Long-Term High-Effort Endurance Exercise in Older Adults: Diminishing Returns for Cognitive and Brain Aging

Jeremy C. Young, Nicholas G. Dowell, Peter W. Watt, Naji Tabet, and Jennifer M. Rusted

(accepted copy before copy-editing)

Author Note:

Jeremy C. Young, Facultad de Ciencias Económicas y Administrativas, Pontificia Universidad Javeriana, Bogotá, Colombia;

Nicholas G. Dowell, Clinical Imaging Sciences Centre, Brighton and Sussex Medical School, Falmer, BN1 9RR, UK;

Peter W. Watt, Centre for Sport and Exercise Science and Medicine, School of Sport and Service Management, University of Brighton, Eastbourne, BN20 7SR, UK;

Naji Tabet, Institute of Postgraduate Medicine, Brighton and Sussex Medical School, University of Brighton, Brighton, BN1 9PH, UK;

Jennifer M. Rusted, School of Psychology, University of Sussex, Brighton, BN1 9QH, UK.

Abstract

While there is evidence that age-related changes in cognitive performance and brain structure can be offset by increased exercise, little is known about the impact on these of long-term high-effort endurance exercise. In a cross-sectional design with 12-month follow-up, we recruited older adults engaging in high-effort endurance exercise over at least twenty years, and compared their cognitive performance and brain structure with a non-sedentary control group similar in age, sex, education, IQ, and lifestyle factors. Our findings showed no differences on measures of speed of processing, executive function, incidental memory, episodic memory, working memory, or visual search for older adults participating in long-term high-effort endurance exercise, when compared without confounds to non-sedentary peers. On tasks that engaged significant attentional control, subtle differences emerged. On indices of brain structure, long-term exercisers displayed higher white matter axial diffusivity than their age-matched peers, but this did not correlate with indices of cognitive performance.

Key terms: Exercise; Aging; Cognition; MRI; Effort

1. Introduction

As we age, our cognitive abilities change and older adults have more difficulty than younger adults with tasks that are more complex, effortful, and strategic (for reviews see Luo & Craik, 2008; Salthouse, 2010). Contributing factors likely include slower processing speed (Salthouse, 1996, 2000), reduced attentional processing resources (Craik, 2006; McAvinue et al., 2012), loss of inhibitory functions (Darowski, Helder, Zacks, Hasher, & Hambrick, 2008; Peltsch, Hemraj, Garcia, & Munoz, 2011; Zanto, Rubens, Thangavel, & Gazzaley, 2011), and decline in controlled processing (Coubard et al., 2011).

These changes in cognitive ability accompany well-documented structural changes in the aging brain, specifically, cortical thinning and regional volumetric loss (for a review see Hedden & Gabrieli, 2004), and poorer white matter integrity (for a review see Madden et al., 2012). Volumetric changes in older adults can be detected over as little as 4 months in the hippocampus (Lövdén et al., 2012) and over just 6 months in many other brain regions (Raz et al. 2013), though there are substantive individual differences in trajectory of these age-related structural changes (Bennett, Madden, Vaidya, Howard, & Howard, 2010; Burzynska et al., 2010; Raz et al., 2013; Raz, Ghisletta, Rodrigue, Kennedy, & Lindenberger, 2010; Zhang et al., 2010).

Many factors influence the rate of cognitive and structural brain changes in older adulthood (Richards, Hardy, & Wadsworth, 2003; Salthouse, 2012; Scherder et al., 2013; Sturman et al., 2013; Thomas & O'Brien, 2008), and a particular target of research interest is the effect of exercise.

Notwithstanding the small effect sizes, the protective effects of exercise against age-related cognitive changes have been evidenced in retrospective (Middleton, Barnes, Lui, & Yaffe, 2010), cross-sectional (Benedict et al., 2012; Eskes et al., 2010; Nemati Karimooy, Hosseini, Nemati, & Esmaily, 2011), medium-term (Weuve et al., 2004) and longer-term studies (Barnes, Yaffe, Satariano, & Tager, 2003; Middleton, Mitnitski, Fallah, Kirkland, & Rockwood, 2008). In normal healthy older populations, moderate aerobic exercise programs leading to small improvements in cognitive ability, particularly executive function, have been reported in systematic reviews of interventions from 2 months to 6 years (Angevaren, Aufdemkampe, Verhaar, Aleman, & Vanhees, 2008; Chang, Labban, Gapin, & Etnier, 2012; Colcombe & Kramer, 2003; Etnier et al., 1997; Hindin & Zelinski, 2012; P. J. Smith et al., 2010). However a recent meta-analysis (Young, Angevaren, Rusted, & Tabet, 2015) that scrutinized and screened the data more closely reported little evidence for significant effects of exercise on cognition in older healthy adults.

Exercise also impacts age-related changes in brain structures (for reviews see Bherer, Erickson, & Liu-Ambrose, 2013; Hayes, Hayes, Cadden, & Verfaellie, 2013), with benefits from exercise reported on grey and white matter volume (Colcombe et al., 2003, 2006) and hippocampal volume (Erickson et al., 2009, 2011). The latter reporting that in groups of older adults, higher levels of aerobic fitness correlated both with higher hippocampal volumes and enhanced spatial memory function.

A limited amount of recent research has explored the sustained impact of long-standing exercise regimes, in older athletes. Winker et al. (2010) compared elderly marathon runners and bicyclists against a sedentary control group similar in age, sex, and

education. From a large battery of cognitive tasks, athletes performed better only in a non-verbal fluency task. Similarly, comparing cognitive performance in a small group of twelve older runners with sedentary older adults similar in age, sex, and education, Tseng, Uh, et al. (2013) reported advantages for their runners only on tasks of category and letter fluency. Structural brain differences were evidenced. Compared to sedentary control volunteers, the older athletes had some regional volumetric advantages in grey and white matter concentrations (Tseng, Gundapuneedi, et al., 2013). They suggested that long-term exercise may preserve white matter integrity from age-related changes.

An issue with many of the aforementioned behavioral and MRI cross-sectional studies is that they compared groups that were not expressly similar on measures of lifestyle, (cognitive activities and social interaction) or diet, which complicates interpretation of the data. In addition, most of the studies either were interventions with sedentary volunteers or used sedentary volunteers as the control group. Comparing groups similar in all key lifestyle measures is necessary to provide an accurate picture of the particular value of long-term exercise. Comparing long-term exercise against non-sedentary control groups is necessary to establish the impact of exercise against a normative "healthy" lifestyle.

Both of these criteria were met in the current study, which compared two older adult populations that were different only in their exercise profiles. The first group, super veteran athletes ("supervets"), comprised individuals whose age qualified them for super veteran categorization by UK Athletics, and who had engaged in long-term (minimum 20 years) high-effort endurance exercise. The second group, non-sedentary control volunteers, was active but not exercising beyond regular levels. The two groups were

similar in age, sex, education level, intelligence, depression scores, social network, cognitive activities, adherence to Mediterranean diet, and potential for physical fitness. More importantly, we collected longitudinal measures of cognitive performance and brain structural changes over a 12-month period, so that in our study we can establish not only cross-sectional differences between the groups but also whether age-related changes over 12 months differentiate them further.

We included a comprehensive battery of cognitive tasks, which have robustly been reported to change with aging and/or correlated with physical fitness or exercise levels. We included speed of processing tasks to explore whether exercise alleviated age-related slower processing speed (Salthouse, 1996, 2000). We included executive function measures since systematic reviews of previous exercise interventions observed differences in executive function (Angevaren et al., 2008; Chang et al., 2012; Colcombe & Kramer, 2003; Etnier et al., 1997; Hindin & Zelinski, 2012; P. J. Smith et al., 2010).

We included a prospective memory tasks because the effect of exercise on prospective memory has not been reported yet to our knowledge. Focal and non-focal prospective memory theoretically utilize different cognitive processes for optimal performance (McDaniel & Einstein, 2000); both were included as exercise may affect these processes differently. In attentional control tasks, such as the Stroop task, which involves suppressing attention and switching attention, older adults have been shown to perform worse than younger adults (Coubard et al., 2011), yet also show exercise-related improvement (Winker et al. 2010; Prakash et al., 2011). However, two meta-analyses of exercise intervention RCTs that included standard Stroop tasks as a measure, observed no differences in this task from the exercise interventions (P. J. Smith et al., 2010; Young et

al., 2015), and in a recent exercise RCT, only in most difficult modified Stroop condition did the exercise improve performance (Predovan, Fraser, Renaud, & Bherer, 2012). Therefore in our study, we included a more difficult Stroop task that included a switching component (Hutchison, Balota, & Ducheck, 2010), and a sustained attention task to test control over time.

Structurally, well-documented changes occur in the aging brain: cortical thinning and regional volumetric loss are observed (for reviews see Hedden & Gabrieli, 2004; Raz, 2000), and white matter integrity deteriorates (for reviews see Gunning-Dixon, Brickman, Cheng, & Alexopoulos, 2009; Madden et al., 2012; Madden, Bennett, & Song, 2009).

Specifically for white matter integrity, it has been observed that mean diffusivity of water molecules within the white matter (MD) increases with age, while the fractional anisotropy of that diffusion (FA) decreases. Decline in FA is greater in anterior regions of the brain than in posterior regions (Madden et al., 2012). More recent studies have started looking more at the specific diffusivity measures, axial diffusivity ($\lambda 1$) and radial diffusivity (RD). Axial diffusivity has been related to the myelin content of white matter tissue, while RD has been related to amount of demyelination and is proposed to be a marker of overall tissue integrity (Klawiter et al., 2011).

Volumetrically, in a typical study, Resnick, Pham, Kraut, Zonderman, & Davatzikos (2003) looking longitudinally at healthy older adults 59-85 observed significant decline in both grey and white matter volume from their first follow-up of 2 years even in a very healthy subgroup that did not develop any medical conditions or cognitive impairments during the 4-year period of evaluation. Declines in brain volume

can be detected in as little as a year (Erickson et al., 2011; Fjell et al., 2009). In longitudinal studies, volumetric changes can be detected over as little as 4 months in both healthy younger (20-30) and older adults (60-70) in the hippocampus (Lövdén et al., 2012). Raz et al. (2013) in a longitudinal study of healthy younger (20-31) and older adults (65-80) over just 6 months, observed significant decline in volume of lateral prefrontal cortex, hippocampus, caudate nucleus, and cerebellum.

Individual differences in trajectory of these brain volume changes vary significantly though, as seen in both Raz et al. (2010) and Raz et al. (2013). Also many factors are likely to influence the rate of structural brain changes in older adulthood (Bickart, Wright, Dautoff, Dickerson, & Barrett, 2011; Kanai, Bahrami, Roylance, & Rees, 2012; Schinka et al., 2010). One such factor is exercise.

Exercise may induce positive change in the aging brain or at least decrease the trajectory of decline in key areas (for reviews see Bherer, Erickson, & Liu-Ambrose, 2013; Hayes, Hayes, Cadden, & Verfaellie, 2013). To highlight, Erickson et al. (2011) reported an RCT with older adults 55-80 years, observing that a 1-year aerobic exercise intervention increased hippocampal volume, while in their stretching control group hippocampal volume decreased. Greater increases in fitness were correlated with greater increases in hippocampal volume, and there was a positive relationship between higher aerobic fitness levels and spatial memory and between increased hippocampal volume and improved spatial memory. Interestingly, while hippocampal volume declined in the stretching group, higher fitness before the intervention was protective against volume loss and partially attenuated the decline in volume.

Recent research has explored longer-term benefits of exercise in older athletes. Tseng, Uh, et al. (2013) reported a cross-sectional study that compared a small group of twelve older runners with sedentary older adults also similar in age, sex, and education. VBM measures of brain volume indicated that their older athletes had more grey matter concentrations in right parietal lobe, cuneus, and the culmen of the cerebellum and more white matter concentrations in precuneus, subgyral occipital lobe, and inferior temporal subgyral temporal lobe. Their runners also did better than sedentary older adults in executive function tasks of category fluency and letter fluency. In a similar cohort comparing ten older runners with twelve sedentary older adults, Tseng, Gundapuneedi, et al. (2013) reported their runners had better white matter integrity, as indexed by TBSS (Tract-based spatial statistics, S. M. Smith et al. 2006). In fractional anisotropy (FA) and mean diffusivity (MD) measures, they observed differences in brain regions associated with motor function; front-and-back connections related to visuospatial function, motor control and coordination; and regions associated with memory function. They also observed that runners had lower deep white matter hyperintensity volume. From this evidence, the authors suggested that long-term exercise may preserve white matter integrity from age-related changes. While participants in these older athlete studies were similar in age, sex, and education level, their cross-sectional design precluded unequivocal attribution of the observed differences to the volunteers' exercise regimen.

The current study set out to explore both brain and behavioral changes associated with long-term high-effort endurance exercise. We anticipated that our supervet group when compared to a non-sedentary socially active control group would show benefits both to cognition and brain structure as a result of their long-term exercise regime. We

hypothesized that long-term high-effort endurance exercise would protect against the age-related decline in executive function, attention, and memory, including prospective memory. We predicted the attenuation of age-related decline in attentional capabilities (Prakash et al., 2011; Winker et al., 2010), and that this would be particularly pronounced in the more difficult Stroop-switch task (Predovan et al., 2012). We predicted that exercise would confer slower decline in tasks requiring greater resource, specifically exposing exercise-related benefits to the resource-needy non-focal PM conditions (McDaniel, Shelton, Breneiser, Moynan, & Balota, 2011). We anticipated exercise-related benefits on measures of episodic memory, incidental memory and working memory, in line with previous literature (Erickson et al , 2009) . We also hypothesized that long-term high-effort endurance exercise would protect against the age-related decline in brain structure. We predicted that supervets would show greater brain volume than controls (Bherer et al., 2013; Colcombe et al., 2003, 2006; Hayes et al., 2013; C Niemann, Godde, Staudinger, & Voelcker-Rehage, 2014; Ruscheweyh et al., 2011), especially in the hippocampus (Erickson et al., 2009, 2011; Lövdén et al., 2012; Claudia Niemann, Godde, & Voelcker-Rehage, 2014; Raz et al., 2010; Uylings & de Brabander, 2002). We also predicted that supervets would have better white matter integrity than controls, since poorer integrity is observed in older adults and is recognized as an index of neural efficiency (eg. Burzynska et al., 2010). In normal aging declines in brain structure can be detected in as little as 12 months (Erickson et al., 2011; Fjell et al., 2009), so the study incorporated a 12-month longitudinal retest of both

performance and brain structure to test for observable differences slope of change between the non-sedentary and supervet groups.

2. Materials and Methods

Ethics Statement

Ethical approval was obtained from the University of Sussex Life Sciences & Psychology Cluster based Research Ethics Committee and the Brighton and Sussex Medical School Research Governance & Ethics Committee. Written consent was obtained from all participants.

2.1 Participants

Participants were non-smokers (never smoked or not smoked in the past 5 years), aged 60-85, on stable or no medication during the past 12 months, with English proficiency equivalent to that of a native speaker, and with no evidence of memory problems (MMSE > 27; Folstein, Folstein, & McHugh, 1975).

Exclusion criteria: history of stroke, myocardial infarction, recently diagnosed diabetes, high blood pressure (systolic above 200 and diastolic above 100), psychiatric or neurological disorders (self-reported), or clinical depression (Geriatric Depression Scale (GDS); Sheikh & Yesavage, 1986).

For our supervet group we required high-effort endurance exercise via running, swimming, and/or cycling (self-paced sports) for 20 or more years. We chose only self-paced sports because athletes of interceptive-dominant sports have been shown to have faster reaction times (Voss, Kramer, Basak, Prakash, & Roberts, 2010) which may affect cognitive performance. The Physical Activity Scale for the Elderly (PASE; Washburn, Smith, Jette, & Janney, 1993) was used to assess physical activity levels in both groups. Supervets were assured to be supervets through recruitment screening (participating

presently in routine high-effort endurance exercise and for at least the past 20 years) but also in the analysis of the PASE which showed a differentiation between our groups on this scale.

We recruited 27 supervets (6 female) and 23 control volunteers (9 female). At 1-year follow-up 23 supervets (6 female) and 19 controls (7 female) returned. From the sample recruited for behavioural testing, 15 supervets (4 female) and 14 controls (5 female) attended at baseline a session for acquisition of structural brain images, and 15 supervets (4 female) and 12 controls (4 females) returned at 1-year follow-up.

2.2 Demographic measures

The following measures were collected:

IQ: Full-scale pre-morbid IQ (National Adult Reading Test; Nelson, 1982).

Physical Activity: Physical Activity Scale for the Elderly (PASE; Washburn, Smith, Jette, & Janney, 1993), for measuring physical activity/exercise.

Diet: EPIC-Norfolk Food Frequency Questionnaire (FFQ; Bingham et al., 2001), from which we derived an adherence to Mediterranean diet score (Trichopoulou, Costacou, Bamia, & Trichopoulos, 2003).

Social Network: Lubben Social Network Scale (LSNS-6; Lubben et al., 2006).

Frequency of Cognitive Activities: Florida Cognitive Activities Scale (FCAS; Schinka et al., 2005).

2.3 Physiological Indices

Lung function measures taken using a spirometer (Microplus, Micro Medical Limited, Kent, UK) included: Forced Expiratory Volume 1st Second (FEV1), Forced Vital Capacity (FVC), Forced Expiratory Ratio (FER), and Peak Expiratory Flow (PEF).

Bioelectrical impedance assay (Bodystat Quadscan 4000 or Bodystat 1500; Douglas, Isle of Man, UK) provided percent body fat. Hand-grip strength was acquired using a hand dynamometer (Grip-D, Takei Scientific Instruments, Japan).

2.4 Cognitive Tasks

Simple Reaction Time (SRT)

The Digit Symbol Substitution Task (DSST, Wechsler, 1981)

Trail Making Test A and B (Reitan, 1958)

Controlled Oral Word Association Test (COWAT, Spreen & Strauss, 1991)

20-Word Item Episodic Memory Task (Rusted & Warburton, 1989)

Backward Digit Span (Wechsler, 1981)

Prospective and Retrospective Memory Questionnaire (PRMQ; Crawford, Smith, Maylor, Della Sala, & Logie, 2003)

Prospective Memory Card Sort Task (Rusted, Sawyer, Jones, Trawley, & Marchant, 2009)

Focal and Non-Focal Prospective Memory Task (McDaniel et al., 2011)

Map Test of Everyday Attention (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996)

Rapid Visual Information Processing (RVIP; Wesnes & Warburton, 1983)

Stroop-Switch Task (Hutchison et al., 2010)

Speed of Processing & Executive Function

Simple Reaction Time (SRT) was used for assessing simple reaction time and calculating intra-individual RT variability (Bunce et al., 2007). Here participants made a button press as soon as they could while maintaining accuracy when they saw a fixation cross turn into an "X" which was presented at a randomly determined interval between 300 and 1000ms. This task consisted of 8 practice trials and 48 test trials. The computer recorded the reaction time to each button press.

The Digit Symbol Substitution Task (DSST, Wechsler, 1981) - Symbol Copy provided a speed of processing component, and subtracting mean time per item in Symbol Copy from the Digit Symbol Substitution yielded a measure of higher mental function (Glosser et al., 1977; Storandt, 1976 as cited in Joy, Fein, & Kaplan, 2003). We also included the incidental memory portion measuring immediate recall of symbol pairings.

Trail Making Test A and B (Reitan, 1958) - Part A provided a speed element, while the difference score provided an estimate of executive function.

Controlled Oral Word Association Test (COWAT, Spreen & Strauss, 1991) was used to measure verbal fluency.

Memory

Episodic Memory:

Episodic memory was measured using a 20-item word list on the computer

screen at the rate of one word every two seconds, followed by an immediate written recall.

Working Memory:

The Backward Digit Span (Wechsler, 1981) was used to assess working memory during the follow-up only.

Prospective memory (PM):

Subjective Memory Rating

Prospective and Retrospective Memory Questionnaire (PRMQ; Crawford et al, 2003).

Event-based PM Tasks

1. Card Sort Task (Rusted et al., 2009)

Participants sorted playing cards into hearts and spades, while ignoring diamonds and clubs. The non-focal PM intention was to press the spacebar for number 7 cards regardless of suit. In each trial the back of the playing card was shown for 1000ms and then the face card was shown for 750ms, automatically advancing between front and back and the next trial. The baseline condition (without PM intention) had 52 trials (one deck), while the PM condition had 104 trials (two decks).

2. Focal and Non-Focal Prospective Memory Task (McDaniel et al., 2011)

The ongoing task was a category decision task. The three counterbalanced PM conditions were: 1) a no PM control condition; 2) a focal PM condition – the PM cue was a specific word, i.e. Tortoise; 3) a non-focal PM condition – the PM cue was part of a word, i.e. “tor” as found in words like his**tor**y. PM targets for the focal and

non-focal conditions were presented 3 times during their appropriate condition and not presented again in the other conditions. Participants practiced the ongoing category decision task first, which included six trials giving speed and accuracy feedback to encourage optimization of both. Then three conditions consisting of 106 word-pair trials (not including prospective memory trials) were used.

Attention

The Map Test of Everyday Attention (Map TEA; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996) provided a visual search measure. Participants circled as many petrol stations as they could find on a complex version of the standard map, which contained other symbols as well, for two minutes, switching colors at the one minute interval.

Rapid Visual Information Processing (RVIP; Wesnes & Warburton, 1983) provided a sustained attention measure. The computer recorded responses and reaction times. In our version we presented one stimulus per second. We had five 80-second intervals, with eight target sequences presented in each interval. The total time for the task was 6:20.

The Stroop-Switch Task (Hutchison et al., 2010) combines the traditional Stroop Task with a task-switching paradigm, intermixing trials requiring color naming and word naming, switching the task every two trials. Trials advanced when participants made a vocal response and there were 24 practice trials and 88 trials. The computer recorded reaction times, while the experimenter recorded errors.

2.5 Procedure

Data collection

Participants' cognitive assessment was completed in the School of Psychology and MRI sessions at the Clinical Imaging and Sciences Centre, at Sussex University. Time intervals between MRI and cognitive sessions at baseline and follow-up were as closely linked as possible, allowing for the individual availability of the volunteers.

Participants were consented by a researcher on each occasion. The cognitive testing sessions lasted approximately 3 hours, including breaks as and when volunteers wanted. MRI scans took less than 30 minutes. Participants were compensated for their transport costs and parking.

MRI protocol

All images were acquired on a Siemens 1.5 Tesla Avanto MRI scanner (Siemens, Erlangen, Germany). High-resolution anatomical images were acquired using a three-dimensional T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence [TR = 1160 ms; TE= 4.44 ms; inversion recovery time (TI) = 600 ms; field of view (FOV), 230 x 230 mm^2; matrix size, 256 x 256; flip angle θ = 15 degrees; voxel dimensions, 0.9 x 0.9 x 0.9 mm^3; acquisition time, 5 min].

Diffusion-weighted images were acquired using an echo planar imaging sequence [TR = 12.4 s; TE = 111 ms; echo spacing, 0.83 ms; FOV, 240 x 240 mm^2; matrix size, 96 x 96; voxel dimensions, 2.5 x 2.5 x 2.5 mm^3; acquisition time, 7 min]. Diffusion gradients were applied along 30 noncollinear directions (b_{max} = 1000 s/mm^2). A nondiffusion-weighted (b ~ 0) volume was also acquired.

2.6 Analyses

For all statistical tests an alpha of .05 was adopted.

2.5.1 Demographics & Experimental Tasks

For demographic comparisons, we used independent t-tests to compare groups (supervets vs. controls) at each time point. For lifestyle factors and cognitive measures at the baseline time point we used independent t-tests between groups to compare means. ANOVAs were used if a task had a repeated measures factor. Follow-up analyses used mixed factor ANOVAs with time point as a repeated measures factor. Any post-hoc tests used a Sidak correction and therefore Sidak-adjusted p-values are reported.

2.5.2 *Image post-processing*

We used whole volume analysis (Erickson et al., 2009, 2011; Resnick et al., 2003), automated segmentation (Resnick et al., 2003), Voxel-Based Morphometry (VBM) (Tseng, Uh, et al., 2013), longitudinal VBM (Colcombe et al., 2003, 2006), and Tract-Based Spatial Statistics (TBSS) (Tseng, Gundapuneedi, et al., 2013).

Raz et al. (2013, 2010) used manual tracing to segment their volumes, but automated segmentation has been shown to be highly correlated in volume and shape with manual tracing especially using Freesurfer (Morey, Petty, & Xu, 2009) with the additional benefits of no operator bias and taking less operator time. For these reasons we used Freesurfer to autosegment our volumes. For volume and cortical thickness analyses we looked specifically at areas most previously seen to be affected by aging and/or were previously shown to be different for those that exercised. Using Freesurfer, we were able to do both these analyses within the same software package. For volumes we chose hippocampus (Erickson et al., 2009, 2011; Lövdén et al., 2012; Raz et al., 2010; Uylings & de Brabander, 2002), cerebellar cortex (Raz et al., 2013, 2010; Tseng, Uh, et al., 2013),

cerebellar white matter (Raz et al., 2013, 2010; Tseng, Uh, et al., 2013), and caudate (Raz et al., 2013, 2010). For cortical thickness we looked at entorhinal cortex (Raz et al., 2010), frontal lobes (Colcombe et al., 2003, 2006; Raz et al., 2013; Resnick et al., 2003; Uylings & de Brabander, 2002), and parietal lobes (Colcombe et al., 2003; Resnick et al., 2003; Tseng, Uh, et al., 2013).

With our Diffusion Tensor Imaging (DTI) measures, we also used the more recently developed histogram analysis, which is sensitive to subtle diffuse differences in the brain in the specific diffusivity measures of mean diffusivity (MD), axial diffusivity (λ1) and radial diffusivity (RD) (Tofts, Davies, & Dehmeshki, 2003), whereas TBSS detects localized differences (S. M. Smith et al., 2006).

DTI

Histogram analysis:

Histogram analysis is sensitive to subtle diffuse disease (Tofts et al., 2003) and aging (Charlton et al., 2006). For example, comparing patients with multiple sclerosis (MS) to controls without MS, Cercignani, Bozzali, Iannucci, Comi, & Filippi (2001) observed diffuse tissue damage in the white matter of patients with MS, in the form of increased overall MD values. Looking at aging, Charlton et al. (2006) observed cross-sectionally middle-aged and older adults and found progressive reduction in FA and increase in MD with age, which correlated with declines in cognitive measures. DTI data were pre-processed using the FMRIB's Software Library (FSL) suite version 5.04 (S. M. Smith et al., 2004; Woolrich et al., 2009). DTI volumes were realigned using the eddy current correction module tool. The diffusion tensor was calculated for each image voxel

to give MD, FA, RD, λ1. To restrict analyses we generated WM and GM regions of interest (ROIs) using the New Segment tool in SPM (Statistical Parametric Mapping, Functional Imaging Laboratory, University College London).

The frequency distribution of each magnetic resonance (MR) parameter was plotted and the resulting histogram was normalized to 20 units and smoothed using a gaussian kernel. Smoothing kernels were: 0.009 and 0.02 for grey matter and white matter FA, respectively; 20×10^{-6} mm^2/s and 10×10^{-6} mm^2/s for grey matter and white matter respectively for MD, λ1, and RD. We then calculated peak height (PH), peak position (PP), and mean for each histogram per ROI per DTI parameter.

MR parameter indices (PH, PP, Mean) were compared at each time point using ANCOVAs between groups (supervet vs. control), $p < .05$, with age and sex (Hsu et al., 2008; Kanaan et al., 2012) as covariates.

Tract-based Spatial Statistics (TBSS):

Tract-based spatial statistics (TBSS) (S. M. Smith et al., 2006), part of FSL, was performed on DTI data following the procedure outlined in (Dowell et al., 2013) to generate TBSS maps of FA, MD, λ1, and RD. TBSS allows for detection of differences in DTI indices in local white matter tracts.

TBSS inference testing was employed to identify differences between supervet and control groups. As the null distribution was not known, a permutation-based inference method was employed (Nichols & Holmes, 2002). A two-sample nonparametric test was carried out with 10000 permutations using a threshold-free cluster enhancement (TFCE) method. In contrast with other such techniques, this method has the advantage that no arbitrary initial cluster-forming threshold needs to be selected. The

computed cluster is tested against a critical cluster size: if the resulting cluster is larger than the critical size, the cluster is significant.

Anatomical

Whole brain volume analysis:

Whole volume analysis allowed us detect global differences. MPRAGE volumes were segmented into white matter (WM), grey matter (GM), and cerebral spinal fluid (CSF) ROIs. This was added together to get the intracranial volume (ICV). Whole brain (WB) volume was calculated by adding WM and GM. Then each ROI – WM, GM, WB – was divided by the ICV to calculate the percentage of the whole volume each ROI took. All data was then compared using two-tailed t-tests between groups, $p < .05$ at baseline. When including the follow-up time point, mixed ANCOVAs were used with time point as a repeated measures factor, groups as an independent measures factor, and age and sex (Good et al., 2001) as covariates.

Voxel-based morphometry (VBM):

VBM allowed an unbiased voxel-by-voxel analysis of the whole brain without the need to specify a priori ROIs. For baseline VBM, we used Christian Gaser's VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm/) for the SPM8 package (Wellcome Department of Cognitive Neurology, London, UK (J Ashburner & Friston, 2000)) in Matlab R2007a (MathWorks, Natick, MA, USA). We skull stripped and warped T1-weighted anatomical images into MNI space and segmented GM, WM and CSF within a unified segmentation model (John Ashburner & Friston, 2005). Modulated normalized segmentations of GM and WM were then smoothed using an 8x8x8-mm^3 kernel. Two-

sample t-tests were defined to detect significant group volume differences. The family-wise error correction was carried out at $p < 0.05$.

Including follow-up, we used VBM longitudinal analysis in VBM8. Follow-up participants' volumes were non-linearly registered to their own baseline volumes and spatial normalization is estimated for baseline volumes only and applied to all time points. Mixed ANOVAs were defined to detect group volume differences, time point volume differences, and interactions between the two. The family-wise error correction was carried out at $p < 0.05$.

Freesurfer volume and cortical thickness:

Freesurfer allowed using specific surface-based and volumetric ROI analysis using a priori determined ROIs. Both time points were automatically processed using the longitudinal stream in Freesurfer 5.3.0 (Reuter, Schmansky, Rosas, & Fischl, 2012).

For volume and cortical thickness analyses we looked specifically at areas most reported to be affected by aging and/or were previously shown to be different for those that exercised. For volumes we chose hippocampus (Erickson et al., 2009, 2011; Lövdén et al., 2012; Raz et al., 2010; Uylings & de Brabander, 2002), cerebellar cortex (Raz et al., 2013, 2010; Tseng, Uh, et al., 2013), cerebellar white matter (Raz et al., 2013, 2010; Tseng, Uh, et al., 2013), and caudate (Raz et al., 2013, 2010). For cortical thickness we looked at entorhinal cortex (Raz et al., 2010), frontal lobes (Colcombe et al., 2003, 2006; Raz et al., 2013; Resnick et al., 2003; Uylings & de Brabander, 2002), and parietal lobes (Colcombe et al., 2003; Resnick et al., 2003; Tseng, Uh, et al., 2013).

Volumes for each hemisphere across subjects and between time points were extracted and divided by ICV; age and sex were used as covariates and groups were

compared using ANCOVAs at the baseline time point, $p < .05$. For the follow-up time point mixed ANCOVAs were used with time point as a repeated measures factor, group as an independent measures factor, and age and sex as covariates.

Baseline cortical thickness and cortical thickness symmetrized percent change (SPC), which divides the rate of change by average thickness over the two points, giving more statistical power than percent change relative to time point 1 thickness (since average thickness is less noisy than thickness at time point 1), were extracted. At baseline, age and sex were used as covariates, and groups were compared using ANCOVAs, $p < .05$. With the follow-up cortical thickness SPC, mixed ANCOVAs were used with time point as a repeated measures factor, group as an independent measures factor, and age and sex as covariates.[1]

Correlations with cognitive data

For the purposes of these analyses, correlations were computed using between group differences found in cognitive data and group differences in MRI data. These were for cognitive data: Reaction Time cost in a PM task, and Switch RT in a modified Stroop task.[2] For MRI measures we entered follow-up White Matter Axial Diffusivity, percentage white matter change, and percentage whole brain change, right hippocampal volume percent change, left and right cerebellar volume percent change, and right frontal lobe cortical thickness symmetrized percent change. Correlations were all corrected for multiple comparisons using a bonferonni correction.

3. Results

3.1 Demographics

3.1.1 Demographics

Results for Demographics are shown in Table I.

There were no significant differences between groups at either the initial and follow-up time points in sex, age, education, IQ, or depression levels.

3.1.2 Physical Activity

Results for Physical Activity are in Table II: Physical Activity.

As expected and desired, on the PASE the supervets took part in more physical activities than the controls at baseline and this difference was maintained in both baseline and follow-up analyses. Specifically, at the baseline time point supervets took part in significantly more strenuous sports, $t(34.58) = 4.59$, $p < .001$, and muscle strength and endurance activities, $t(30.41) = 2.78$, $p = .009$. Including the follow-up time point in the analysis, there was a group by leisure activity interaction, $F(3.09, 117.33) = 6.20$, $p = .001$. Again supervets ($M = 1.08$ hours per day) participated in more strenuous sports than controls ($M = .20$), $F(1, 38) = 27.93$, $p < .001$. Also supervets ($M = .12$) participated in more muscle strength and endurance activities than controls ($M = .03$), $F(1, 38) = 8.88$, $p = .005$. The groups did not differ in time spent on other leisure activities. For both groups, there was an increase in moderate activity at follow-up ($M = .28$ hours per day) vs. baseline ($M = .12$), $F(1, 38) = 4.31$, $p = .045$, while other activities did not change.[3]

3.1.3 Diet

Results for Diet are in Table II: Diet.

There were no differences between groups in adherence to Mediterranean diet at the initial or follow-up time point.

3.1.4 Lung Function

Results for Lung Function are in Table II: Lung Function.

At initial and including follow-up there were no significant differences between groups in lung function measures: FEV1, FVC, FER, and PEF. Lung function measures are not affected by exercise training (Dempsey, 1986), and comparability of groups indicates that both have similar potential for cardiovascular performance and that both groups were capable of achieving the same level of fitness.

At follow-up, FVC had declined in both groups, $F(1, 30) = 5.62$, $p = .024$, and PEF increased, $F(1, 33) = 4.97$, $p = .033$.

3.1.5 Physiological

Results for Physiological Measures are in Table II: Physiological.

Because of physiological differences between males and females, these measures were analyzed by gender.

In physiological measures, expected differences between our groups were confirmed at the initial time point and at follow-up. Supervets had significantly less percent body fat in both sexes. Male supervets trended towards stronger hand-grip than male controls at the initial time point and this difference was significant at follow-up. Females did not differ in hand-grip strength at either time point. At follow-up, percentage fat increased in both groups for both sexes: males, $F(1, 24) = 23.25$, $p < .001$; females, $F(1, 11) = 5.46$, $p = .039$.

3.1.6 Social and Cognitive Activities

Results for Social and Cognitive Activities are in Table II: Social and Cognitive Activities.

There were no significant differences between the groups in LSNS score or FCAS score either at the initial time point or at follow-up, $Qs < 1$.

3.2 Experimental Tasks

3.2.1 Speed of Processing

Results are shown in Table III: Speed of Processing.

At the initial time point there were no significant differences between groups for the mean reaction time for correct trials in the Simple Reaction Time task. When including the follow-up time point in analyses, there was still no main effect of group, $F(1, 38) = 3.59$, $p = .066$.

For both Intra-Individual Variability (IIV) and Intra-Individual Coefficient of Variability (ICV) there were no significant differences between groups at both the initial time point and including the follow-up time point. IIV was significantly higher at follow-up than at the initial time point, $F(1, 38) = 4.21$, $p = .047$.

There were no significant differences between groups for the Trail Making Test part A (Trails A) or Symbol Copy task at both the initial time point and including follow-up, $Qs < 1$.

3.2.2 Executive Function

Results are shown in Table III: Executive Function.

For DSST higher mental function, there were no differences between groups at the initial time point (power = .29) and when including follow-up, $F < 1$; power: between factors = .16, within factors = .51, within-between interaction = .11. For Trails B, accounting for motor speed, there were no differences between groups at the initial time point (power = .58) or when including follow-up, $Q < 1$.

In the COWAT, there were no significant differences between supervets and controls in verbal fluency at the initial time point (power = .24) nor when including the follow-up time point, $F < 1$; power: between factors = .09, within factors = .98, within-between interaction = .99.

3.2.3 Memory

Results are shown in Table III: Memory.

Incidental and Episodic Memory:

There were no main effects between groups for incidental memory performance (from the DSST) nor for episodic memory performance at the initial time point nor when including the follow-up time point, $F(1, 34) = 1.46$, $p = .236$ and $F < 1$ respectively.

Working Memory:

At follow-up, there was no main effect between groups for the Backwards Digit Span.

<u>Prospective Memory:</u>

Subjective Rating of Memory

Results for this measure are shown in Table III: Memory.

For the Prospective and Retrospective Memory Questionnaire, there were no differences between groups for ratings by supervets and controls of their own prospective and retrospective memory. When including the follow-up time point, there were also no differences.

Card Sort PM Task:

Results for this task are shown in Table IV.

There were no differences between groups at either time point on PM accuracy, but at the initial time point, performance on the ongoing (card-sorting) task was differentially affected by the requirement to carry a PM intention, with supervets making more errors than controls, $t(30.94) = 2.17$, $p = .038$: supervets, $M = 3.66\%$, controls, $M = 1.09\%$, $M_{diff} = 2.66$, 95% CI [0.33, 4.86], $p = .031$. Reaction time to complete the ongoing task was not different between groups in either condition at either time point.

McDaniel PM Task:

Results for this task are shown in Table V.

There were no differences between groups in either condition at either time point on PM accuracy. There were no differences between groups in any condition at either time points in accuracy of the ongoing task (category decision) but there were differences in reaction times. A condition by group linear interaction trend, $F(1, 38) = 3.61$, $p = .065$, was explored with Sidak-corrected pairwise comparisons. As in the card sort PM, supervets, were more impacted when required to carry a PM intention, with longer RTs when maintaining either focal ($p = .011$) or non-focal ($p < .001$) PM intentions. For control volunteers, ongoing task RTs were longer only when maintaining a non-focal PM intention ($ps < .001$).

3.2.4 Attention

Map TEA:

Results are shown in Table III: Attention.

There were no significant differences between groups at initial time point or when including follow-up.

Stroop Switch Task:

Results for this task are shown in Table VII.

The standard Stroop effects were replicated but did not differentiate our groups.

Errors

There was a group by time point interaction, $F(1, 36) = 4.37$, $p = .044$, $n^2p = .108$, qualified by a group by time point by cue interaction, $F(1, 36) = 8.04$, $p = .007$, $n^2p = .183$: sidak-corrected pairwise comparisons confirmed that only in supervets at the follow-up time point did color naming produce more errors than word naming ($p = .008$, $d_z = 0.63$, power = .79).

Reaction Time

There was a cue by switch interaction, $F(1, 46) = 4.99$, $p = .030$, $\eta^2p = .098$, qualified by a group by cue by switch interaction, $F(1, 46) = 5.92$, $p = .019$, $n^2p = .114$: sidak-corrected pairwise comparisons confirmed that reaction time difference were limited to word-naming trials, where 1) supervets were faster than controls in switch trials only ($p = .046$, $d = 0.61$, power = .67) and 2) switch trials were slower

than non-switch trials for controls only, $F(1, 21) = 30.89$, $p < .001$, $n^2p = .595$. However this didn't appear to have any obvious theoretical significance.

Including follow-up, there was a cue by switch interaction, $F(1, 36) = 9.60$, $p = .004$, $n^2p = .210$, qualified by a group by cue by switch interaction, $F(1, 36) = 4.65$, $p = .038$, $n^2p = .114$, qualified by group by cue by switch by time point interaction, $F(1, 36) = 8.03$, $p = .007$, $n^2p = .182$, attributable to the initial time point already described; the interaction at the initial time point was not replicated at follow-up.

RVIP:

Results for this task are shown in Table VI.

Number of Correct Responses

There were no group differences; for both groups, performance deteriorated over time at initial, $F(4, 184) = 8.11$, $p < .001$, and including follow-up time point, $F(4, 140) = 12.37$, $p < .001$. At follow-up, a time-bin by group interaction, $F(4, 140) = 2.78$, $p = .029$, was explored using Sidak-corrected pairwise comparisons: Supervets correctly identified significantly more targets in the first time-bin than any of the other time-bins, ($ps = .001$ or less). Controls correctly identified significantly more sequences in the second and fourth time-bin than in the fifth time-bin, $p = .031$ and $p = .009$ respectively.

Reaction Time

At the initial time point and including follow-up there was a significant group by time-bin interaction, $F(4, 180) = 2.55$, $p = .041$, $F(4, 136) = 2.87$, $p = .025$ respectively. Sidak-corrected pairwise comparisons showed that at both time points, for supervets only, reaction times in the first time-bin were significantly faster than the second bin ($p = .004$,

$p < .001$, respectively), which did not differ from subsequent bins; for controls the first 3 time-bins were faster than the fifth time-bin ($p = .019$, $p = .042$, $p = .030$ respectively).

3.3 MRI data

3.3.1 Initial Time Point

At the initial time point there were no significant differences between groups for any of the structural indices: DTI histogram, $ps > .05$; TBSS; Whole Brain Volume in GM, WM, and WB, $Fs < 1$; and VBM.

3.3.2 Follow-up Time Point

DTI

Supervets ($M = 1004.00 \times 10^{-6}$ mm^2/s, $SD = 33.12$) had significantly higher white matter axial diffusivity (λ1) than controls ($M = 971.67 \times 10^{-6}$ mm^2/s, $SD = 28.87$), $F(1, 23) = 5.96$, $p = .023$. There were no other differences between groups for any of the other indices in all parameters. No differences were found between groups using TBSS, indicating that these differences were not localized to any specific white matter tracts.

Trending Anatomical Differences

At follow-up trends towards group differences emerged on a number of measures; exploratory analyses applying a more liberal threshold ($p < 0.05$ without correction for multiple comparison) are reported.

Whole Volume Analysis:

For white matter percentage, there was a group by time point interaction trend, $F(1, 22) = 4.14$, $p = .054$: relative to supervets ($M = 0.03\%$, $SD = 0.14$) controls

decreased in percentage volume (M = -0.09%, SD = 0.11). This interaction trend was reflected in the whole brain analysis, where there was a marginal group by time point interaction trend, $F(1, 22) = 3.98$, $p = .058$: relative to supervets (M = -0.02%, SD = 0.10) controls decreased in whole brain volume (M = 0.17%, SD = 0.21).

Freesurfer Volume:

Including follow-up, for right hippocampal percentage there was a significant effect of time point, $F(1, 22) = 5.29$, $p = .031$: both groups decreased in volume from initial (M = 0.28%ICV, SD = 0.03) to follow-up time point (M = 0.27%ICV, SD = 0.03).

For left and right cerebellar cortex percentage volume there were group by time point interactions, $F(1, 22) = 5.49$, $p = .032$, $F(1, 22) = 4.31$, $p = .050$, respectively: relative to supervets (left: M = -0.04%, SD = 0.006; right: M = -0.04 %ICV, SD = 0.07) controls increased in percentage volume (left: M = 0.09%, SD = 0.17; right: M = 0.09%, SD = 0.21).

Freesurfer Cortical Thickness:

There was a significant difference between groups for symmetrized percent change in frontal lobe in the left hemisphere, $F(1, 22) = 6.75$, $p = .016$: relative to supervets (M = -0.15%, SD = 1.29), controls increased in thickness (M = 1.43%, SD = 1.36); no significant difference was observed in right hemisphere, $F(1, 22) = 3.63$, $p = .070$.

VBM Analysis:

There were no significant differences.

Correlations with Cognitive Data

We found no correlations between any cognitive and MRI measure entered.

4. Discussion

In this study, we explored cognitive and brain structural differences between superveteran athletes (older adults engaging in long-term high-effort endurance exercise) and non-sedentary volunteers of similar age and lifestyle; we also completed a 12 month follow-up on these same measures. Our study avoided confounds present in many previous studies by using groups that were statistically similar in age, sex, education, IQ, depression levels, cognitive activities, social network, adherence to Mediterranean diet, and potential for fitness.

Both groups had the same potential for fitness, as demonstrated by lung function measures, but our supervet group was indeed more physically fit than our control group, as demonstrated by lower percentage body fat in all supervets and stronger hand-grip in males. The decrease in FVC, although statistically significant, was within the error of the instrument (±3% from Micro Medical website) and the amount of decrease was not a physiologically significant amount as well. As for PEF, there was a small increase (~15% in supervets, less in controls), but with a very wide range of values at both time points (T0: 106 - 756 L/min, T1: 247 – 703 L/min). There is no real reason why there would be a change in PEF and in physiological terms, it is unlikely that this would occur.

Cognitively, we found surprisingly few differences between our two groups in our tasks addressing key domains.[4]

Speed of Processing

We found no differences between supervets and non-sedentary controls in simple reaction time and speed of processing tasks (Trails A and DSST).

Executive Function

We found no differences between supervets and non-sedentary controls in executive function tasks (DSST higher mental function, Trails B, COWAT).

Memory

We found no differences between supervets and non-sedentary controls in working memory (Backwards Digit Span), incidental memory (from DSST), and episodic memory.

In prospective memory tasks, carrying in mind a non-focal PM intention requires increased attentional resource compared to focal PM intention (Einstein, Mcdaniel, Manzi, Cochran, & Baker, 2000). Thus, we anticipated that supervets would show smaller costs in carrying a PM intention, based on the evidence that exercise increases attentional resources used for monitoring (Barnes et al., 2003; Voss et al., 2010; Angevaren et al., 2008, Smith et al. 2010; Erikson et al., 2011; Winker et al., 2010; Voss et al., 2010a). In contrast, in both tasks, supervets showed greater cost in carrying a PM intention. In the Card Sort Task, this was indexed by poorer ongoing task accuracy, and in the McDaniel PM Task, by longer reaction times under the focal PM condition. Einstein & McDaniel's (2000) multiprocess theory suggests that the less complex focal PM condition should rely more on spontaneous retrieval than monitoring, and therefore should show no cost in reaction time. Results from our controls are consistent with this. The significant cost in reaction time in the focal PM condition for supervets suggests this group may engage in active monitoring when there is any PM component, regardless of focality. The fact that this did not improve PM performance is consistent with findings of Einstein et al. (2005) and those of Marsh, Hicks, & Cook (2005).

In both PM tasks completed here a parsimonious explanation of our results is that the supervets were applying higher effort in the form of active monitoring, regardless of whether or not the added effort actually resulted in a performance benefit. Anecdotally, it was clear that the supervets were a highly motivated group who were a) much more keen to attend the 4 hour testing sessions and b) individually committed to the view that their exercise regimen improved their cognitive health and wellbeing. Previous studies have reported that motivational differences in participants elicit better performance on lab-based cognitive tasks (Deaton & Parasuraman, 1993; Tomporowski & Tinsley, 1996). It is therefore doubly surprising that, with such motivation, the supervets did not demonstrate superior performance across the task battery compared to the control group.

Attention

We found no differences between supervets and non-sedentary controls in visual search (Map TEA).

On tasks that engaged significant attentional control, small and subtle differences emerged. These seem best interpreted as differences in strategy between supervets and controls. On the RVIP task, supervets started with faster reaction times and greater accuracy, but rapidly adjusted to a slower pace to sustain accuracy. Controls maintained a more paced performance from the start, with a shallower fatigue curve. These results correspond well with reports from the sports psychology literature. Elite marathoners are reported to employ an "effort sense", a strategy to adjust their effort and achieve relatively stable running pace (Morgan & Pollock, 1977). The RVIP results may reflect this greater metacognitive awareness and attentional control. This is line with previous research as well observing that older expert orienteers, older cyclists, and older athletes

having better attentional skills to modulate their attention and reaction time than non-athletes (Pesce, Cereatti, Casella, Baldari, & Capranica, 2007; Pesce, Cereatti, Forte, Crova, & Casella, 2011) and children participating in exercise interventions having increased meta-cognition as compared to their peers (Tomporowski, McCullick, Pendleton, & Pesce, 2015).

Our results on this comprehensive battery of cognitive tasks complement outcomes from other recent studies of senior athletes (Tseng, Uh, et al., 2013; Winker et al., 2010). Both reported benefits only on fluency tasks, though substantive test batteries were administered in each of these studies. All results could be simply interpreted as advantages emerging only under conditions where strategic effort can influence outcome measures.

This study provides valuable data on an under-researched group. Long-term high-effort endurance athletes who are continuing their regimes beyond the sixth decade of life are a specialist and elite group. Their numbers are limited, and our group sizes, while larger than those reported previously (Tseng, Gundapuneedi, et al. (2013) and Tseng, Uh, et al. (2013) reflect this.

A strength of the present study was the inclusion of longitudinal structural imaging data in the same group of volunteers who contributed cognitive data. This MRI data similarly recorded very few differences between groups. Notably, however, supervets showed higher white matter axial diffusivity than controls at the follow-up time point. Differences in this MR parameter have been interpreted in terms of myelin content (Klawiter et al., 2011), with higher levels reflecting greater neural integrity (Madden et al., 2009; Zhang et al., 2007). Age-related decreases in axial diffusivity have been

reported in many regions (Bennett et al., 2010; Burzynska et al., 2010); higher diffusivity in supervets would suggest they are showing less age-related decline. Potentially higher axial diffusivity in supervets, while not producing functional differences now, may advantage supervets in the longer term.

In terms of hippocampal volume, we did not observe higher volumes in supervets compared to controls at baseline nor were there differences in hippocampal volume changes over the 12-month follow-up. This contrasts with the results from studies by Erickson et al. (2009, 2011) measuring effects of exercise on sedentary older adults. Similarly, we did not observe differences in baseline or longitudinal VBM analyses, in contrast to Colcombe et al. (2003, 2006) and Tseng, Uh, et al. (2013), nor localized differences when using TBSS, in contrast to Tseng, Gundapuneedi, et al. (2013). We did not observe any significant differences or differences in rate of change in volume or cortical thickness ROIs.

As already stated, a key difference between our study and previous ones was our selection of controls. Unlike Tseng et al. (2013a&b) and Winker et al. (2010), we explicitly used non-sedentary controls. In addition, and unlike the earlier studies, our controls were overall similar to supervets in social network, cognitive activity, and depression levels; they were physically mobile and engaging in similar amounts of moderate physical activity as our supervets. Differences between comparator groups would certainly represent potential confounds in the earlier studies (Kim et al., 2011; Lee et al., 2014; Richards et al., 2003; Scherder et al., 2013; Sturman et al., 2005; Thomas & O'Brien, 2008) Overall, our regular moderately-exercising controls do not seem to be at a disadvantage when compared to superveteran athletes.

A second difference from previous studies is the exercise time frame. It is plausible that the cognitive and brain structural benefits of exercise reflect acute effects of relatively short-term interventions in previously sedentary groups, and that benefits are not sustained in extreme long-term exercisers.

The use of a longitudinal design improved the power of our results and the consistency of the null results across the two time points lends further credence to our assertion that these populations don't differ much at this point in time. Where differences were observed at follow-up, mainly in our MRI measures, these were found either using ANOVAs including both time points or using change values, indicating a rate of change difference between the two groups. Exercise may be retarding age-related change but only after a certain threshold is reached. Therefore a longer-term follow-up may have found larger differences beginning to emerge, as older adults exercising at lower levels would reach that threshold more quickly. Alternatively a more holistic view can be taken that the trajectory of age-related change reflects the cumulative impact of all lifestyle factors that affect cognitive function. In this view, the key to good cognitive health in older aging would then be to participate more frequently in cognitive activities, have larger social networks, adhere more to Mediterranean-style diets, and to exercise a bit but not necessarily to the level of an athlete. Indeed, by the logic of additive factors, moderate adherence to a number of moderators would produce greater overall benefits than singular adherence to one.

Our sample also had high mean IQs and high mean education; these two may be protective factors, promoting higher cognitive reserve in both groups (Alexander et al., 1997; Stern, Alexander, Prohovnik, Mayew, & Stern, 1992; Tucker & Stern, 2011). This

may also account for why we did not see much change in either group over the one-year follow-up. Longer term observation may have differentiated our sample, and this was somewhat indicated in the scanning indices where we observed at follow-up a supervet advantage in white matter axial diffusivity and some trends towards emerging volumetric differences also.

We measured physical activity using a questionnaire, but it has become more common for recent interventions and epidemiological studies to use accelerometers for a more objective assessment (Guiraud et al., 2012; Motl, McAuley, & Dlugonski, 2012; Opdenacker, Boen, Coorevits, & Delecluse, 2008). The accelerometers allow one to monitor the number of steps one takes as well as the intensity level of their activity. The only drawback to accelerometers is that they can be quite expensive.

Conclusion

The results across the wide battery of cognitive tests and brain structure indices show that supervets participating in long-term high-effort endurance exercise display only small cognitive and brain structural differences relative to non-sedentary volunteers of similar age, social, cognitive, and neuropsychiatric profile. Cognitively, these small differences manifest in tasks requiring application of effort and strategy, and potentially index differences in effort, metacognitive awareness, or motivation in the supervets. That there are diminishing returns for exercise on cognition, and that these effects do not compare with the benefits seen in a sedentary population, is not surprising. That effects may emerge as brain structural benefits over the longer term allows for the possibility that cognitive benefits too will increase with age. To confirm this, however, it would be

necessary to follow this elite cohort beyond the 6th decade and for longer than the 12-month period tested here.

Conflict of Interest

The authors declare no conflicts of interest.

References

Alexander, G. E., Furey, M. L., Grady, C. L., Pietrini, P., Brady, D. R., Mentis, M. J., & Schapiro, M. B. (1997). Association of premorbid intellectual function with cerebral metabolism in Alzheimer's disease: implications for the cognitive reserve hypothesis. *American Journal of Psychiatry*, *154*(2), 165–172. JOUR. Retrieved from http://journals.psychiatryonline.org/data/Journals/AJP/3673/165.pdf

Angevaren, M., Aufdemkampe, G., Verhaar, H., Aleman, A., & Vanhees, L. (2008). Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment. *Cochrane Database of Systematic Reviews*, (3), Art. No.: CD005381. http://doi.org/10.1002/14651858.CD005381.pub3

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry--the methods. *NeuroImage*, *11*(6 Pt 1), 805–21. http://doi.org/10.1006/nimg.2000.0582

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–51. http://doi.org/10.1016/j.neuroimage.2005.02.018

Barnes, D. E., Yaffe, K., Satariano, W. a, & Tager, I. B. (2003). A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. *Journal of the American Geriatrics Society*, *51*(4), 459–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12657064

Benedict, C., Brooks, S. J., Kullberg, J., Nordenskjöld, R., Burgos, J., Le Grevès, M., … Schiöth, H. B. (2012). Association between physical activity and brain health in older adults. *Neurobiology of Aging*. http://doi.org/10.1016/j.neurobiolaging.2012.04.013

Bennett, I. J., Madden, D. J., Vaidya, C. J., Howard, D. V, & Howard, J. H. (2010). Age-related differences in multiple measures of white matter integrity: A diffusion tensor imaging study of healthy aging. *Human Brain Mapping*, *31*(3), 378–90. http://doi.org/10.1002/hbm.20872

Bherer, L., Erickson, K. I., & Liu-Ambrose, T. (2013). A Review of the Effects of Physical Activity and Exercise on Cognitive and Brain Functions in Older Adults. *Journal of Aging Research*, *2013*, 657508. http://doi.org/10.1155/2013/657508

Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., & Barrett, L. F. (2011). Amygdala volume and social network size in humans. *Nature Neuroscience*, *14*(2), 163–4. http://doi.org/10.1038/nn.2724

Bingham, S. A., Welch, A. A., McTaggart, A., Mulligan, A. A., Runswick, S. A., Luben, R., … Day, N. E. (2001). Nutritional methods in the European prospective investigation of cancer in Norfolk. *Public Health Nutrition*, *4*(3), 847–858. JOUR. http://doi.org/10.1079/PHN2000102

Bunce, D. J., Anstey, K. J., Christensen, H., Dear, K., Wen, W., & Sachdev, P. (2007). White matter hyperintensities and within-person variability in community-dwelling adults aged 60-64 years. *Neuropsychologia*, *45*(9), 2009–15. http://doi.org/10.1016/j.neuropsychologia.2007.02.006

Burzynska, a Z., Preuschhof, C., Bäckman, L., Nyberg, L., Li, S.-C., Lindenberger, U., & Heekeren, H. R. (2010). Age-related differences in white matter microstructure: region-specific patterns of diffusivity. *NeuroImage*, *49*(3), 2104–12. http://doi.org/10.1016/j.neuroimage.2009.09.041

Cercignani, M., Bozzali, M., Iannucci, G., Comi, G., & Filippi, M. (2001). Magnetisation transfer ratio and mean diffusivity of normal appearing white and grey matter from patients with multiple sclerosis. *Journal of Neurology, Neurosurgery, and Psychiatry*, *70*(3), 311–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1737283&tool=pmcentrez&rendertype=abstract

Chang, Y. K., Labban, J. D., Gapin, J. I., & Etnier, J. L. (2012). The effects of acute exercise on cognitive performance: a meta-analysis. *Brain Research*, *1453*(250), 87–101. http://doi.org/10.1016/j.brainres.2012.02.068

Charlton, R. a, Barrick, T. R., McIntyre, D. J., Shen, Y., O'Sullivan, M., Howe, F. a, … Markus, H. S. (2006). White matter damage on diffusion tensor imaging correlates with age-related cognitive decline. *Neurology*, *66*(2), 217–22. http://doi.org/10.1212/01.wnl.0000194256.15247.83

Colcombe, S. J., Erickson, K. I., Raz, N., Webb, A. G., Cohen, N. J., McAuley, E., & Kramer, A. F. (2003). Aerobic fitness reduces brain tissue loss in aging humans. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *58*(2), 176–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12586857

Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., … Kramer, A. F. (2006). Aerobic exercise training increases brain volume in aging humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *61*(11), 1166–1170. JOUR. Retrieved from http://biomedgerontology.oxfordjournals.org/content/61/11/1166.short

Colcombe, S. J., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: a meta-analytic study. *Psychological Science : A Journal of the American Psychological Society / APS*, *14*(2), 125–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12661673

Coubard, O. a, Ferrufino, L., Boura, M., Gripon, A., Renaud, M., & Bherer, L. (2011). Attentional control in normal aging and Alzheimer's disease. *Neuropsychology*, *25*(3), 353–67. http://doi.org/10.1037/a0022058

Craik, F. I. M. (2006). Age-related changes in human memory: Practical consequences., 181–197. JOUR.

Crawford, J. R., Smith, G., Maylor, E. a, Della Sala, S., & Logie, R. H. (2003). The Prospective and Retrospective Memory Questionnaire (PRMQ): Normative data and latent structure in a large non-clinical sample. *Memory (Hove, England)*, *11*(3), 261–75. http://doi.org/10.1080/09658210244000027

Darowski, E. S., Helder, E., Zacks, R. T., Hasher, L., & Hambrick, D. Z. (2008). Age-related differences in cognition: the role of distraction control. *Neuropsychology*,

22(5), 638–44. http://doi.org/10.1037/0894-4105.22.5.638

Deaton, J., & Parasuraman, R. (1993). Sensory and cognitive vigilance: Effects of age on performance and subjective workload. *Human Performance*, 37–41. http://doi.org/10.1207/s15327043hup0601

Dempsey, J. A. (1986). JB Wolffe memorial lecture. Is the lung built for exercise? *Medicine and Science in Sports and Exercise*, *18*(2), 143–155. JOUR.

Dowell, N. G., Ruest, T., Evans, S. L., King, S. L., Tabet, N., Tofts, P. S., & Rusted, J. M. (2013). MRI of carriers of the apolipoprotein E e4 allele-evidence for structural differences in normal-appearing brain tissue in e4+ relative to e4- young adults. *NMR in Biomedicine*, *26*(6), 674–82. http://doi.org/10.1002/nbm.2912

Einstein, G. O., Mcdaniel, M. A., Manzi, M., Cochran, B., & Baker, M. (2000). Prospective Memory and Aging : Forgetting Intentions Over Short Delays. *Psychology and Aging*, *15*(4), 671–683. http://doi.org/10.1037//0882-7974.15.4.I

Einstein, G. O., McDaniel, M. a, Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology. General*, *134*(3), 327–42. http://doi.org/10.1037/0096-3445.134.3.327

Erickson, K. I., Prakash, R. S., Voss, M. W., Chaddock, L., Hu, L., Morris, K. S., … Kramer, A. F. (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus*, *19*(10), 1030–9. http://doi.org/10.1002/hipo.20547

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., … Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3017–3022. http://doi.org/10.1073/pnas.1015950108

Eskes, G. a, Longman, S., Brown, A. D., McMorris, C. a, Langdon, K. D., Hogan, D. B., & Poulin, M. (2010). Contribution of physical fitness, cerebrovascular reserve and cognitive stimulation to cognitive function in post-menopausal women. *Frontiers in Aging Neuroscience*, *2*(October), 137. http://doi.org/10.3389/fnagi.2010.00137

Etnier, J. L., Salazar, W., Landers, D. M., Petruzzello, S. J., Han, M., & Nowell, P. (1997). The influence of physical fitness and exercise upon cognitive functioning: A meta-analysis. *Journal of Sport & Exercise Psychology*, *19*(3), 249–277. JOUR.

Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., … Dale, A. M. (2009). One-Year Brain Atrophy Evident in Healthy Aging. *Journal of Neuroscience*, *29*(48), 15223–15231. http://doi.org/10.1523/JNEUROSCI.3252-09.2009

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. JOUR. http://doi.org/http://dx.doi.org/10.1016/0022-3956(75)90026-6

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak,

R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, *14*(1 Pt 1), 21–36. http://doi.org/10.1006/nimg.2001.0786

Guiraud, T., Granger, R., Gremeaux, V., Bousquet, M., Richard, L., Soukarie, L., … Pathak, A. (2012). Accelerometer as a tool to assess sedentarity and adherence to physical activity recommendations after cardiac rehabilitation program. *Annals of Physical and Rehabilitation Medicine*, *55*(5), 312–321. JOUR.

Gunning-Dixon, F. M., Brickman, A. M., Cheng, J. C., & Alexopoulos, G. S. (2009). Aging of cerebral white matter: a review of MRI findings. *International Journal of Geriatric Psychiatry*, *24*(2), 109–117. JOUR. http://doi.org/10.1002/gps

Hayes, S. M., Hayes, J. P., Cadden, M., & Verfaellie, M. (2013). A review of cardiorespiratory fitness-related neuroplasticity in the aging brain. *Frontiers in Aging Neuroscience*, *5*(July), 31. http://doi.org/10.3389/fnagi.2013.00031

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature Reviews. Neuroscience*, *5*(2), 87–96. http://doi.org/10.1038/nrn1323

Hindin, S. B., & Zelinski, E. M. (2012). Extended Practice and Aerobic Exercise Interventions Benefit Untrained Cognitive Outcomes in Older Adults :, 136–141. http://doi.org/10.1111/j.1532-5415.2011.03761.x

Hsu, J.-L., Leemans, A., Bai, C.-H., Lee, C.-H., Tsai, Y.-F., Chiu, H.-C., & Chen, W.-H. (2008). Gender differences and age-related white matter changes of the human brain: a diffusion tensor imaging study. *NeuroImage*, *39*(2), 566–77. http://doi.org/10.1016/j.neuroimage.2007.09.017

Hutchison, K. a, Balota, D. A., & Ducheck, J. M. (2010). The utility of Stroop task switching as a marker for early-stage Alzheimer's disease. *Psychology and Aging*, *25*(3), 545–59. http://doi.org/10.1037/a0018498

Joy, S., Fein, D., & Kaplan, E. (2003). Decoding Digit Symbol Speed, Memory, and Visual Scanning. *Assessment*. Retrieved from http://asm.sagepub.com/content/10/1/56.short

Kanaan, R. a, Allin, M., Picchioni, M., Barker, G. J., Daly, E., Shergill, S. S., … McGuire, P. K. (2012). Gender differences in white matter microstructure. *PloS One*, *7*(6), e38272. http://doi.org/10.1371/journal.pone.0038272

Kanai, R., Bahrami, B., Roylance, R., & Rees, G. (2012). Online social network size is reflected in human brain structure. *Proceedings. Biological Sciences / The Royal Society*, *279*(1732), 1327–34. http://doi.org/10.1098/rspb.2011.1959

Kim, J. H., Hwang, K. J., Kim, J.-H., Lee, Y. H., Rhee, H. Y., & Park, K.-C. (2011). Regional white matter hyperintensities in normal aging, single domain amnestic mild cognitive impairment, and mild Alzheimer's disease. *Journal of Clinical Neuroscience : Official Journal of the Neurosurgical Society of Australasia*, *18*(8), 1101–6. http://doi.org/10.1016/j.jocn.2011.01.008

Klawiter, E. C., Schmidt, R. E., Trinkaus, K., Liang, H.-F., Budde, M. D., Naismith, R. T., … Benzinger, T. L. (2011). Radial diffusivity predicts demyelination in ex vivo multiple sclerosis spinal cords. *NeuroImage*, *55*(4), 1454–60. http://doi.org/10.1016/j.neuroimage.2011.01.007

Lee, T., Lipnicki, D. M., Crawford, J. D., Henry, J. D., Trollor, J. N., Ames, D., … Sachdev, P. S. (2014). Leisure activity, health, and medical correlates of neurocognitive performance among monozygotic twins: The older Australian twins study. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *69*, 514–522. http://doi.org/10.1093/geronb/gbt031

Lövdén, M., Schaefer, S., Noack, H., Bodammer, N. C., Kühn, S., Heinze, H.-J., … Lindenberger, U. (2012). Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. *Neurobiology of Aging*, *33*(3), 620.e9-620.e22. http://doi.org/10.1016/j.neurobiolaging.2011.02.013

Lubben, J., Blozik, E., Gillmann, G., Iliffe, S., von Rentein Kruse, W., Beck, J. C., & Stuck, A. E. (2006). Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. *The Gerentologist*, *46*(4), 503–513.

Luo, L., & Craik, F. I. M. (2008). Aging and memory: a cognitive approach. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, *53*(6), 346–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19646789

Madden, D. J., Bennett, I. J., Burzynska, A., Potter, G. G., Chen, N.-K., & Song, A. W. (2012). Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta*, *1822*(3), 386–400. http://doi.org/10.1016/j.bbadis.2011.08.003

Madden, D. J., Bennett, I. J., & Song, A. W. (2009). Cerebral white matter integrity and cognitive aging: contributions from diffusion tensor imaging. *Neuropsychology Review*, *19*(4), 415–35. http://doi.org/10.1007/s11065-009-9113-2

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(1), 68–75. http://doi.org/10.1037/0278-7393.31.1.68

McAvinue, L. P., Habekost, T., Johnson, K. a, Kyllingsbæk, S., Vangkilde, S., Bundesen, C., & Robertson, I. H. (2012). Sustained attention, attentional selectivity, and attentional capacity across the lifespan. *Attention, Perception & Psychophysics*, *74*(8), 1570–82. http://doi.org/10.3758/s13414-012-0352-6

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, *144*(September), 127–144. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/acp.775/abstract

McDaniel, M. A., Shelton, J. T., Breneiser, J. E., Moynan, S., & Balota, D. a. (2011). Focal and nonfocal prospective memory performance in very mild dementia: a

signature decline. *Neuropsychology*, *25*(3), 387–96. http://doi.org/10.1037/a0021682

Middleton, L. E., Barnes, D. E., Lui, L.-Y., & Yaffe, K. (2010). Physical activity over the life course and its association with cognitive performance and impairment in old age. *Journal of the American Geriatrics Society*, *58*(7), 1322–6. http://doi.org/10.1111/j.1532-5415.2010.02903.x

Middleton, L. E., Mitnitski, A., Fallah, N., Kirkland, S. a, & Rockwood, K. (2008). Changes in cognition and mortality in relation to exercise in late life: a population based study. *PloS One*, *3*(9), e3124. http://doi.org/10.1371/journal.pone.0003124

Morey, R., Petty, C., & Xu, Y. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage*, *45*(3), 855–866. http://doi.org/10.1016/j.neuroimage.2008.12.033.A

Morgan, W. P., & Pollock, M. L. (1977). PSYCHOLOGIC CHARACTERIZATION OF THE ELITE. *Annals of the New York Academy of Sciences*, *301*(1), 382–403.

Motl, R. W., McAuley, E., & Dlugonski, D. (2012). Reactivity in baseline accelerometer data from a physical activity behavioral intervention. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *31*(2), 172–5. http://doi.org/10.1037/a0025965

Nelson, H. E. (1982). *National Adult Reading Test (NART): For the Assessment of Premorbid Intelligence in Patients with Dementia: Test Manual*. BOOK, NFER-Nelson.

Nemati Karimooy, H., Hosseini, M., Nemati, M., & Esmaily, H. O. (2011). Lifelong physical activity affects mini mental state exam scores in individuals over 55 years of age. *Journal of Bodywork and Movement Therapies*, 6–11. http://doi.org/10.1016/j.jbmt.2011.08.003

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, *15*(1), 1–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11747097

Niemann, C., Godde, B., Staudinger, U. M., & Voelcker-Rehage, C. (2014). Exercise-induced changes in basal ganglia volume and cognition in older adults. *Neuroscience*, *281C*(September), 147–163. http://doi.org/10.1016/j.neuroscience.2014.09.033

Niemann, C., Godde, B., & Voelcker-Rehage, C. (2014). Not only cardiovascular, but also coordinative exercise increases hippocampal volume in older adults. *Frontiers in Aging Neuroscience*, *6*(JUN), 1–24. http://doi.org/10.3389/fnagi.2014.00170

Opdenacker, J., Boen, F., Coorevits, N., & Delecluse, C. (2008). Effectiveness of a lifestyle intervention and a structured exercise intervention in older adults. *Preventive Medicine*, *46*(6), 518–24. http://doi.org/10.1016/j.ypmed.2008.02.017

Peltsch, a, Hemraj, a, Garcia, a, & Munoz, D. P. (2011). Age-related trends in saccade characteristics among the elderly. *Neurobiology of Aging*, *32*(4), 669–79. http://doi.org/10.1016/j.neurobiolaging.2009.04.001

Pesce, C., Cereatti, L., Casella, R., Baldari, C., & Capranica, L. (2007). Preservation of visual attention in older expert orienteers at rest and under physical effort. *Journal of Sport and Exercise Psychology*, *29*(1), 78. JOUR.

Pesce, C., Cereatti, L., Forte, R., Crova, C., & Casella, R. (2011). Acute and chronic exercise effects on attentional control in older road cyclists. *Gerontology*, *57*(2), 121–128. JOUR.

Prakash, R. S., Voss, M. W., Erickson, K. I., Lewis, J. M., Chaddock, L., Malkowski, E., … Kramer, A. F. (2011). Cardiorespiratory fitness and attentional control in the aging brain. *Frontiers in Human Neuroscience*, *4*(January), 229. http://doi.org/10.3389/fnhum.2010.00229

Predovan, D., Fraser, S. a, Renaud, M., & Bherer, L. (2012). The effect of three months of aerobic training on stroop performance in older adults. *Journal of Aging Research*, *2012*, 269815. http://doi.org/10.1155/2012/269815

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. JOUR.

Raz, N., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: regional and individual differences. *NeuroImage*, *51*(2), 501–11. http://doi.org/10.1016/j.neuroimage.2010.03.020

Raz, N., Schmiedek, F., Rodrigue, K. M., Kennedy, K. M., Lindenberger, U., & Lövdén, M. (2013). Differential brain shrinkage over 6 months shows limited association with cognitive practice. *Brain and Cognition*, *82*(2), 171–80. http://doi.org/10.1016/j.bandc.2013.04.002

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*(3), 271–276. JOUR.

Resnick, S. M., Pham, D. L., Kraut, M. a, Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(8), 3295–301. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12716936

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*(4), 1402–18. http://doi.org/10.1016/j.neuroimage.2012.02.084

Richards, M., Hardy, R., & Wadsworth, M. E. J. (2003). Does active leisure protect cognition? Evidence from a national birth cohort. *Social Science & Medicine (1982)*, *56*(4), 785–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12560011

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society : JINS*, *2*(6), 525–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9375156

Ruscheweyh, R., Willemer, C., Krüger, K., Duning, T., Warnecke, T., Sommer, J., …

Flöel, a. (2011). Physical activity and memory functions: an interventional study. *Neurobiology of Aging*, *32*(7), 1304–19. http://doi.org/10.1016/j.neurobiolaging.2009.08.001

Rusted, J. M., Sawyer, R., Jones, C., Trawley, S. L., & Marchant, N. L. (2009). Positive effects of nicotine on cognition: the deployment of attention for prospective memory. *Psychopharmacology*, *202*(1–3), 93–102. http://doi.org/10.1007/s00213-008-1320-7

Rusted, J. M., & Warburton, D. M. (1989). Cognitive models and cholinergic drugs. *Neuropsychobiology*, *21*(1), 31–36. JOUR.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8759042

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*(1–3), 35–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11035219

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society : JINS*, *16*(5), 754–60. http://doi.org/10.1017/S1355617710000706

Salthouse, T. A. (2012). Does the direction and magnitude of cognitive change depend on initial level of ability? *Intelligence*, *40*(4), 352–361. http://doi.org/10.1016/j.intell.2012.02.006

Scherder, E., Scherder, R., Verburgh, L., Königs, M., Blom, M., Kramer, A. F., & Eggermont, L. (2013). Executive Functions of Sedentary Elderly May Benefit from Walking: A Systematic Review and Meta-Analysis. *The American Journal of Geriatric Psychiatry : Official Journal of the American Association for Geriatric Psychiatry*, 1–10. http://doi.org/10.1016/j.jagp.2012.12.026

Schinka, J. A., McBride, A., Vanderploeg, R. D., Tennyson, K., Borenstein, A. R., & Mortimer, J. A. (2005). Florida Cognitive Activities Scale: initial development and validation. *J Int Neuropsychol Soc*, *11*(1), 108–116.

Schinka, J. A., Raj, A., Loewenstein, D. A., Small, B. J., Duara, R., & Potter, H. (2010). Cross-validation of the Florida Cognitive Activities Scale (FCAS) in an Alzheimer's disease research center sample. *Journal of Geriatric Psychiatry and Neurology*, *23*(1), 9–14. http://doi.org/10.1177/0891988709342724

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol*, *37*, 819–820. JOUR. Retrieved from http://www.tandfonline.com/doi/abs/10.1300/J018v05n01_09

Smith, P. J., Blumenthal, J. a, Hoffman, B. M., Cooper, H., Strauman, T. a, Welsh-Bohmer, K., … Sherwood, A. (2010). Aerobic exercise and neurocognitive performance: a meta-analytic review of randomized controlled trials. *Psychosomatic Medicine*, *72*(3), 239–52. http://doi.org/10.1097/PSY.0b013e3181d14633

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C.

E., … Behrens, T. E. J. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*(4), 1487–505. http://doi.org/10.1016/j.neuroimage.2006.02.024

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208-19. http://doi.org/10.1016/j.neuroimage.2004.07.051

Spreen, O., & Strauss, E. (1991). A compendium of neuropsychological tests: Administration, norms, and commentary. GEN, Oxford University Press (New York).

Stern, Y., Alexander, G. E., Prohovnik, I., Mayew, R., & Stern, Y. (1992). Inverse Rehtionshp Between Education and Parietotemporal Perfusion Deficit in Alzheimer ' s Disease, 371–375.

Sturman, M. T., Morris, M. C., Leon, C. F. M. De, Bienias, J. L., Wilson, R. S., & Evans, D. A. (2013). Physical Activity, Cognitive Activity, and Cognitive Decline in a Biracial Community Population, *62*.

Sturman, M. T., Morris, M. C., Mendes de Leon, C. F., Bienias, J. L., Wilson, R. S., & Evans, D. a. (2005). Physical activity, cognitive activity, and cognitive decline in a biracial community population. *Archives of Neurology*, *62*(11), 1750–4. http://doi.org/10.1001/archneur.62.11.1750

Thomas, A. J., & O'Brien, J. T. (2008). Depression and cognition in older adults. *Current Opinion in Psychiatry*, *21*(1), 8–13. JOUR.

Tofts, P. S., Davies, G. R., & Dehmeshki, J. (2003). Histograms: measuring subtle diffuse disease. *Quantitative MRI of the Brain: Measuring Changes Caused by Disease*, 581–610. JOUR.

Tomporowski, P. D., McCullick, B., Pendleton, D. M., & Pesce, C. (2015). Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. *Journal of Sport and Health Science*, *4*(1), 47–55. JOUR.

Tomporowski, P. D., & Tinsley, V. F. (1996). Effects of memory demand and motivation on sustained attention in young and older adults. *The American Journal of Psychology*, *109*(2), 187–204. http://doi.org/10.2307/1423272

Trichopoulou, A., Costacou, T., Bamia, C., & Trichopoulos, D. (2003). Adherence to a Mediterranean diet and survival in a Greek population. *New England Journal of Medicine*, *348*(26), 2599–2608. JOUR.

Tseng, B. Y., Gundapuneedi, T., Khan, M. a. a, Diaz-Arrastia, R., Levine, B. D. D., Lu, H., … Zhang, R. (2013). White Matter Integrity in Physically Fit Older Adults. *NeuroImage*, *82*, 510–6. http://doi.org/10.1016/j.neuroimage.2013.06.011

Tseng, B. Y., Uh, J., Rossetti, H. C., Cullum, C. M., Diaz-Arrastia, R. F., Levine, B. D., … Zhang, R. (2013). Masters athletes exhibit larger regional brain volume and better cognitive performance than sedentary older adults. *Journal of Magnetic Resonance*

Imaging : JMRI, *0*(5), 1–8. http://doi.org/10.1002/jmri.24085

Tucker, A., & Stern, Y. (2011). Cognitive reserve in aging. *Current Alzheimer Research*, *8*(4), 354–360. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3135666/

Uylings, H. B. M., & de Brabander, J. M. (2002). Neuronal Changes in Normal Human Aging and Alzheimer's Disease. *Brain and Cognition*, *49*(3), 268–276. http://doi.org/10.1006/brcg.2001.1500

Voss, M. W., Kramer, A. F., Basak, C., Prakash, R. S., & Roberts, B. (2010). Are Expert Athletes " Expert " in the Cognitive Laboratory ? A Meta-Analytic Review of Cognition and Sport Expertise. *Applied Cognitive Psychology*, *826*(June 2009), 812–826. http://doi.org/10.1002/acp

Washburn, R. A., Smith, K. W., Jette, A. M., & Janney, C. A. (1993). The Physical Activity Scale for the Elderly (PASE): development and evaluation. *Journal of Clinical Epidemiology*, *46*(2), 153–162. JOUR.

Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised*. BOOK, Psychological Corporation.

Wesnes, K., & Warburton, D. M. (1983). Effects of smoking on rapid information processing performance. *Neuropsychobiology*, *9*(4), 223–229. JOUR. http://doi.org/10.1159/000117969

Weuve, J., Kang, J. H., Manson, J. E., Breteler, M. M. B., Ware, J. H., & Grodstein, F. (2004). Physical activity, including walking, and cognitive function in older women. *JAMA : The Journal of the American Medical Association*, *292*(12), 1454–61. http://doi.org/10.1001/jama.292.12.1454

Winker, R., Lukas, I., Perkmann, T., Haslacher, H., Ponocny, E., Lehrner, J., … Dal-Bianco, P. (2010). Cognitive function in elderly marathon runners: Cross-sectional data from the marathon trial (apsoem). *Wiener Klinische Wochenschrift*, *122*(23–24), 704–716. http://doi.org/10.1007/s00508-010-1485-z

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., … Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, *45*(1 Suppl), S173-86. http://doi.org/10.1016/j.neuroimage.2008.10.055

Young, J. C., Angevaren, M., Rusted, J. M., & Tabet, N. (2015). Aerobic exercise to improve cognitive function in older people without known cognitive impairment. *The Cochrane Library*. JOUR.

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience*, *14*(5), 656–61. http://doi.org/10.1038/nn.2773

Zhang, Y., Du, A.-T., Hayasaka, S., Jahng, G.-H., Hlavin, J., Zhan, W., … Schuff, N. (2010). Patterns of age-related water diffusion changes in human brain by concordance and discordance analysis. *Neurobiology of Aging*, *31*(11), 1991–2001. http://doi.org/10.1016/j.neurobiolaging.2008.10.009

Zhang, Y., Schuff, N., Jahng, G.-H., Bayne, W., Mori, S., Schad, L., … Weiner, M. W. (2007). Diffusion tensor imaging of cingulum fibers in mild cognitive impairment and Alzheimer disease. *Neurology*, *68*(1), 13–9. http://doi.org/10.1212/01.wnl.0000250326.77323.01

Footnotes

[1] Between baseline and 12-month follow-up MRI sessions the MRI scanner gradients were upgraded yielding an increase in gradient amplitude (from 33mT/m to 44 mT/m) and slew rates. For our DTI measures we ensured diffusion gradient amplitudes were identical for both time points, but differences remained in the imaging gradients. As a result, the DTI data between sessions were not directly compared. Daily independent QA data using the EPI sequence also showed there were no SNR step changes as a result of the upgrade. Any potential differences in T1-weighted MPRAGE scans as a result of the hardware upgrade were accounted for in analyses: participants' volumes at each time point were normalized to templates (Freesurfer longitudinal); normalized to baseline time point volumes (VBM longitudinal), or intracranial volumes at each time point were used to normalize the data as is standard practice.

[2] PM task was a Focal PM task, switch reaction time was from the Word-Neutral condition, which was the easiest switch trial.

[3] Note there were violations of normality and some homogeneity of variance violations in follow-up ANOVAs for the follow-up time point, so these must be interpreted with caution.

[4] Multiple regression analyses using the cohort as a single continuous group and physical activity / physiological measures as predictors did not alter these outcomes. Medium to low power for detection also has to be considered when interpreting some of the non-significant findings.

Table I

Means, Standard Deviations, and Significances between Supervet and Control Group for Demographics

Measure	Supervets Mean (SD)	Controls Mean (SD)	Significance at T0	Supervets Mean (SD)	Controls Mean (SD)	Significance at T1
Demographics:						
Sex[a]	6 female	9 female	$p = .158$	6 female	7 female	$p = .577$
Age (years)	67.88 (5.45)	68.35 (5.81)	$p = .771$ $d = 0.09$	69.29 (5.07)	69.21 (5.83)	$p = .965$ $d = 0.02$
Education (years)[b]	15.26 (3.59)	16.18 (3.38)	$p = .363$‡ $d = 0.27$	16.00 (3.59)	16.16 (3.45)	$p = .888$ $d = 0.05$
IQ[c]	117.66 (3.92)	119.93 (4.51)	$p = .077$‡ $d = 0.55$	118.37 (5.72)	120.70 (4.34)	$p = .170$‡ $d = 0.47$
Depression[d]	0.52 (0.89)	1.18 (1.71)	$p = .139$‡ $d = 0.52$	0.71 (1.10)	1.11 (1.82)	$p = .428$‡ $d = 0.27$

Note: a. Chi-squared test used. b. Years of full-time education. c. Full Scale Pre-morbid Intelligence Quotient derived from National Adult Reading Test. d. From Geriatric Depression Scale score. ‡ - From 1000-sample bootstrap results.

Table II

Means, Standard Deviations, and Significances between Supervet and Control Groups Physiological, Diet, Social, and Cognitive Activities Characteristics

Measure	Supervets Mean (SD)	Controls Mean (SD)	Significance at T0	Supervets Mean (SD)	Controls Mean (SD)	Significance with T0 and T1
Physical Activity:						
PASE[a]	221.70 (71.24)	169.39 (66.03)	$p = .011$* $d = 0.77$	230.60 (60.39)	160.87 (55.12)	$p = .001^{\Omega}$*
Walking (Avg Hours / Day)	1.09 (0.76)	0.92 (0.96)		0.99 (0.70)	0.77 (0.43)	
Light Sports (Avg Hours / Day)	0.45 (0.59)	0.46 (0.92)		0.60 (0.70)	0.38 (0.67)	
Moderate Sports (Avg Hours / Day)	0.20 (0.29)	0.10 (0.18)		0.37 (0.58)	0.19 (0.33)	
Strenuous Sports (Avg Hours / Day)	1.14 (1.00)	0.18 (0.38)		1.02 (0.69)	0.19 (0.31)	
Muscle Strength / Endurance Exercise (Avg Hours / Day)	0.15 (0.24)	0.03 (0.06)		0.11 (0.14)	0.02 (0.06)	
Diet:						
Mediterranean Diet Score	4.81 (1.62)	4.64 (1.81)	$p = .690^{\ddagger}$ $d = 0.10$	4.84 (1.89)	4.78 (1.86)	$p = .494^{\Omega}$
Lung Function:						
Forced Expiratory Volume 1st Second (L)	2.87 (0.72)	2.78 (0.78)	$p = .686$ $d = 0.12$	2.97 (0.66)	2.86 (0.72)	$p = .999$ $n^2p < .001$
Forced Vital Capacity (L)	3.78 (0.87)	3.38 (0.93)	$p = .153$ $d = 0.46$	3.77 (0.84)	3.36 (0.81)	$p = .235$ $n^2p = .047$
Forced Expiratory Ratio (%)	78.01 (8.93)	83.20 (8.88)	$p = .079^{\ddagger}$ $d = 0.60$	77.25 (8.92)	84.75 (9.20)	$p = .085^{\Omega}$
Peak Expiratory Flow (L/min)	411.89 (143.83)	421.95 (155.24)	$p = .875$ $d = 0.07$	481.37 (119.32)	458.25 (132.21)	$p = .963$ $n^2p < .001$
Physiological:						
Body Fat Males (%)	22.05 (4.07)	26.77 (4.07)	$p = .003$* $d = 1.20$	25.03 (4.38)	28.14 (4.09)	$p = .031$* $n^2p = .180$
Body Fat Females (%)	32.50 (4.19)	39.38 (6.59)	$p = .042$* $d = 1.28$	35.45 (2.94)	40.84 (3.44)	$p = .030$* $n^2p = .361$
Hand-grip Strength Male (kg)	40.65 (7.00)	36.12 (5.68)	$p = .079$ $d = 0.71$	43.69 (5.66)	37.34 (5.27)	$p = .006$* $n^2p = .339$
Hand-grip Strength Female (kg)	27.17 (4.71)	25.22 (3.28)	$p = .361$ $d = 0.54$	26.17 (4.70)	26.08 (3.44)	$p = .768$ $n^2p = .009$
Social and Cognitive Activities:						
Social Network[b]	20.07 (4.74)	20.09 (6.24)	$p = .992$ $d < 0.01$	20.95 (4.30)	19.00 (7.19)	$p = .728^{\Omega}$
Cognitive Activities[c]	46.48 (10.49)	50.89 (6.81)	$p = .114^{\ddagger}$ $d = 0.50$	48.86 (8.24)	50.48 (7.60)	$p = .434^{\Omega}$

Note: a. Physical Activity Scale for the Elderly score. b. From Lubben Social Network Scale. c. From Florida Cognitive Activity Scale. ‡ - From 1000-sample bootstrap results. Ω - From mixed ANOVA on trimmed means *- Statistically significant

Table III

Means, Standard Deviations, and Significances between Supervet and Control Groups on Speed of Processing and Cognitive Tasks

	Time Point 0			Time Point 1		
Measure	**Supervets Mean (SD)**	**Controls Mean (SD)**	**Significance at T0**	**Supervets Mean (SD)**	**Controls Mean (SD)**	**Significance with T0 and T1**
Speed of Processing:						
Simple Reaction Time (ms)	301.71 (48.65)	326.29 (52.62)	$p = .093$ $d = 0.50$	300.78 (48.80)	333.95 (92.05)	$p = .066$ $n^2p = .086$
IIV (SD, ms)[a]	78.64 (26.10)	81.75 (29.03)	$p = .695$ $d = 0.12$	94.09 (39.30)	86.59 (41.92)	$p = .734$ $n^2p = .003$
ICV (SD/mean * 100)[b]	25.63 (7.21)	24.55 (6.98)	$p = .599$ $d = 0.16$	30.39 (10.92)	25.23 (9.06)	$p = .090$ $n^2p = .074$
Symbol Copy (secs/item)[c]	0.91 (0.19)	0.91 (0.17)	$p = .964$ $d = 0.01$	0.89 (0.14)	0.85 (0.14)	$p = .909$ $n^2p < .001$
Trails A (secs)[d]	30.95 (6.88)	31.90 (11.33)	$p = .971$ $d = 0.11$	31.74 (6.67)	32.17 (13.04)	$p = .963^{\Omega}$
Executive Function:						
DSST Higher Mental Function (secs/item)[e]	1.12 (0.28)	1.00 (0.31)	$p = .167$ $d = 0.42$	1.09 (0.32)	1.03 (0.27)	$p = .385$ $n^2p = .020$
Trails B - A (secs)[f]	39.90 (27.03)	26.32 (12.55)	$p = .093$ $d = 0.64$	35.63 (14.34)	23.14 (12.35)	$p = .500^{\Omega}$
Verbal Fluency (total correct)[g]	45.89 (9.82)	50.23 (14.27)	$p = .657$ $d = 0.37$	48.86 (9.74)	50.63 (15.31)	$p = .580$ $n^2p = .008$
Memory:						
Incidental Memory (percent recalled)[g]	41.20 (21.85)	48.15 (27.74)	$p = .382^{\ddagger}$ $d = 0.29$	49.71 (25.76)	48.37 (23.55)	$p = .720$ $n^2p = .004$
Episodic Memory (percent recalled)	28.42 (13.44)	30.71 (11.91)	$p = .635^{\ddagger}$ $d = 0.01$	28.33 (11.13)	26.15 (14.02)	$p = .859$ $n^2p = .001$
Backward Digit Spans (number of digits recalled)				5.76 (1.04)	5.95 (1.55)	$p = .673^{\ddagger}$
Subjective Prospective Memory[h]	19.01 (3.21)	19.00 (3.53)	$p = .991^{\ddagger}$ $d < 0.01$	18.76 (2.86)	19.95 (3.11)	$p = .667^{\Omega}$
Subjective Retrospective Memory[h]	18.06 (3.95)	17.18 (3.17)	$p = .393^{\ddagger}$ $d = 0.25$	18.14 (3.07)	18.59 (3.56)	$p = .603^{\Omega}$
Attention:						
Visual Search (total no. correct)[i]	57.89 (12.93)	56.09 (9.32)	$p = .587^{\ddagger}$ $d = 0.16$	56.90 (12.28)	54.05 (10.84)	$p = .936^{\Omega}$

Note: a. Intra-individual Variability. b. Intra-individual Coefficient of Variability. c. From Digit Symbol Substitution Task for 30 seconds. d. Trail Making Test part A time taken. e. Digit Symbol Substitution Task, time per item accounting for copying speed. f. Trail Making Test time taken for part B with part A subtracted. g. From Digit Symbol Substitution Task. h. From the Prospective and Retrospective Memory Questionnaire. I. From Map Test of Everyday Attention. ‡ - From 1000-sample bootstrap result. Ω – From mixed ANOVA on trimmed means. * - Statistically significant.

Table IV

Means, Standard Deviations, and Significances Between Supervet and Control Groups on the Card Sort PM Task

	Time Point 0			Time Point 1		
Condition	**Supervets Mean (SD)**	**Controls Mean (SD)**	**Significance at T0**	**Supervets Mean (SD)**	**Controls Mean (SD)**	**Significance with T0 and T1**
PM Accuracy (%)	77.98 (28.75)	84.38 (18.53)	$p = .410^{‡}$ $d = 0.27$	92.19 (12.81)	90.41 (11.29)	$p = .993$ $n^2p = .001$
Card Sort Accuracy without PM (%)	97.36 (2.27)	97.44 (2.38)	$p = .924^{‡}$ $d = 0.04$	97.16 (2.89)	98.18 (1.68)	$p = .568^{Ω}$
Card Sort Accuracy with PM (%)	93.87 (4.34)	96.22 (2.33)	$p = .038$* $d = 0.69$	95.51 (3.12)	95.75 (4.19)	$p = .156^{Ω}$
Card Sort Reaction Time without PM (ms)	557.55 (77.23)	548.99 (61.13)	$p = .722$ $d = 0.12$	566.23 (77.46)	566.01 (62.64)	$p = .547^{Ω}$
Card Sort Reaction Time with PM (ms)	697.37 (78.85)	676.20 (77.62)	$p = .335$ $d = 0.28$	713.42 (69.25)	702.67 (102.32)	$p = .847^{Ω}$

Note: ‡- From 1000-sample bootstrap result. Ω – From mixed ANOVA on trimmed means. * - Statistically significant.

Table V

Means and Standard Deviations on the McDaniel PM Task for PM Accuracies, Ongoing Category Decision Task Accuracies, and Ongoing Category Decision Task Reaction Times

	Time Point 0		Time Point 1	
Task Condition	**Supervet Mean (SD)**	**Control Mean (SD)**	**Supervet Mean (SD)**	**Control Mean (SD)**
		PM Accuracy		
Focal PM (%)	89.39 (26.00)	88.89 (25.57)	91.67 (19.25)	97.62 (8.91)
Non-Focal PM (%)	62.12 (38.89)	70.37 (41.05)	56.25 (35.94)	76.19 (33.15)
		Ongoing Category Decision Task Accuracy		
No PM (%)	95.54 (2.31)	95.96 (2.42)	96.29 (2.54)	95.62 (2.73)
Focal PM (%)	95.83 (2.07)	95.75 (1.95)	95.92 (1.78)	93.64 (8.15)
Non-Focal PM (%)	95.62 (1.92)	95.12 (2.33)	95.17 (2.55)	95.48 (2.23)
		Ongoing Category Decision Task Reaction Time		
No PM (ms)	1396.81 (291.31)	1404.57 (316.10)	1239.52 (159.75)	1359.82 (323.64)
Focal PM (ms)	1519.76 (316.99)	1472.48 (319.86)	1395.84 (179.76)	1513.19 (422.99)
Non-Focal PM (ms)	1887.80 (431.72)	1732.57 (405.14)	1726.73 (269.58)	1728.35 (424.93)

Table VI

Means and Standard Deviations for RVIP Task, Number of Correct Responses and Reaction Times (ms)

	Time Point 0				Time Point 1			
	Supervets		Controls		Supervets		Controls	
Bin	Correct (SD)	RT (SD)	Correct (SD)	RT (SD)	Correct (SD)	RT (SD)	Correct (SD)	RT (SD)
1	6.73 (1.37)	492.42 (63.46)	6.59 (1.79)	519.36 (80.17)	7.19 (0.75)	491.61 (61.21)	6.31 (1.89)	511.30 (72.91)
2	5.50 (2.18)	530.88 (64.54)	6.36 (2.13)	508.41 (75.77)	5.81 (1.40)	530.41 (66.31)	6.31 (2.21)	529.26 (56.10)
3	5.23 (1.86)	523.36 (77.11)	5.59 (1.92)	511.91 (72.07)	6.00 (1.58)	522.30 (75.73)	6.19 (2.34)	518.85 (64.30)
4	5.73 (1.37)	524.55 (52.13)	5.82 (1.76)	530.25 (75.12)	5.95 (1.36)	533.53 (76.36)	6.19 (2.26)	529.65 (75.14)
5	5.31 (2.09)	523.00 (52.81)	5.55 (2.26)	537.39 (87.84)	5.48 (1.57)	529.61 (72.74)	5.06 (2.29)	569.38 (90.75)

Table VII

Means and Standard Deviations for Stroop-Switch Task Reaction Times (ms) and Percent Errors

	Time Point 0				Time Point 1			
	Supervets		Controls		Supervets		Controls	
Condition	Mean RT (SD)	Mean % Error (SD)	Mean RT (SD)	Mean % Error (SD)	Mean RT (SD)	Mean % Error (SD)	Mean RT (SD)	Mean % Error (SD)
Color:								
Incongruent:								
Switch	1020.13 (303.15)	14.38 (16.06)	1018.68 (268.01)	18.63 (18.87)	925.13 (181.06)	12.56 (17.08)	986.26 (169.35)	11.17 (11.34)
Non-Switch	1035.75 (306.19)	14.96 (15.81)	1037.82 (219.88)	16.47 (15.07)	952.75 (186.47)	19.93 (18.61)	997.00 (202.16)	10.14 (8.02)
Neutral:								
Switch	912.16 (255.06)	6.65 (10.49)	926.89 (191.37)	9.52 (9.42)	811.12 (153.16)	8.26 (17.80)	876.61 (129.08)	1.93 (4.45)
Non-Switch	870.95 (210.24)	4.83 (10.18)	940.57 (219.85)	4.77 (9.02)	852.27 (168.91)	8.00 (13.75)	881.00 (193.76)	2.36 (4.58)
Word:								
Incongruent:								
Switch	795.40 (182.97)	3.55 (7.16)	878.62 (211.71)	2.34 (5.46)	755.82 (191.29)	3.75 (6.33)	829.84 (209.95)	5.03 (5.60)
Non-Switch	778.62 (240.93)	1.08 (4.00)	780.26 (153.86)	1.52 (4.18)	725.83 (162.13)	1.36 (4.45)	740.47 (124.01)	2.40 (4.90)
Neutral:								
Switch	739.40 (170.82)	8.29 (13.13)	872.67 (191.09)	6.16 (7.65)	752.27 (180.34)	5.93 (9.53)	804.11 (197.82)	7.15 (7.18)
Non-Switch	755.37 (160.76)	5.81 (10.67)	768.33 (157.91)	4.79 (6.98)	717.98 (160.54)	3.06 (7.39)	776.00 (156.15)	1.73 (3.98)

Funding

This work was supported by a studentship for Jeremy Young from the Sussex NHS Partnership Trust and the University of Sussex. The imaging costs were covered by a grant from Ageing Research Centre Sussex.

Acknowledgments

Nick Dowell contributed significantly to the development of all protocols for the structural image acquisition and analysis, and this side of the work would not have been possible without his input. Thanks to David Bunce, Keith Hutchison, Jennifer Breneiser, Mark McDaniel, and Jill Shelton for generously allowing us to use their tasks and guidance on implementation and analysis of these tasks. Thanks to Rebecca Graber and Nicolas Rothen for helpful comments on early drafts of this manuscript.