm$ 0.03
Math Accuracy (%)*	89.6 $\pm$ 1.04	91.6 $\pm$ 1.05	86.9 $\pm$ 1.92
Math RT (ms)	3806.0 $\pm$ 106.70	3732.3 $\pm$ 130.00	3902.5 $\pm$ 178.98

Note: SES is classified as 'low' (score below 2), 'moderate' (score between 2 and 3), and 'high' (score greater than or equal 3); IQ = intelligent quotient; BMI = body mass index; RT = median reaction time. \* $p \leq .05$ , independent *t*-test between male and female groups.

performed more accurately than females (86.9  $\pm$  1.91%) only for math accuracy,  $t(95) = 2.29$ ,  $p \leq .05$ .

#### Bivariate correlations

Correlations (i.e. collective and parsed by sex) revealed that higher age, grade, and IQ were related to greater OSPAN accuracy, math accuracy, and shorter math RT ( $|r's| \geq .28$ ). Thus, age, grade, and IQ were included as covariates in the regression models along with sex, fitness, and the interaction product of sex  $\times$  fitness.

#### Regression analysis

The linear regression analysis for OSPAN word recall performance revealed a main effect of sex such that males performed better than females,  $\beta = -1.05$ ,  $t(90) = 2.21$ ,  $p \leq .05$ , 95% CI =  $[-.870, .024]$ , IQ,  $\beta = 0.28$ ,  $t(90) = 3.05$ ,  $p \leq .05$ , 95% CI =  $[.002, .008]$ , and a significant sex  $\times$  fitness interaction,  $\beta = 1.30$ ,  $t(90) = 2.52$ ,  $p \leq .05$ , 95% CI =  $[.001, .022]$ . Decomposition of the interaction revealed that for males, Step 1 was significant,  $R^2 = .23$ ,  $F(3, 51) = 6.47$ ,  $p \leq .05$ , revealing an association of higher IQ with higher OSPAN accuracy,  $\beta = 0.30$ ,  $t(51) = 2.31$ ,  $p \leq .05$ ,  $pr = .31$ , 95% CI  $[.001, .010]$ . Step 2 was also significant,  $\Delta R^2 = .07$ ,  $F(4, 50) = 6.52$ ,  $p \leq .05$ , such that higher fitness levels were associated with greater OSPAN accuracy with fitness accounting for an incremental amount of variance in OSPAN performance beyond associated descriptive variables,  $\beta = 0.26$ ,  $t(50) = 2.26$ ,  $p \leq .05$ ,  $pr = .31$ , 95% CI  $[.001, .013]$  (see Table 2 and Figure 1). For females, although Step 1 was significant,  $R^2 = .39$ ,  $F(3, 38) = 9.59$ ,  $p \leq .05$ , revealing an association of higher IQ with higher OSPAN accuracy,  $\beta = 0.26$ ,  $t(38) = 1.83$ ,  $p = .08$ ,  $pr = .29$ , 95% CI  $[.001,$

.009], Step 2 did not achieve significance for fitness,  $\beta = -0.18$ ,  $t(37) = 1.40$ ,  $p = .17$ ,  $pr = -.22$ , 95% CI  $[-.017, .006]$ .

The observed main effect of sex for OSPAN accuracy suggests that males outperform females at mean levels of fitness after controlling for demographic factors. However, simple slope analysis revealed no main effect of sex when re-centering the fitness values 1 *SD* below the mean,  $\beta = -0.07$ ,  $t(90) = 0.56$ ,  $p = .58$ ,  $pr = -.06$ , 95% CI  $[-0.77, 0.36]$ , suggesting that sexually dimorphic differences in working memory performance emerge only at higher fitness levels and are eliminated at lower fitness levels. However, since males demonstrate greater fitness levels on average compared to females, sexually dimorphic differences might be due to the lack of higher-fit females in this sample. Thus, planned comparisons were performed to evaluate groups matched for fitness. This was accomplished by binning participants into quartiles based on female fitness measures. All participants within the upper quartile range (43.5–59.8 ml/kg/min; excluding males with fitness levels above the quartile limits) were retained for OSPAN performance comparisons. Results revealed greater accuracy for males (OSPAN = .36) compared to females (OSPAN = .18),  $t(37) = 2.33$ ,  $p < .05$ , suggesting that even when matching groups based on fitness,  $t(37) = 1.29$ ,  $p = .21$ , dimorphic sex differences remain for working memory performance at higher fitness levels.

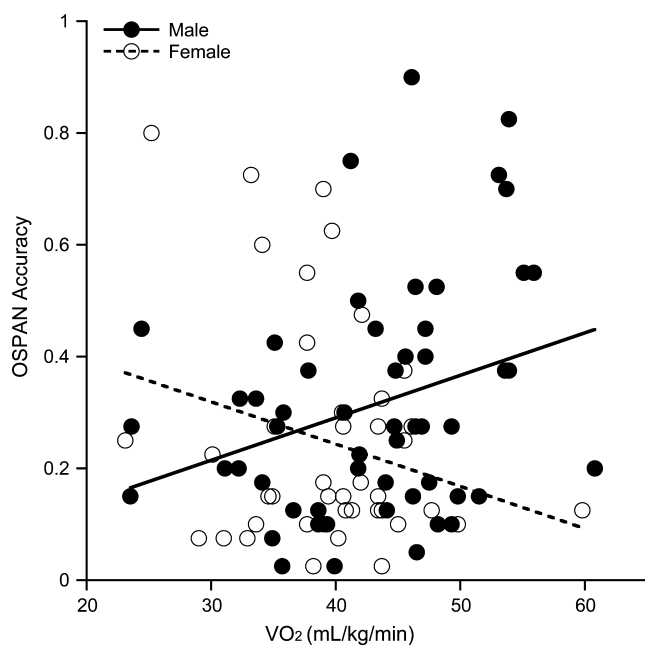
#### Discussion

We examined whether cardiorespiratory fitness differentially modulated working memory performance in male and female children using the OSPAN task. Collectively, results revealed that higher fitness was significantly associated with greater working memory performance

**Table 2** Study 1: summary of hierarchical regression analysis for variables predicting OSPAN accuracy in male ( $n = 55$ ) children

Measure	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	95% CI
Step 1					
Age	.082	.066	.33	1.24	(-.024, .220)
Grade	.000	.046	.00	0.01	(-.113, .068)
IQ	.005	.002	.30*	2.31	(.001, .010)*
Step 2					
Age	.078	.063	.31	1.23	(-.028, .200)
Grade	.002	.044	.01	0.05	(-.106, .070)
IQ	.005	.002	.28*	2.23	(.001, .009)*
Fitness	.006	.003	.26*	2.26	(.001, .013)*

Note:  $R^2 = .23$  for Step 1;  $\Delta R^2 = .07$  for Step 2 ( $p \leq .05$ ). 95% CI = confidence interval of unstandardized *B* coefficients derived from bootstrap analysis. \* $p \leq .05$ .

**Figure 1** Scatter plot for the sexual dimorphic bivariate relation between cardiorespiratory fitness and OSPAN accuracy ( $r = -.10$ , females;  $r = 0.26^*$ , males); \* $p \leq .05$  with two-tailed test.

in males but not for females. Thus, we demonstrated a sexually dimorphic pattern such that higher fitness levels are related to greater working memory capacity in developing male children, but not for female children.

## Study 2

The second study investigated the same relations with the exception of employing an *n*-back task as the index

of working memory performance. The *n*-back measures several components of working memory that are not shared with WMS tasks such as the OSPAN. To illustrate, Kane, Conway, Miura and Colflesh (2007) investigated the relation between the *n*-back and various WMS tasks. Consonant with previous investigations (Oberauer, 2005; Roberts & Gibson, 2002), a weak correlation was observed between the tasks, suggesting that although the WMS tasks and the *n*-back share some constructs such as maintenance, controlled attention under conditions of interference, and updating, they are not measuring identical constructs (Kane *et al.*, 2007). Therefore, if the sexually dimorphic patterns of fitness are related to general constructs of working memory, then similar results would be expected for the *n*-back. Lastly, unlike the OSPAN, the *n*-back provides a clear representation of incremental increases in working memory demand (i.e. 0-back to 2-back), allowing for a more detailed replication of fitness effects on conditions that require greater working memory demand.

## Methods

### Participants

One hundred and eight (54 females) preadolescent children were recruited from the East-Central Illinois area. All participants stemmed from a larger clinical trial (FITKids; Hillman, Pontifex, Castelli, Khan, Raine *et al.*, 2014). The purpose of the FITKids Trial was to investigate the effects of a 9-month physical activity intervention on changes in cognitive and brain function, and academic achievement in preadolescent children. The data presented herein were derived from the baseline assessment. Participants were excluded from Study 2 due to incomplete demographic data ( $n = 5$ ), task performance below zero on the  $d'$  measures ( $n = 3$ ), and failure to achieve maximal aerobic capacity criterion during the cardiorespiratory fitness assessment ( $n = 5$ ; see Cardiorespiratory Fitness Assessment methods in Study 1). Thus, analyses were conducted on the remaining 95 (45 female) preadolescent children (age range: 7.7–10 years). All participants provided written assent and their legal guardians provided written informed consent in accordance with the Institutional Review Board (IRB) of the University of Illinois at Urbana-Champaign. Prior to testing, legal guardians completed a battery of health history and demographic questionnaires on behalf of the participant. Health history documentation and prescreening information were collected utilizing the protocol employed in Study 1.

### *n*-back task

Participants performed the *n*-back task while seated 1 meter from a computer screen. Five different shapes of various colors (i.e. green square, red circle, blue triangle, purple star, and orange cross) were presented focally in a sequential manner. Three different conditions were performed (0-, 1-, and 2-back), with each condition presented in a fixed order for all participants. For the 0-back condition, participants were instructed to respond as quickly and accurately as possible with a right button press when the cross shape (i.e. target) appeared on the screen and a left button press when any of the other remaining four shapes (i.e. non-target) appeared. For the 1-back and 2-back conditions, instructions were similar to the 0-back with the exception of correctly matching and identifying the currently presented shape with the shape presented previously either one (1-back) or two (2-back) trials. A left button press indicated that the shape was not the same (i.e. non-target) as 1 or 2 trials back, and a right button press indicated that the shape was the same (i.e. target) as 1 or 2 trials back. Prior to each condition, 10 practice trials were administered to ensure that participants achieved a sufficient level of accuracy prior to actual testing. Each trial was presented for 2900 ms with a fixed inter-stimulus interval of 3000 ms on a black background. Each condition was presented in one block containing 80 trials with 16 targets (20% probability) for the 0-back condition and 20 targets (25% probability) for the 1- and 2-back conditions. Outcome variables included hit rate (the probability of correctly identifying a target), hit rate median reaction time (RT), correct rejections (CR; the probability of correctly identifying a non-target), CR median RT, false alarm rate (FA; the probability of incorrectly identifying a non-target as a target), and *d*' prime (*d*') accuracy. Calculation of *d*' followed the formula provided by Sorokin (1999),  $z(\text{adjusted hit rate}) - z(\text{adjusted false alarm rate})$ . Adjustments were implemented for perfect scores, such that if the probability of hits was 1.0 then an adjustment of  $2^{-(1/n)}$  ( $n$  = number of trials) would replace the maximum probability, and if the probability of false alarm rate was 0.0 then the adjustment of  $1 - (2^{-(1/n)})$  would replace the minimum probability. Higher values of *d*' indicate increased ability to discriminate between targets and non-targets with the highest possible score after adjustment equal to 4.02 for the 0-back condition, and 4.1 for the 1- and 2-back conditions. Prior to analysis, behavioral screening excluded participants with a *d*' score  $\leq 0$  for any of the three *n*-back conditions.

### Cardiorespiratory fitness assessment

See Study 1 methods for cardiorespiratory fitness protocol.

### Procedure

Experimental procedures were similar to Study 1, with the exception of administering an *n*-back task in place of the OSPAN task.

### Statistical analysis

Statistical analysis was performed following the procedure outlined in Study 1.

### Results

Means and standard error of the mean (SEM) are reported for demographic and fitness variables in Table 3, and *n*-back behavioral variables in Table 4. As expected, comparison of sex revealed higher mean fitness for males ( $41.2 \pm 0.91$  ml/kg/min) compared to females ( $37.6 \pm 0.80$  ml/kg/min),  $t(93) = 2.97$ ,  $p \leq .01$ . However, ranges of fitness levels were similar for males (28.2–53.2 ml/kg/min) and females (27.6–49.1 ml/kg/min). Lastly, analysis revealed greater response accuracy for males ( $83.7 \pm 2.34\%$ ) compared to females ( $73.1 \pm 2.90\%$ ) only for 2-back CR trials,  $t(93) = 2.87$ ,  $p = .01$ .

### Bivariate correlations

Correlations (i.e. collective and parsed by sex) revealed that higher age and grade were consistently related to greater accuracy and shorter RT across multiple trial types ( $|r's| \geq .27$ ). Thus, age and grade were included as covariates in the regression models along with sex, fitness, and the interaction product of sex  $\times$  fitness.

**Table 3** Study 2: Mean  $\pm 1$  SEM values for demographic and fitness measures

Measure	Combined	Male	Female
<i>n</i>	95	50	45
Age (years)	8.8 $\pm$ 0.06	8.6 $\pm$ 0.08	8.8 $\pm$ 0.09
Grade	2.9 $\pm$ 0.08	2.8 $\pm$ 0.10	3.0 $\pm$ 0.12
SES	1.9 $\pm$ 0.08	1.9 $\pm$ 0.12	2.0 $\pm$ 0.12
IQ	114.2 $\pm$ 1.44	114.4 $\pm$ 2.12	114.1 $\pm$ 1.94
BMI (kg/m <sup>2</sup> )	18.3 $\pm$ 0.37	18.5 $\pm$ 0.59	18.1 $\pm$ 0.42
BMI Percentile (%)	64.9 $\pm$ 2.80	64.0 $\pm$ 4.19	65.9 $\pm$ 3.72
VO <sub>2max</sub> (ml/kg/min)*	39.5 $\pm$ 0.63	41.2 $\pm$ 0.91	37.6 $\pm$ 0.80
VO <sub>2max</sub> Percentile (%)	21.3 $\pm$ 2.29	20.4 $\pm$ 3.15	22.2 $\pm$ 3.37

Note: SES is classified as 'low' (score below 2), 'moderate' (score between 2 and 3), and 'high' (score greater than or equal 3); IQ = intelligent quotient; BMI = body mass index. \* $p \leq .05$ , independent *t*-test between male and female groups.



**Table 4** Study 2: Mean  $\pm$  1 SEM values for *n*-back behavior data

Measure	Male			Female		
	0-back	1-back	2-back	0-back	1-back	2-back
Hit median RT (ms)	702.8 $\pm$ 19.03	967.7 $\pm$ 26.87	1049.0 $\pm$ 40.24	696.1 $\pm$ 20.01	926.6 $\pm$ 25.54	967.5 $\pm$ 46.80
Hit rate (%)	91.6 $\pm$ 0.97	76.6 $\pm$ 1.93	58.9 $\pm$ 2.18	88.8 $\pm$ 1.73	73.4 $\pm$ 2.39	55.8 $\pm$ 2.57
CR median RT (ms)	657.8 $\pm$ 16.62	965.7 $\pm$ 27.62	1090.0 $\pm$ 35.30	624.9 $\pm$ 18.02	948.8 $\pm$ 30.86	1148.1 $\pm$ 52.32
CR accuracy (%)	97.1 $\pm$ 0.45	91.1 $\pm$ 1.44	83.6 $\pm$ 2.34*	97.7 $\pm$ 0.4	88.5 $\pm$ 1.90	73.1 $\pm$ 2.90*
FA (%)	1.3 $\pm$ 0.23	5.8 $\pm$ 1.38	9.3 $\pm$ 1.74	1.2 $\pm$ 0.26	4.5 $\pm$ 0.89	12.6 $\pm$ 1.71
<i>d'</i>	3.5 $\pm$ 0.07	2.6 $\pm$ 0.10	1.8 $\pm$ 0.10	3.4 $\pm$ 0.07	2.6 $\pm$ 0.10	1.5 $\pm$ 0.11

Note: RT = reaction time; CR = correct reject; FA = false alarms. \* $p \leq .05$ . Independent *t*-test between groups on the same *n*-back measure.

### Regression analysis

For 2-back *d'* accuracy, analysis revealed no main effect of sex, age, grade, or fitness; however, results revealed a significant sex  $\times$  fitness interaction,  $\beta = 1.55$ ,  $t(89) = 2.13$ ,  $p \leq .05$ , 95% CI = [.003, .104]. Decomposition of the interaction revealed that for males, although Step 1 was non-significant,  $R^2 = .04$ ,  $F(2, 47) = 1.95$ ,  $p = .15$ , Step 2 was significant,  $\Delta R^2 = .08$ ,  $F(3, 46) = 2.82$ ,  $p \leq .05$ , such that higher fitness levels were associated with greater 2-back *d'* accuracy, with fitness accounting for an incremental amount of variance in working memory beyond associated descriptive variables,  $\beta = 0.28$ ,  $t(46) = 2.07$ ,  $p \leq .05$ ,  $pr = .29$ , 95% CI [.001, .059] (see Table 5 and Figure 2). For females, results revealed no significant relation between fitness and 2-back *d'* accuracy,  $\Delta R^2 = .04$ ,  $F(3, 41) = 1.98$ ,  $p = .13$ , 95% CI = [-.067, .018].

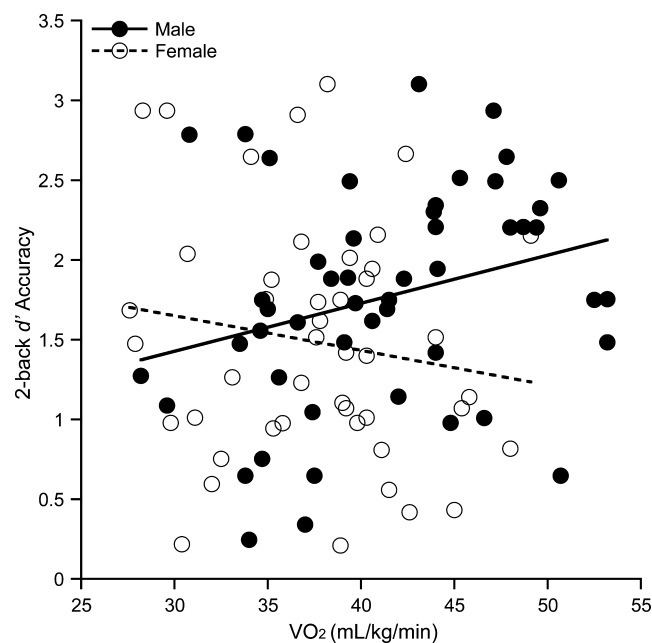
Since the linear regression analysis revealed no main effect of sex, males and females performed statistically equivalently at mean fitness levels. However, a simple slope analysis was conducted that re-centered the fitness values 1 *SD* above the mean, which revealed a main effect of sex,  $\beta = 0.45$ ,  $t(89) = 2.87$ ,  $p \leq .05$ ,  $pr = .29$ , 95% CI [0.25, 1.51], suggesting that sexually dimorphic

**Table 5** Study 2: Summary of hierarchical regression analysis for variables predicting 2-back *d'* accuracy in male ( $n = 50$ ) children

Measure	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	95% CI
Step 1					
Age	0.29	0.22	.24	1.29	(-.145, .757)
Grade	0.05	0.19	.05	0.29	(-.362, .419)
Step 2					
Age	0.32	0.22	.26	1.47	(-.106, .795)
Grade	0.02	0.18	.02	0.12	(-.351, .386)
Fitness	0.03	0.02	.28*	2.07	(.000, .059)*

Note:  $R^2 = .06$  for Step 1;  $\Delta R^2 = .08$  for Step 2 ( $p \leq .05$ ). 95% CI = confidence interval of unstandardized *B* coefficients derived from bootstrap analysis. \* $p \leq .05$ .

differences in working memory performance are evident at higher fitness levels, but not at lower or mean fitness levels. Similar to the OSPAN, since males demonstrated higher fitness levels on average compared to females, these differences might be due to the lack of higher-fit females in this sample. Thus, planned comparisons were performed (see Regression Analysis in Study 1 for methods) to evaluate groups matched for fitness (40.8–49.1 ml/kg/min). Results revealed greater accuracy for males (2-back *d'* = 2.06) compared to females (2-back *d'* = 1.25),  $t(29) = 3.29$ ,  $p < .05$ , even when matching the highest fit females with their male counterparts. This suggests that even when matching groups based on

**Figure 2** Scatter plot for the sexual dimorphic bivariate relation between cardiorespiratory fitness and 2-back *d'* accuracy ( $r = -.08$ , females;  $r = 0.29^*$ , males); \* $p \leq .05$  with two-tailed test.

fitness,  $t(29) = 1.07$ ,  $p = .29$ , dimorphic sex differences remain in working memory performance at higher fitness levels.

### Discussion

Results of Study 2 replicate the main findings observed in Study 1. First, both studies demonstrate that the association between higher fitness and better working memory performance is sexually dimorphic, such that the effect is evident for males and not females. Second, Study 2 provides evidence that aerobic fitness modulates general working memory constructs, such as maintenance, controlled attention under conditions of interference, and updating or active manipulation. Lastly, these sexual dimorphic relations observed in Study 2 (selective for males) were only observed for the 2-back condition, suggesting selective benefits for more demanding working memory conditions.

### Study 3

The third study extends the findings of the first two studies by examining the Sternberg task. Unlike the working memory tasks in the first two studies (i.e. the  $n$ -back and OSPAN tasks) which require maintenance, controlled attention under conditions of interference, and updating; the Sternberg task places greater demand on storage and maintenance operations. As such, some researchers argue that the demands of the Sternberg task match those similar to a short-term memory task (Klein, Rauh & Biscaldi, 2010; Schooler, Caplan, Revell, Salazar & Grafman, 2008) rather than a working memory task. Regardless, the Sternberg task is advantageous for the present investigation in that interference control and active manipulation are attenuated to a greater degree compared to the tasks employed in the first two studies, allowing for further examination of shared aspects of working memory between the various tasks.

Therefore, the aim of Study 3 was to extend the sexually dimorphic relationship between fitness and working memory observed in the previous studies and further assess the general versus selective nature among shared constructs of working memory across the different tasks. If fitness modulation for males is generalized across shared working memory constructs then we would predict that the results of Study 3 would replicate the findings of the first two studies. Lastly, given that the Sternberg task incorporates various levels of difficulty in trial types (i.e. 1-letter, 3-letter, and 5-letter arrays), providing discrete representation of increasing working memory demand as letter-string length increases (from 1-letter to 5-letter), it

was predicted that observed differences in sex and fitness would replicate Study 2 and demonstrate selective results for conditions that require greater working memory demand.

### Methods

#### Participants

One hundred and twenty (50 females) preadolescent children were recruited from the East-Central Illinois area. Similar to Study 2, all participants stemmed from baseline measures of the FITKids clinical trial (Hillman *et al.*, 2014), but represent an entirely separate sample of children compared to Studies 1 and 2. Participants were excluded from the present investigation due to incomplete demographic data ( $n = 5$ ), task performance below zero on the  $d'$  measures ( $n = 15$ ), and failure to achieve maximal aerobic capacity criterion during the cardiorespiratory fitness assessment ( $n = 17$ ; see Cardiorespiratory Fitness Assessment methods in Study 1). Thus, analyses were conducted on the remaining 83 (37 female) preadolescent children (age range: 7.8–9.9 years). All participants provided written assent and their legal guardians provided written informed consent in accordance with the Institutional Review Board (IRB) of the University of Illinois at Urbana-Champaign. Prior to testing, legal guardians completed a battery of health history and demographic questionnaires on behalf of the participant. Health history documentation and prescreening information were collected utilizing the protocol employed in Study 1.

#### Sternberg task

Stimuli were memory sets comprising uppercase consonants (e.g. KRM), and probe letters comprising lowercase consonants flanked by “?” to match the perceptual size of the memory sets (e.g. ?m?). Participants were seated 1 meter from a computer screen as memory sets and probes, which contained an array of 1, 3, or 5 letters, were presented in a sequential and focal manner. Participants were asked to respond as quickly and accurately as possible with a right button press if the probe letter correctly matched any letter contained in the previous memory set (i.e. target), or a left button press if the probe letter did not correctly match any letter contained in the previous memory set (i.e. non-target). Prior to each task condition, 10 practice trials were given to ensure that participants achieved sufficient accuracy prior to testing. The stimulus durations were 2500 ms for encoded array (S1) and 250 ms for probe letter (S2), with a 2000 ms inter-stimulus interval (from S1 offset to S2

onset) and a 2500 ms response window (from S2 onset to S1 onset). Set sizes were presented randomly within four separate blocks containing a combined total of 180 trials with equiprobable presentation for targets and non-targets. Outcome variables were identical to the *n*-back task (see serial *n*-back task in Study 2 methods section). The highest possible score for *d'*, following hit rate and FA corrections, was 3.9 for all memory set sizes. Prior to analysis, behavioral screening excluded participants with a *d'* score  $\leq 0$  in any of the three set sizes.

#### Cardiorespiratory fitness assessment

See Study 1 methods for cardiorespiratory fitness protocol.

#### Procedure

Experimental procedures were similar to Studies 1 and 2, with the exception of administering the Sternberg task.

#### Statistical analysis

Statistical analysis was performed following the procedure outlined in Study 1.

#### Results

Means and standard error of the mean (SEM) are reported for demographic and fitness variables in Table 6, and Sternberg behavioral variables in Table 7. Consonant with Studies 1 and 2, comparison of sex revealed higher mean fitness levels for males ( $40.1 \pm 0.99$  ml/kg/min; range: 26.4–52.8) compared to females ( $34.2 \pm 1.16$  ml/kg/min; range: 19.9–48.2),  $t(82) = 3.93$ ,  $p \leq .01$ . In addition, females demonstrated greater BMI ( $20.7 \pm 0.8$  kg/m<sup>2</sup>) compared to males

**Table 6** Study 3: Mean  $\pm 1$  SEM values for demographic and fitness measures

Measure	Combined	Male	Female
<i>n</i>	84	47	37
Age (years)	8.9 $\pm$ 0.06	8.9 $\pm$ 0.08	8.9 $\pm$ 0.10
Grade	3.5 $\pm$ 0.07	3.4 $\pm$ 0.10	3.5 $\pm$ 0.10
SES	2.0 $\pm$ 0.10	2.0 $\pm$ 0.13	2.0 $\pm$ 0.16
IQ	110.0 $\pm$ 1.38	109.6 $\pm$ 1.68	110.4 $\pm$ 2.33
BMI (kg/m <sup>2</sup> )*	19.5 $\pm$ 0.46	18.6 $\pm$ 0.54	20.7 $\pm$ 0.75
BMI Percentile (%)	71.1 $\pm$ 3.08	66.8 $\pm$ 4.11	76.5 $\pm$ 4.56
VO <sub>2max</sub> (ml/kg/min)*	37.4 $\pm$ 0.81	40.0 $\pm$ 0.97	34.2 $\pm$ 1.16
VO <sub>2max</sub> Percentile (%)	18.2 $\pm$ 2.29	18.7 $\pm$ 2.98	17.5 $\pm$ 3.60

Note: SES is classified as 'low' (score below 2), 'moderate' (score between 2 and 3), and 'high' (score greater than or equal 3); IQ = intelligent quotient; BMI = body mass index. \* $p \leq .05$ , independent *t*-test between male and female groups.

( $18.5 \pm 0.5$  kg/m<sup>2</sup>),  $t(82) = 2.35$ ,  $p \leq .05$ . Lastly, males and females did not differ in performance across any measure of the Sternberg,  $t_s(82) \leq 1.52$ ,  $p_s \geq .13$ .

#### Bivariate correlations

Correlations (i.e. collective and parsed by sex) revealed that higher grade, SES, and IQ were consistently related to greater accuracy and shorter RT across multiple trial types ( $r$ 's  $\geq .23$ ). Thus, grade, SES, and IQ were included as covariates in the regression models along with sex, fitness, and the interaction product of sex  $\times$  fitness.

#### Regression analysis

For 5-letter *d'* accuracy, analysis revealed a main effect of grade,  $\beta = 0.23$ ,  $t(91) = 2.31$ ,  $p \leq .05$ , (bootstrap) 95% CI = [.039, .429], IQ,  $\beta = 0.44$ ,  $t(91) = 4.06$ ,  $p \leq .05$ , 95% CI = [.013, .036], and a significant sex  $\times$  fitness interaction,  $\beta = 1.15$ ,  $t(76) = 1.95$ ,  $p \leq .05$ , 95% CI = [.001, .068]. Decomposition of the interaction revealed that for males, Step 1 was significant,  $R^2 = .31$ ,  $F(3, 42) = 7.61$ ,  $p \leq .05$ , revealing an association of higher IQ with greater *d'* accuracy,  $\beta = 0.02$ ,  $t(42) = 2.77$ ,  $p \leq .05$ ,  $pr = .39$ , 95% CI = [.007, .036]. Step 2 was also significant,  $\Delta R^2 = .07$ ,  $F(4, 41) = 7.45$ ,  $p \leq .05$ , such that higher grade, IQ, and fitness were associated with greater 5-letter *d'* accuracy with fitness accounting for an incremental amount of variance in working memory beyond associated descriptive variables,  $\beta = 0.29$ ,  $t(41) = 2.21$ ,  $p \leq .05$ ,  $pr = .33$ , 95% CI [.004, .049] (see Table 8 and Figure 3). For females, although Step 1 was significant,  $R^2 = .14$ ,  $F(1, 35) = 7.03$ ,  $p \leq .05$ , revealing an association of higher IQ with greater 5-letter *d'* accuracy,  $\beta = 0.02$ ,  $t(35) = 2.65$ ,  $p \leq .05$ ,  $pr = .41$ , 95% CI [.006, .030], Step 2 did not achieve significance for fitness,  $\beta = -0.00$ ,  $t(34) = 0.14$ ,  $p = .89$ ,  $pr = -.02$ , 95% CI [-.035, .028].

Furthermore, the linear regression analysis revealed no main effect of sex, suggesting equivalent working memory performance between males and females at mean fitness levels. However, a simple slope analysis in which the fitness values were re-centered 1 *SD* above the mean revealed a main effect of sex,  $\beta = 0.31$ ,  $t(76) = 2.04$ ,  $p \leq .05$ ,  $pr = .26$ , 95% CI [-0.08, 1.22], suggesting that sexually dimorphic differences in working memory performance were evident at higher fitness levels, but not at mean or low fitness levels. Finally, similar to the analysis performed for the OSPAN and *n*-back tasks, planned comparisons (see Regression Analysis in Study 1 for methods) were performed to evaluate groups matched for fitness (40.5–48.2 ml/kg/min). Results revealed greater accuracy for males (5-letter *d'* = 1.49) compared to females (5-letter *d'* = 0.84),  $t(23) = 2.66$ ,

**Table 7** Study 3: Mean  $\pm$  1 SEM values for Sternberg behavior data

Measure	Male			Female		
	1-letter	3-letter	5-letter	1-letter	3-letter	5-letter
Hit median RT (ms)	1041.5 $\pm$ 53.28	1352.7 $\pm$ 50.38	1361.3 $\pm$ 53.45	1121.6 $\pm$ 55.83	1385.4 $\pm$ 46.87	1407.3 $\pm$ 63.15
Hit rate (%)	67.9 $\pm$ 2.74	66.6 $\pm$ 2.73	57.4 $\pm$ 2.51	65.0 $\pm$ 2.45	66.4 $\pm$ 2.73	51.4 $\pm$ 2.82
CR median RT (ms)	1123.1 $\pm$ 45.95	1332.2 $\pm$ 39.36	1370.7 $\pm$ 52.92	1223.2 $\pm$ 56.00	1427.3 $\pm$ 49.30	1453.1 $\pm$ 50.36
CR (%)	77.5 $\pm$ 2.73	72.43 $\pm$ 2.96	64.7 $\pm$ 2.93	71.7 $\pm$ 2.71	67.3 $\pm$ 3.73	58.2 $\pm$ 2.98
FA (%)	11.8 $\pm$ 1.66	9.9 $\pm$ 1.50	19.0 $\pm$ 1.91	10.8 $\pm$ 1.30	11.9 $\pm$ 2.14	19.1 $\pm$ 2.03
$d'$	1.8 $\pm$ 0.12	1.9 $\pm$ 0.13	1.2 $\pm$ 0.09	1.8 $\pm$ 0.12	1.8 $\pm$ 0.16	1.0 $\pm$ 0.10

Note: RT = reaction time; CR = correct reject; FA = false alarms. \* $p \leq .05$ . Independent  $t$ -test between groups on the same Sternberg measure.

$p \leq .05$ , even when matching the highest fit females with their male counterparts. This suggests that even when matching groups based on fitness,  $t(23) = 0.64$ ,  $p = .53$ , dimorphic sex differences remain in working memory performance at higher fitness levels.

### Discussion

The results of Study 3 replicate the main findings of Studies 1 and 2. Specifically, the results revealed selective fitness and working memory relations for male children and only for trial types that place greater demand on working memory (i.e. 5-letter  $d'$ ). No such relationship was observed among female children. In addition, given the nature of the Sternberg task, results further disclose that fitness relations for males appear selective for storage and active maintenance constructs of working memory.

### General discussion

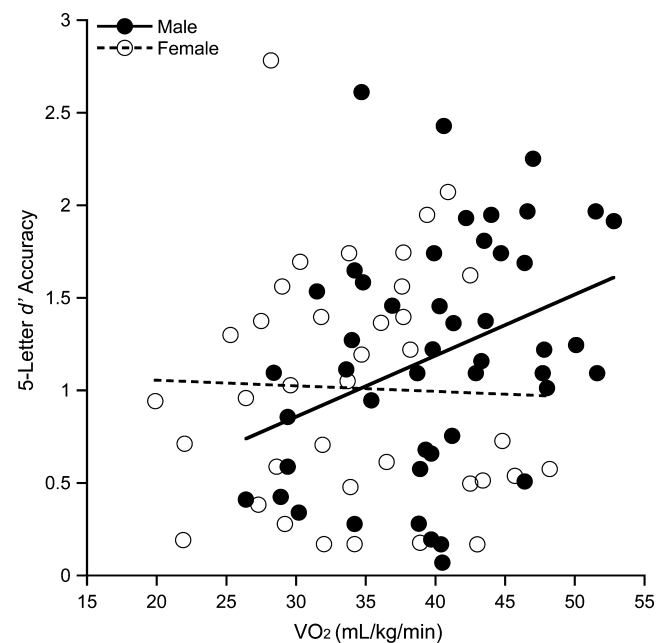
The findings across all three studies provide compelling evidence to suggest that the beneficial relations

**Table 8** Study 3: Summary of hierarchical regression analysis for variables predicting Sternberg 5-letter  $d'$  accuracy in male ( $n = 47$ ) children

Measure	$B$	$SE B$	$\beta$	$t$	95% CI
Step 1					
Grade	0.22	0.13	.23	1.79	(-.034, .551)
SES	0.12	0.10	.17	1.22	(-.091, .311)
IQ	0.02	0.01	.39**	2.77	(.007, .036)**
Step 2					
Grade	0.25	0.12	.26*	2.10	(-.011, .550)
SES	0.04	0.10	.05	0.37	(-.237, .249)
IQ	0.02	0.01	.40**	2.94	(.009, .039)**
Fitness	0.03	0.01	.29*	2.21	(.004, .049)*

Note:  $R^2 = .31$  for Step 1;  $\Delta R^2 = .07$  for Step 2 ( $p \leq .05$ ). 95% CI = confidence interval of unstandardized  $B$  coefficients derived from bootstrap analysis. \* $p \leq .05$ ; \*\* $p \leq .01$ .

between fitness and working memory are selective for prepubertal male children. Specifically, results demonstrate that increased working memory performance is related to greater fitness for males but not females. In addition, these effects were generalized across several constructs of working memory (i.e. storage and active maintenance) and selective for conditions that placed greater demand on working memory. Greater performance was also observed for males compared to females but only at higher fitness levels, providing further evidence that sexual dimorphic patterns of cardiorespiratory health differentially influence aspects of working memory.

**Figure 3** Scatter plot for the sexual dimorphic bivariate relation between cardiorespiratory fitness and 5-letter  $d'$  accuracy ( $r = -.03$ , females;  $r = 0.35^*$ , males); \* $p \leq .05$  with two-tailed test.



Collectively, these sex-related differences of fitness on working memory are novel and may provide a link for future research aimed at elucidating the benefits of cardiorespiratory fitness on mechanisms of brain health and cognition. As described earlier, research has demonstrated volumetric differences in the hippocampus related to fitness (Chaddock *et al.*, 2010; Erickson *et al.*, 2011). Such changes are supported by underlying healthy development of neurotrophic factors, specifically brain-derived neurotrophic factors (BDNF), within these subcortical regions. BDNF plays a key role in growth (i.e. neurogenesis) and longevity of synaptic efficacy, neuronal connectivity, and neural plasticity (Altar & Distefano, 1998; Lu & Chow, 1999; McAllister, Katz & Lo, 1999; Schinder & Poo, 2000). Research has demonstrated that following several days of voluntary wheel-running in mice, levels of BDNF expressed in the hippocampus increased (Neeper, Gomez-Pinilla, Choi & Cotman, 1995) with accompanying improvements in spatial learning (van Praag, Christie, Sejnowski & Gage, 1999). Interestingly, sex steroid hormones, such as estrogen, mediate this relation of exercise and hippocampal expression of BDNF and neurogenesis (Berchtold, Kesslak, Pike, Adlard & Cotman, 2001; Singh, Meyer & Simpkins, 1995). That is, when low levels of hippocampal estrogen are available in female rats, exercise fails to increase BDNF levels. Conversely, when estrogen levels were artificially increased through estrogen replacement, exercise increased BDNF levels to a greater degree than by estrogen replacement alone (Berchtold *et al.*, 2001).

Research further demonstrates that sexually dimorphic differences in brain structure and function described earlier may be dependent on circulating sex steroid hormones during critical developmental changes (Cooke, Hegstrom, Villeneuve & Breedlove, 1998; Kawata, 1995). For example, increasing levels of testosterone (males) and estrogen (females) during puberty was related to emerging sex differences in the hippocampus and amygdala in males (Neufang, Specht, Hausmann, Güntürkün, Herpertz-Dahlmann *et al.*, 2009), and grey matter changes in females (Peper, Brouwer, Schnack, van Baal, van Leeuwen *et al.*, 2009). Further, animal studies have demonstrated that the production of new cells (i.e. neurogenesis) in the hippocampus is highest during peak levels of estrogen in females (Tanapat, Hastings, Reeves & Gould, 1999). In addition, not only has estrogen been shown to modulate volume in regions of the hippocampus (for review see Wnuk, Korol & Erickson, 2012), but researchers have also observed concurrent changes in cognitive behavior on tasks such as the spatial navigation task in animals (Isgor & Sengelaub, 1998). In humans, these behavioral patterns appear reciprocal in a manner

that parallels estrogen and testosterone fluctuations. That is, females demonstrate poorer cognitive performance (only for tasks that typically demonstrate a male advantage) during peak estrogen levels (Hampson, 1990; Hampson & Kimura, 1988; Maki, Rich & Rosebaum, 2002; but see Kimura & Hampson, 1994, for review), whereas for males working memory performance improves when testosterone levels are high in relation to estrogen (Janowsky, Chavez & Orwoll, 2000).

Taken together, the overall pattern of potential mechanisms discussed previously reveals multiple robust similarities. Interestingly, sex steroid hormones appear to be a common variable in (1) sex dimorphic development of the hippocampus, (2) the reciprocal relation with cognitive performance in males and females, and (3) mediating fitness effects on BDNF gene expression in the hippocampus. Although the present investigation did not assess circulating sex hormones in children, previous research has demonstrated that prepubertal females have significantly higher levels of estrogen compared to age-matched males (Courant, Aksglaede, Antignac, Monteau, Sorensen *et al.*, 2010). These increased levels of estrogen are indicative of pubertal onset in females resulting in a cascade of events associated with maturation. Although highly speculative, estrogen might be a common variable related to the differential findings of the present investigation given its particular contribution to pubertal onset in females at this age, its moderating effects of fitness on BDNF gene expression, and its acute interaction with cognitive outcomes associated with working memory. However, it is unclear as to whether the presence of sex steroid hormones interact in isolation from fitness to influence cognitive performance or whether estrogen-dependent hippocampal development differentiates pathways selective for fitness modulation in males. Therefore, future investigations should incorporate methods to assess sex hormones to determine whether any interactions are observed that might further our understanding of why the relation of fitness and working memory performance is selective for males and not females.

Despite replicating this effect across all studies, certain limitations should be considered. First, our samples represent a distinct age group of children (~7–10 years old), who in some cases are at the threshold of pubertal development. Future research may benefit by evaluating these relations across the developmental spectrum to better understand whether the present findings are selective for prepubertal children. Second, working memory was the only aspect of cognitive control investigated in the present investigation. Future research could benefit by incorporating other aspects of cognitive control to further parse the influence of sex on fitness for this aspect of cognition.

Third, although our samples represented a wide distribution of fitness levels, the majority of children in each study were lower-fit in comparison to national norms of fitness for this age group (less than the 30th percentile). Thus, the relation between cognitive control and fitness may be curvilinear at higher fitness levels (i.e. threshold; see Sibley & Etnier, 2003) and that the tolerance for a threshold effect may be greater for females compared to males. Regardless, given that the present investigation evaluated samples of children who are considered lower-fit, our results further demonstrate the beneficial effects of fitness on cognition among children who might benefit the most from physical activity programs. Fourth, the present results might be influenced by the inability to maintain attention (i.e. cognitive vigilance) for the duration of the task especially for lower-fit males. Therefore, future research may benefit by evaluating the impact of such factors. Lastly, our investigation was cross-sectional, thus limiting causal conclusions from the observed outcomes. Future investigations may benefit by performing randomized controlled physical activity interventions to better understand sex-related differences in programs aimed at improving fitness.

The present investigation provides new evidence that not only add to the ever-growing literature demonstrating brain health benefits associated with cardiorespiratory fitness, but further suggests that sex may modify this relationship. Future research should consider the unique contribution of sexual dimorphic patterns when evaluating benefits of fitness to brain health. These results not only provide support for the importance of regular physical activity for developing prepubertal children, but also reveal the need for individualized physical activity programs among institutions and organizations intimately involved in children's daily activities.

## Acknowledgement

This research was supported by a grant from the National Institute of Child Health and Human Development (NIH HD055352) to Charles Hillman.

## Author disclosure statement

No conflicting financial interests exist.

## References

Aiken, L.S., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications.

- Altar, A., & DiStefano, P.S. (1998). Neurotrophin trafficking by anterograde transport. *Trends in Neuroscience*, **21**, 433–437. doi:10.1016/S0166-2236(98)01273-9
- American College of Sports Medicine (ACSM) (2010). *ACSM's guidelines for exercise testing and prescription* (8th edn.). New York: Lippincott Williams & Wilkins.
- Andersen, R.E., Crespo, C.J., Bartlett, S.J., Cheskin, L.J., & Pratt, M. (1998). Relationship of physical activity and television watching with body weight and level of fatness among children: Results from the Third National Health and Nutrition Examination Survey. *Journal of American Medical Association*, **279**, 938–942. doi:10.1001/jama.279.12.938
- Astur, R.S., Tropp, J., Sava, S., Constable, R.T., & Markus, E.J. (2004). Sex differences and correlations in a virtual Morris water task, a virtual radial arm maze, and mental rotation. *Behavioral Brain Research*, **151**, 103–115. doi:10.1016/j.bbr.2003.08.024
- Axmacher, N., Henseler, M.M., Jensen, O., Weinreich, I., Elger, C.E. *et al.* (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences, USA*, **107** (7), 3228–3233. doi:10.1073/pnas.1301653110
- Axmacher, N., Mormann, F., Fernández, G., Cohen, M.X., Elger, C.E. *et al.* (2007). Sustained neural activity patterns during working memory in the human medial temporal lobe. *Journal of Neuroscience*, **27**, 7807–7816. doi:10.1523/JNEUROSCI.0962-07.2007
- Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–89). New York: Academic Press.
- Bar-Or, O. (1983). *Pediatric sports medicine for the practitioner: From physiologic principles to clinical applications*. New York: Springer-Verlag.
- Bell, E.C., Willson, M.C., Wilman, A.H., Dave, S., & Silverstone, P.H. (2006). Males and females differ in brain activation during cognitive tasks. *NeuroImage*, **30**, 529–538. doi:10.1016/j.neuroimage.2005.09.049
- Berchtold, N.C., Kesslak, J.P., Pike, C.J., Adlard, P.A., & Cotman, C.W. (2001). Estrogen and exercise interact to regulate brain-derived neurotrophic factor mRNA and protein expression in the hippocampus. *European Journal of Neuroscience*, **14**, 1992–2002. doi:10.1046/j.0953-816x.2001.01825.x
- Birnbaum, A.S., Lytle, L.A., Murray, D.M., Story, M., Perry, C.L. *et al.* (2002). Survey development for assessing correlates of young adolescents' eating. *American Journal of Health Behavior*, **26**, 284–295. doi:10.5993/AJHB.26.4.5
- Blakemore, S.-J., Burnett, S., & Dahl, R.E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping*, **31**, 926–933. doi:10.1002/hbm.21052
- Bramen, J.E., Hranilovich, J.A., Dahl, R.E., Forbes, E.E., Chen, J. *et al.* (2011). Puberty influences medial temporal lobe and cortical gray matter maturation differently in boys than girls matched for sexual maturity. *Cerebral Cortex*, **21**, 636–646. doi:10.1093/cercor/bhq137
- Castelli, D.M., Hillman, C.H., Buck, S.M., & Erwin, H.E. (2007). Physical fitness and academic achievement in third- and fifth-grade students. *Journal of Sport & Exercise Psychology*, **29**, 239–252.

- Centers for Education Policy (2007). *Choices, changes, and challenges: Curriculum and instruction in the NCLB era*. Washington, DC: Center for Education Policy.
- Chaddock, L., Erickson, K.I., Prakash, R.S., Kim, J.S., Voss, M.W. *et al.* (2010). A neuroimaging investigation of the association between aerobic fitness, hippocampal volume, and memory performance in preadolescent children. *Brain Research*, **1358**, 172–183. doi:10.1016/j.brainres.2010.08.049
- Chein, J.M., Moore, A.B., & Conway, A.R.A. (2011). Domain-general mechanisms of complex working memory span. *NeuroImage*, **54**, 550–559. doi:10.1016/j.neuroimage.2010.07.067
- Chomitz, V.R., Slining, M.M., McGowan, R.J., Mitchell, S.E., Dawson, G.F. *et al.* (2009). Is there a relationship between physical fitness and academic achievement? Positive results from public school children in the northeastern United States. *Journal of School Health*, **79**, 30–37. doi:10.1111/j.1746-1561.2008.00371.x
- Christakou, A., Halari, R., Smith, A.B., Ifkovits, E., Brammer, M. *et al.* (2009). Sex-dependent age modulation of frontostriatal and temporo-parietal activation during cognitive control. *NeuroImage*, **48** (1), 223–236. doi:10.1016/j.neuroimage.2009.06.070
- Colcombe, S.J., Erickson, K.I., Raz, N., Webb, A.G., Cohen, N.J. *et al.* (2003). Aerobic fitness reduces brain tissue loss in aging humans. *Journal of Gerontology: Medical Sciences*, **58A**, 176–180. doi:10.1093/gerona/58.2.M176
- Colcombe, S.J., Erickson, K.I., Scalf, P.E., Kim, J.S., Prakash, R. *et al.* (2006). Aerobic exercise training increases brain volume in aging humans. *Journal of Gerontology: Medical Sciences*, **61A**, 1166–1170.
- Colcombe, S.J., Kramer, A.F., Erickson, K.I., Scalf, P., McAuley, E. *et al.* (2004). Cardiovascular fitness, cortical plasticity, and aging. *Proceedings of the National Academy of Sciences, USA*, **101**, 3316–3321. doi:10.1073/pnas.0400266101
- Colom, R., Stein, J.L., Rajagopalan, P., Martínez, K., Hermel, D. *et al.* (2013). Hippocampal structure and human cognition: key role of spatial processing and evidence supporting the efficiency hypothesis in females. *Intelligence*, **41**, 129–140. doi:10.1016/j.intell.2013.01.002
- Conway, A.R.A., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O. *et al.* (2005). Working memory span tasks: a methodological review and user's guide. *Psychonomic Bulletin & Review*, **12** (5), 769–786.
- Cooke, B., Hegstrom, C.D., Villeneuve, L.S., & Breedlove, S.M. (1998). Sexual differentiation of the vertebrate brain: principles and mechanisms. *Frontiers in Neuroendocrinology*, **19**, 323–362. doi:10.1006/frne.1998.0171
- Courant, F., Aksglaede, L., Antignac, J.-P., Monteau, F., Sorensen, K. *et al.* (2010). Assessment of circulating sex steroid levels in prepubertal and pubertal boys and girls by a novel ultrasensitive gas chromatography-tandem mass spectrometry method. *Journal of Clinical Endocrinology & Metabolism*, **95**, 82–92. doi:10.1210/jc.2009-1140
- Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series, No. 26. New York: Oxford University Press.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.
- Davies, P.L., Segalowitz, S.J., & Gavin, W.J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology*, **25**, 355–376. doi:10.1207/s15326942dn2503\_6
- De Bellis, M.D., Keshavan, M.S., Beers, S.R., Hall, J., Frustaci, K. *et al.* (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex*, **11**, 552–557. doi:10.1093/cercor/11.6.552
- D'Esposito, M., Detre, J.A., Alsop, D.C., Shin, R.K., Atlas, S. *et al.* (1995). The neural basis of the central executive system of working memory. *Nature*, **378**, 279–281. doi:10.1038/378279a0
- D'Esposito, M., Postle, B.R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Experimental Brain Research*, **133**, 3–11. doi:10.1007/s002210000395
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry. In D.T. Stuss & R.T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). New York: Oxford University Press.
- Diamond, A., Barnett, W.S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, **318**, 1387–1388. doi:10.1126/science.1151148
- Donnelly, J.E., Greene, J.L., Gibson, C.A., Smith, B.K., Washburn, R.A. *et al.* (2009). Physical activity across the curriculum (PAAC): a randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine*, **49**, 336–341. doi:10.1016/j.ypmed.2009.07.022
- Drollette, E.S., Scudder, M.R., Raine, L.B., Moore, R.D., Saliba, B.J. *et al.* (2014). Acute exercise facilitates brain function and cognition in children who need it most: an ERP study of individual differences in inhibitory control capacity. *Developmental Cognitive Neuroscience*, **7**, 53–64. doi:10.1016/j.dcn.2013.11.001
- Drollette, E.S., Shishido, T., Pontifex, M.B., & Hillman, C.H. (2012). Maintenance of cognitive control during and after walking in preadolescent children. *Medicine & Science in Sports & Exercise*, **10**, 2017–2024. doi:10.1249/MSS.0b013e318258bcd5
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.
- Eisenmann, J. (2003). Secular trends in variables associated with the metabolic syndrome of North American children and adolescents: a review and synthesis. *American Journal of Human Biology*, **15**, 786–794. doi:10.1002/ajhb.10214
- Engle, R.W., & Kane, M.J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation*, Vol. **44** (pp. 145–199). New York: Elsevier.



- Erickson, K.I., Prakash, R.S., Voss, M.W., Chaddock, L., Hu, L. *et al.* (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus*, **19**, 1030–1039. doi:10.1002/hipo.20547
- Erickson, K.I., Voss, M.W., Prakash, R.S., Basak, C., Szabo, A. *et al.* (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences, USA*, **108** (7), 3017–3022. doi:10.1073/pnas.1015950108
- Eriksson, P.S., Perfilieva, E., Björk-Eriksson, T., Alborn, A.-M., Nordborg, C. *et al.* (1998). Neurogenesis in the adult human hippocampus. *Nature Medicine*, **4**, 1313–1317. doi:10.1038/3305
- Faraco, C.C., Unsworth, N., Langley, J., Terry, D., Li, K. *et al.* (2011). Complex span tasks and hippocampal recruitment during working memory. *NeuroImage*, **55**, 773–787. doi:10.1016/j.neuroimage.2010.12.033
- Filipek, P.A., Richelme, C., Kennedy, D.N., & Caviness, V.S. (1994). The young adult human brain: an MRI-based morphometric analysis. *Cerebral Cortex*, **4**, 344–360. doi:10.1093/cercor/4.4.344
- Freedson, P.S., & Goodman, T.L. (1993). Measurement of oxygen consumption. In T.W. Rowland (Ed.), *Pediatric laboratory exercise testing: Clinical guidelines* (pp. 91–113). Champaign, IL: Human Kinetics.
- Frings, L., Wagner, K., Unterrainer, J., Spreer, J., Halsband, U. *et al.* (2006). Gender-related differences in lateralization of hippocampal activation and cognitive strategy. *NeuroReport*, **17** (4), 417–421. doi:10.1097/01.wnr.0000203623.02082.e3
- Gathercole, S.E., & Pickering, S.J. (2000). Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology*, **92**, 377–390. doi:10.1037/0022-0663.92.2.377
- Geary, D.C., Hoard, M.K., Byrd-Craven, J., & DeSoto, M.C. (2004). Strategy choices in simple and complex addition: contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, **88**, 121–151. doi:10.1016/j.jecp.2004.03.002
- Geiger, J.F., & Litwiller, R.M. (2005). Spatial working memory and gender differences in science. *Journal of Instructional Psychology*, **32** (1), 49–57.
- Giedd, J.N., Vaituzis, A.C., Hamburger, S.D., Lange, N., Rajapakse, J.C. *et al.* (1996). Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: ages 4–18 years. *Journal of Comparative Neurology*, **366**, 223–230. doi:10.1002/(SICI)1096-9861(19960304)366:2 < 223:AID-CNE3 > 3.0.CO;2-7
- Goldstein, J.M., Jerram, M., Poldrack, R., Anagnoson, R., Breiter, H.C. *et al.* (2005). Sex differences in prefrontal cortical brain activity during fMRI of auditory verbal working memory. *Neuropsychology*, **19**, 509–519. doi:10.1037/0894-4105.19.4.509
- Grissom, J.B. (2005). Physical fitness and academic achievement. *Journal of Exercise Physiology Online*, **8**, 11–25.
- Hallal, P.C., Andersen, L.B., Bull, F.C., Guthold, R., Haskell, W. *et al.* (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, **380** (9838), 247–257. doi:10.1016/S0140-6736(12)60646-1
- Hampson, E. (1990). Estrogen-related variations in human spatial and articulatory-motor skills. *Psychoneuroendocrinology*, **15**, 97–111. doi:10.1016/0306-4530(90)90018-5
- Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behavioral Neuroscience*, **102**, 456–459. doi:10.1037/0735-7044.102.3.456
- Hannula, D.E., Tranel, D., & Cohen, N.J. (2006). The long and the short of it: relational memory impairments in amnesia, even at short lags. *Journal of Neuroscience*, **26**, 8352–8359. doi:10.1523/JNEUROSCI.5222-05.2006
- 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. doi:10.1038/nrn2298
- Hillman, C.H., Pontifex, M.B., Castelli, D.M., Khan, N.A., & Raine, L.B. *et al.* (2014). Physical activity intervention improves cognitive and brain health in children. Manuscript submitted for publication.
- Hillman, C.H., Pontifex, M.B., Raine, L.B., Castelli, D.M., Hall, E.E. *et al.* (2009). The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. *Neuroscience*, **159**, 1044–1054. doi:10.1016/j.neuroscience.2009.01.057
- Hugdahl, K., Thomsen, T., & Ersland, L. (2006). Sex differences in visuo-spatial processing: an fMRI study of mental rotation. *Neuropsychologia*, **44**, 1575–1583. doi:10.1016/j.neuropsychologia.2006.01.026
- Institute of Medicine of the National Academies (2013). *Educating the student body: Taking physical activity and physical education to school*. H.W. Kohl III & H.D. Cook (Eds.). Washington, DC: The National Academies Press.
- Isgor, C., & Sengelaub, D.R. (1998). Prenatal gonadal steroids affect adult spatial behavior, CA1 and CA3 pyramidal cell morphology in rats. *Hormones and Behavior*, **34**, 183–198. doi:10.1006/hbeh.1998.1477
- Jaccard, J., & Turrissi, R. (2003). *Interaction effects in multiple regression* (2nd edn.). Thousand Oaks, CA: Sage Publications.
- Janowsky, J.S., Chavez, B., & Orwoll, E. (2000). Sex steroids modify working memory. *Journal of Cognitive Neuroscience*, **12** (3), 407–414. doi:10.1162/089892900562228
- Kamijo, K., Pontifex, M.B., O’Leary, K.C., Scudder, M.R., Wu, C.T. *et al.* (2011). The effects of an afterschool physical activity program on working memory in preadolescent children. *Developmental Science*, **14**, 1046–1058. doi:10.1111/j.1467-7687.2011.01054.x
- Kane, M.J., Conway, A.R.A., Miura, T.K., & Colflesh, G.J.H. (2007). Working memory, attention control, and the n-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **33** (3), 615–622. doi:10.1037/0278-7393.33.3.615
- Kane, M.J., & Engle, R. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general



- fluid intelligence: an individual-differences perspective. *Psychonomic Bulletin & Review*, **9**, 637–671.
- Kaufman, A.S., & Kaufman, N.L. (1990). *Kaufman Brief Intelligence Test manual*. Circle Pines, MN: American Guidance Service.
- Kaufman, A.S., & Kaufman, N.L. (2004). *Kaufman Brief Intelligence Test* (2nd edn.). Circle Pines, MN: AGS Publishing.
- Kaufman, S.B. (2007). Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity? *Intelligence*, **35**, 211–223. doi:10.1016/j.intell.2006.07.009
- Kawata, M. (1995). Roles of steroid hormones and their receptors in structural organization in the nervous system. *Neuroscience Research*, **24**, 1–46. doi:10.1016/0168-0102(96)81278-8
- Kimura, D., & Hampson, E. (1994). Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Current Directions in Psychological Science*, **3** (2), 57–61. doi:10.1111/1467-8721.ep10769964
- Klein, C., Rauh, R., & Biscaldi, M. (2010). Cognitive correlates of anti-saccade task performance. *Experimental Brain Research*, **203**, 759–764. doi:10.1007/s00221-010-2276-5
- Kohl, H.W. III, Craig, C.L., Lambert, E.V., Inove, S., Alkandari, J.R. *et al.* (2012). The pandemic of physical inactivity: global action for public health. *Lancet*, **380**, 294–305. doi:10.1016/S0140-6736(12)60898-8
- Kornack, D.R., & Rakic, P. (1999). Continuation of neurogenesis in the hippocampus of the adult macaque monkey. *Proceedings of the National Academy of Sciences, USA*, **96**, 5768–5773. doi:10.1073/pnas.96.10.5768
- Kramer, A.F., & Erickson, K.I. (2007). Capitalizing on cortical plasticity: influence of physical activity on cognition and brain function. *Trends in Cognitive Sciences*, **11** (8), 342–348. doi:doi.org/10.1016/j.tics.2007.06.009
- Kwak, L., Kremers, S.P.J., Bergman, P., Ruiz, J.R., Rizzo, N.S. *et al.* (2009). Associations between physical activity, fitness, and academic achievement. *Journal of Pediatrics*, **155**, 914–918. doi:10.1016/j.jpeds.2009.06.019
- Levine, S.C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, **35** (4), 940–949. doi:10.1037/0012-1649.35.4.940
- Linn, M.C., & Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Development*, **56**, 1479–1498. doi:10.2307/1130467
- Lu, B., & Chow, A. (1999). Neurotrophins and hippocampal synaptic transmission and plasticity. *Journal of Neuroscience Research*, **58**, 76–87. doi:10.1002/(SICI)1097-4547(19991001)58:1 < 76:AID-JNR8 > 3.0.CO;2-0
- McAllister, A.K., Katz, L.C., & Lo, D.C. (1999). Neurotrophins and synaptic plasticity. *Annual Reviews Neuroscience*, **22**, 295–318. doi:10.1146/annurev.neuro.22.1.295
- Maki, P.M., Rich, J.B., & Rosenbaum, R.S. (2002). Implicit memory varies across the menstrual cycle: estrogen effects in young women. *Neuropsychologia*, **40**, 518–529. doi:10.1016/S0028-3932(01)00126-9
- Masters, M.S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics*, **23**, 337–341.
- Moore, D.S., & Johnson, S.P. (2008). Mental rotation in human infants: a sex difference. *Psychological Science*, **19**, 1063–1066. doi:10.1111/j.1467-9280.2008.02200.x
- Murphy, D.G.M., DeCarli, C., McIntosh, A.R., Daly, E., Mentis, M.J. *et al.* (1996). Sex differences in human brain morphometry and metabolism: an in vivo quantitative magnetic resonance imaging and positron emission tomography study on the effect of aging. *Archives of General Psychiatry*, **53** (7), 585–594. doi:10.1001/archpsyc.1996.01830070031007
- National Association for Sport and Physical Education (2008). *Comprehensive school physical activity programs [position statement]*. Reston, VA: Author.
- Neeper, S.A., Gomez-Pinilla, F., Choi, J., & Cotman, C. (1995). Exercise and brain neurotrophins. *Nature*, **373**, 109. doi:10.1038/373109a0
- Neufang, S., Specht, K., Hausmann, M., Güntürkün, O., Herpertz-Dahlmann, B. *et al.* (2009). Sex differences and the impact of steroid hormones on the developing human brain. *Cerebral Cortex*, **19**, 464–473. doi:10.1093/cercor/bhn100
- Norman, D.A., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R.J. Davidson, G.E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory*, Vol. 4 (pp. 1–18). New York: Plenum Press.
- Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, **134**, 368–387. doi:10.1037/0096-3445.134.3.368
- Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H. *et al.* (2003). The neural basis of individual differences in working memory capacity: an fMRI study. *NeuroImage*, **18**, 789–797. doi:10.1016/S1053-8119(02)00032-0
- Owen, A.M., Evans, A.C., & Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. *Cerebral Cortex*, **6**, 31–38. doi:10.1093/cercor/6.1.31
- Öztekin, I., McElree, B., Staresina, B.P., & Davachi, L. (2009). Working memory retrieval: contributions of the left prefrontal cortex, the left posterior parietal cortex, and the hippocampus. *Journal of Cognitive Neuroscience*, **21** (3), 581–593. doi:10.1162/jocn.2008.21016
- Parsons, T.D., Larson, P., Kratz, K., Thiebaut, M., Bluestein, B. *et al.* (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia*, **42**, 555–562. doi:10.1016/j.neuropsychologia.2003.08.014
- Peper, J.S., Brouwer, R.M., Schnack, H.G., van Baal, G.C., van Leeuwen, M. *et al.* (2009). Sex steroids and brain structure in pubertal boys and girls. *Psychoneuroendocrinology*, **34**, 332–342. doi:10.1016/j.psyneuen.2008.09.012
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R. *et al.* (1995). A redrawn Vandenberg and Kuse mental rotations test – different versions and factors that affect performance. *Brain and Cognition*, **28**, 39–58. doi:10.1006/brcg.1995.1032
- Petrides, M. (1989). Frontal lobes and memory. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 601–614). Amsterdam and New York: Elsevier.

- 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. doi:10.1016/j.jpeds.2012.08.036
- Postma, A., Jager, G., Kessels, R.P.C., Koppeschaar, H.P.F., & Van Honk, J. (2004). Sex differences for selective forms of spatial memory. *Brain and Cognition*, **54**, 24–34. doi:10.1016/S0278-2626(03)00238-0
- Quinn, P.C., & Liben, L.S. (2008). A sex difference in mental rotation in young infants. *Psychological Science*, **19**, 1067–1070. doi:10.1111/j.1467-9280.2008.02201.x
- Roberts, R., & Gibson, E. (2002). Individual differences in sentence memory. *Journal of Psycholinguistic Research*, **31**, 573–598.
- Rogers, R.D., & Monsell, S. (1995). Cost of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, **124**, 207–231. doi:10.1037/0096-3445.124.2.207
- Rubia, K., Hyde, Z., Halari, R., Giampietro, V., & Smith, A. (2010). Effects of age and sex on developmental neural networks of visual-spatial attention allocation. *NeuroImage*, **51**, 817–827. doi:10.1016/j.neuroimage.2010.02.058
- Sallis, J.F. (1993). Epidemiology of physical activity and fitness in children and adolescents. *Critical Reviews in Food Science and Nutrition*, **33**, 403–408. doi:10.1080/10408399309527639
- Schinder, A.F., & Poo, M. (2000). The neurotrophin hypothesis for synaptic plasticity. *Trends in Neurosciences*, **23**, 639–645. doi:10.1016/S0166-2236(00)01672-6
- Schooler, C., Caplan, L.J., Revell, A.J., Salazar, A.M., & Grafman, J. (2008). Brain lesion and memory functioning: short-term memory deficit is independent of lesion location. *Psychonomic Bulletin & Review*, **15** (3), 521–527. doi:10.3758/PBR.15.3.521
- Scudder, M.R., Lambourne, K., Drollette, E.S., Herrmann, S.D., Washburn, R.A. *et al.* (2014). Aerobic capacity and cognitive control in elementary school-age children. *Medicine & Science in Sports & Exercise*, **46**, 1025–1035. doi:10.1249/MSS.0000000000000199
- Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, **171**, 701–703.
- Shrout, P.E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological Methods*, **7**, 422–445. doi:10.1037/1082-989X.7.4.422
- Sibley, B.A., & Etnier, J.L. (2003). The relationship between physical activity and cognition in children: a meta-analysis. *Pediatric Exercise Science*, **15**, 243–256.
- Sigfúsdóttir, I.D., Kristjánsson, A.L., & Allegrante, J.P. (2007). Health behaviour and academic achievement in Icelandic school children. *Health Education Research*, **22**, 70–80. doi:10.1093/her/cyl044
- Singh, M., Meyer, E.M., & Simpkins, J.W. (1995). The effect of ovariectomy and estradiol replacement on brain-derived neurotrophic factor messenger ribonucleic acid expression in cortical and hippocampal brain regions of female Sprague-Dawley rats. *Endocrinology*, **136**, 2320–2324.
- Sisson, S.B., Church, T.S., Martin, C.K., Tudor-Locke, C., Smith, S.R. *et al.* (2009). Profiles of sedentary behavior in children and adolescents: the US National Health and Nutrition Examination Survey, 2001–2006. *International Journal of Pediatric Obesity*, **4** (4), 353–359. doi:10.3109/17477160902934777
- Smith, E.E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, **283**, 1657–1661. doi:10.1126/science.283.5408.1657
- Sorkin, R.D. (1999). Spreadsheet signal detection. *Behavior Research Methods, Instruments, & Computers*, **31**, 46–54. doi:10.3758/BF03207691
- Speck, O., Ernst, T., Braun, J., Koch, C., Miller, E., & Chang, L. (2000). Gender differences in the functional organization of the brain for working memory. *NeuroReport*, **11**, 2581–2585. doi:10.1097/00001756-200008030-00046
- Tanapat, P., Hastings, N.B., Reeves, A.J., & Gould, E. (1999). Estrogen stimulates a transient increase in the number of new neurons in the dentate gyrus of the adult female rat. *Journal of Neuroscience*, **19** (14), 5792–5801.
- Tanner, J.M. (1962). *Growth at adolescence: With a general consideration of the effects of hereditary and environmental factors upon growth and maturation from birth to maturity*. Oxford: Blackwell Scientific Publishing.
- Taylor, S.J.C., Whincup, P.H., Hindmarsh, P.C., Lampe, F., Odoki, K. *et al.* (2001). Performance of a new pubertal self-assessment questionnaire: a preliminary study. *Paediatric and Perinatal Epidemiology*, **15**, 88–94. doi:10.1046/j.1365-3016.2001.00317.x
- Thomas, S., Reading, J., & Shephard, R.J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sport Science*, **17**, 338–345.
- Thomsen, T., Hugdahl, K., Ersland, L., Barndton, R., Lundervold, A. *et al.* (2000). Functional magnetic resonance imaging (fMRI) study of sex differences in a mental rotation task. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, **6**, 1186–1196.
- Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, **28**, 127–154. doi:10.1016/0749-596X(89)90040-5
- Utter, A.C., Robertson, R.J., Nieman, D.C., & Kang, J. (2002). Children's OMNI scale of perceived exertion: walking/running evaluation. *Medicine & Science in Sports & Exercise*, **34**, 139–144.
- van Praag, H., Christie, B.R., Sejnowski, T.J., & Gage, F.H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proceedings of the National Academy of Sciences, USA*, **96**, 13427–13431. doi:10.1073/pnas.96.23.13427
- van Vugt, M.K., Schulze-Bonhage, A., Litt, B., Brandt, A., & Kahana, M.J. (2010). Hippocampal gamma oscillations increase with memory load. *Journal of Neuroscience*, **30** (7), 2694–2699. doi:10.1523/JNEUROSCI.0567-09.2010
- Vandenberg, S.G., & Kruse, A.R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, **47**, 599–604. doi:10.2466/pms.1978.47.2.599

- Vecchi, T., & Girelli, L. (1998). Gender differences in visuospatial processing: the importance of distinguishing between passive storage and active manipulation. *Acta Psychologica*, **99**, 1–16. doi:10.1016/S0001-6918(97)00052-8
- Warren, D.E., Duff, M.C., Tranel, D., & Cohen, N.J. (2011). Observing degradation of visual representations over short intervals when medial temporal lobe is damaged. *Journal of Cognitive Neuroscience*, **23** (12), 3862–3873. doi:10.1162/jocn\_a\_00089
- Weiss, E., Siedentopf, C.M., Hofer, A., Deisenhammer, E.A., Hoptman, M.J. *et al.* (2003). Sex differences in brain activation pattern during a visuospatial cognitive task: a functional magnetic resonance imaging study in healthy volunteers. *Neuroscience Letters*, **344**, 169–172. doi:10.1016/S0304-3940(03)00406-3
- Wnuk, A., Korol, D.L., & Erickson, K.I. (2012). Estrogens, hormone therapy, and hippocampal volume in postmenopausal women. *Maturitas*, **73**, 186–190. doi:10.1016/j.maturitas.2012.07.001
- Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Rolling Meadows, IL: Riverside Publishing.
- World Health Organization (WHO) (2009). *Global health risks: Mortality and burden of disease attributable to selected major risks*. Retrieved 4 January 2014 from: [http://www.who.int/healthinfo/global\\_burden\\_disease/GlobalHealthRisks\\_report\\_full.pdf](http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf)

Received: 29 September 2014

Accepted: 22 December 2014