# Improving fitness increases dentate gyrus/CA3 volume in the hippocampal head and enhances memory in young adults

**Rachel K. Nauer**[a,b,c,d], **Matthew F. Dunne**[d], **Chantal E. Stern**[a,b,c], **Thomas W. Storer**[e], **Karin Schon**[a,b,c,d]

[a]Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215

[b]Center for Memory and Brain, Boston University, Boston, MA 02215

[c]Center for Systems Neuroscience, Boston University, Boston, MA 02215

[d]Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, MA 02118

[e]Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

## Abstract

Converging evidence suggests a relationship between aerobic exercise and hippocampal neuroplasticity that interactively impacts hippocampally-dependent memory. The majority of human studies have focused on the potential for exercise to reduce brain atrophy and attenuate cognitive decline in older adults, whereas animal studies often center on exercise-induced neurogenesis and hippocampal plasticity in the dentate gyrus of young adult animals. In the present study, initially sedentary young adults (18–35 years) participated in a moderate-intensity randomized controlled exercise intervention trial (ClinicalTrials.gov; NCT02057354) for a duration of twelve weeks. The aims of the study were to investigate the relationship between change in cardiorespiratory fitness (CRF) as determined by estimated $\dot{V}O_2^{MAX}$, hippocampally-dependent mnemonic discrimination, and change in hippocampal subfield volume. Results show that improving CRF after exercise training is associated with an increased volume in the left dentate gyrus/CA3 subregion in young adults. Consistent with previous studies that found exercise-induced increases in anterior hippocampus in older adults, this result was specific to the hippocampal head, or most anterior portion, of the subregion. Our results also demonstrate a positive relationship between change in CRF and change in corrected accuracy for trials requiring the highest level of discrimination on a putative behavioral pattern separation task. This relationship was observed in individuals who were initially lower-fit, suggesting that individuals who show greater improvement in their CRF may receive greater cognitive benefit. This work extends animal models by providing evidence for exercise-induced neuroplasticity specific to the neurogenic zone of the human hippocampus.

**Corresponding author:** Karin Schon, PhD, Department of Anatomy and Neurobiology, Boston University School of Medicine, 72 E Concord, R-1008B, Boston, MA 02118; kschon@bu.edu; Tel: 617-358-2118.

The authors declare no conflict of interest.

**Keywords**

aerobic exercise; human; memory; mnemonic discrimination; pattern separation

## 1 | Introduction

In the past several decades, studies in both rodents (Creer et al., 2010; Pereira et al., 2007; van Praag et al., 1999, 2005; Vivar et al., 2016) and humans (Chaddock et al., 2010; Erickson et al., 2009, 2011; Maass et al., 2015) have demonstrated evidence of dynamic structural and functional plasticity in the hippocampus that is positively related to aerobic fitness. Aerobic exercise powerfully influences hippocampal neuroplasticity by stimulating a cascade of changes in the local circuitry. In rodents, this includes increases in vascularization (van Praag et al., 2005), neurotrophic expression (Neeper et al., 1995; Park & Poo, 2013), dendritic spine density and arborization (Eadie et al., 2005; Stranahan et al., 2007), and synaptic plasticity (Christie et al., 2008; Vivar et al., 2016). In turn, these molecular and cellular changes lead to robust upregulation of adult hippocampal neurogenesis (AHN) in the dentate gyrus (DG) subregion (van Praag et al., 1999), and functional integration of newborn neurons into the existing neural network (Ramirez-Amaya et al., 2006).

Recent studies in humans have leveraged multimodal neuroimaging techniques to corroborate the positive relationship between aerobic exercise and markers of hippocampal integrity. Changes in hippocampal volume (Erickson et al., 2011; Smith et al., 2014; Thomas et al., 2016), vasoelasticity (Schwarb et al., 2017), white matter integrity (Burzynska et al., 2014), gray matter density (Kleemeyer et al., 2016) and cerebral blood flow (Maass et al., 2015; Pereira et al., 2007) are enhanced by aerobic exercise. Further examination has localized specified changes to the anterior division of the hippocampus in humans (Erickson et al., 2011; Maass et al., 2015; Thomas et al., 2016; Wagner et al., 2015, 2017). While this suggests the anterior hippocampus, or the hippocampal head, may be more receptive to exercise-induced neuroplasticity than more posterior regions in humans, possible disparate effects on the subfields have not been consistent across studies (Rosano et al., 2017; Varma et al., 2016), and delineation of specific effects have not been fully elucidated along the hippocampal long axis. Across species, this axis (termed dorsoventral in rodents) shows functional and structural gradations (Duvernoy, 2005; Poppenk et al., 2013; Strange et al., 2014; Thompson et al., 2008), including different effects on neurogenesis along this axis in rodents (Amrein et al., 2015). Although in rodents, basal levels of newborn neurons have been shown to be greater in the dorsal compared to ventral DG (Piatti et al., 2011; Snyder et al., 2009), altered neurogenesis via pharmacological agents or genetic manipulation has a greater effect on increasing neurogenesis in the ventral DG (Tanti & Belzung, 2013). This suggests that perhaps the ventral DG, which corresponds to the most anterior portion of the DG in humans, may demonstrate heightened exercise-induced neuroplasticity.

The specificity of the cognitive impact from exercise-induced hippocampal plasticity remains unclear. Although greater physical activity in humans has been linked to better global cognition including executive functioning (Colcombe & Kramer, 2003), evidence to

whether this association exists for hippocampally-dependent memory formation is mixed, as some studies but not others found such a link (Barnes et al., 2003; Déry et al., 2013; Gomez-Pinilla & Hillman, 2013; Griffin et al., 2011; Suwabe, Hyodo, Byun, Ochi, Yassa, et al., 2017; Voss, Prakash, et al., 2010). One explanation for these inconsistent findings may be that studies that have not found a relationship between exercise and memory mechanisms (Barnes et al., 2003; Gomez-Pinilla & Hillman, 2013; Voss, Prakash, et al., 2010) have used memory tasks that may not be reliant on the hippocampus. In contrast, studies that have used established hippocampally-dependent tasks (Déry et al., 2013; Griffin et al., 2011; Suwabe, Hyodo, Byun, Ochi, Yassa, et al., 2017) have found a positive association between cardiorespiratory fitness (CRF) or change in CRF and memory performance. Detection of exercise-induced cognitive enhancement will likely require focused, neurogenesis-dependent memory tasks that rely on newborn neurons within the DG, such as mnemonic discrimination tasks.

Data from computational models highlight an important functional role of newborn neurons in actively orthogonalizing the representations of similar stimuli and reducing memory interference (Aimone et al., 2011, 2009), a process known as pattern separation (Hasselmo & Wyble, 1997; O'Reilly & McClelland, 1994). These immature cells are particularly poised to perform this function due to their heightened excitability compared to their mature counterparts (Kee et al., 2007) and the unique sparse innervation from entorhinal cortex to the DG. Indeed, upregulation of adult hippocampal neurogenesis either by aerobic exercise (Creer et al., 2010) or exogenous induction via genetic manipulation (Sahay et al., 2011) facilitates behavioral efficiency on tasks requiring fine discrimination in rodents. Thus, mnemonic discrimination tasks that tap into pattern separation mechanisms may be sensitive to detect these effects in humans. Some evidence suggests a positive relationship between performance on mnemonic discrimination tasks following acute bouts of exercise (Suwabe et al., 2018; Suwabe, Hyodo, Byun, Ochi, Yassa, et al., 2017) and within the context of depression (Déry et al., 2013) in young adults. However, there remains a need for evidence of a clear link between exercise-induced changes on hippocampal subregion volume and memory performance measured using a behavioral paradigm sensitive to pattern separation.

In the present study, we enrolled healthy young adults (ages 18–35, N = 38) to participate in a twelve-week moderate-intensity exercise intervention designed to increase CRF, undergo two treadmill tests to estimate $\dot{V}O_2^{MAX}$ (a measure of CRF), two MRIs to measure hippocampal subfield volumes, and perform a mnemonic discrimination memory task that putatively taxes pattern separation mechanisms (for timeline, see Figure 1). To address how improving CRF specifically alters volume in hippocampal subregions that support mnemonic discrimination, participants underwent a structural MRI both prior to and following the intervention. Given previous work demonstrating a potential selectivity of an anterior hippocampal volume in response to exercise, it is critical to study exercise-induced hippocampal plasticity both by subregion and along the anterior-posterior axis of the hippocampus. Here, we partitioned out the hippocampal subfields and subdivided by head, body, and tail to test the prediction that exercise-induced changes in hippocampal volume is specific to the head of the DG/CA3 region in humans. We hypothesized that enhanced CRF

would predict an increase in hippocampal volume, specifically in the DG/CA3 subregion within the hippocampal head.

To address the question of whether improving CRF enhances mnemonic discrimination of highly similar stimuli, we developed a delayed-matching-to-sample task using novel face stimuli that participants performed while in the MRI scanner. All stimuli were rendered more similar to one another by objective manipulation and morphed to a trial-unique template at 10%, 30%, or 50% overlap, creating three parametrically manipulated similarity conditions. The stimuli at the highest level of similarity have objectively greater representational overlap with one another. Therefore, this condition in particular would theoretically rely on the DG for successful resolution of interference. Based on the converging literature across species, we predicted improvement in CRF following a twelve-week exercise training program would relate to enhanced mnemonic discrimination performance, particularly in trials with high levels of overlap between stimuli.

## 2 |   Materials and Methods

### 2.1 |   Participants and study design

This study was designed as a controlled twelve-week exercise intervention program and registered as a randomized controlled clinical trial (ClinicalTrials.gov; NCT02057354). Healthy young adults were recruited from Boston University and the surrounding community (N=110), and seventy-eight were enrolled. Eligible participants were English language speakers between the ages of 18 and 35 years with normal or corrected-to-normal vision, reported no past or current neurological or psychiatric disorders, and had no musculoskeletal, circulatory, or pulmonary conditions at the time of the study. Participants reported being sedentary prior to commencement of the study, defined as less than 30 minutes three times per week of moderate intensity physical activity over the previous three months (ACSM, 2014). Additional exclusionary criteria included MRI contra-indicators (e.g. ferro-magnetic metal in or on the body that could not be removed, claustrophobia), obesity, eating disorder, diabetes mellitus, pregnancy, presence of infection, diagnosis of kidney failure, liver  disease, thyroid disease, cancer, use of cardio- or psycho-active medication at time of study, alcohol or drug misuse or abuse, or recreational smoking. Pre-screening was completed over the phone, followed by an informed consent and full screening visit. All participants provided signed, informed consent prior to participation in the study. Protocols and study procedures were approved by the Boston University Medical Campus Institutional Review Board. Following determination of eligibility, participants were invited to participate in the twelve-week exercise intervention program and asked to complete baseline and follow-up visits to assess CRF via estimated $\dot{V}O_2^{MAX}$, and undergo MR imaging, and cognitive testing (see Figure 1 for study timeline). In addition to the MRI and cognitive task performed in the scanner, participants were also given the WAIS-IV digit span task to assess working memory performance (Strauss et al., 2006; see Supplemental Information for methods and results related to WAIS-IV).

Seven participants voluntarily withdrew following an initial screening visit due to time constraints. An additional five withdrew following baseline fitness and MRI visits, but prior

to randomization into an exercise group. For the exercise intervention, participants were randomized into either an endurance training (ET) program or resistance training (RT) program. Participants were given a completion bonus of an additional \$50 if they regularly attended the exercise sessions (30/36 sessions attended) and completed all MRI visits and fitness tests. Thirty-eight participants completed the entire intervention (mean age = 25.7 ± 3.3, 81.5% female), and are included in our analyses. No participant was excluded due to poor attendance of the exercise training program. Overall attendance at the exercise sessions for the included subjects was 81%, with a median of 30 sessions attended out of 36, indicating moderately high adherence to the exercise program. Demographic information for the participants included in the analyses can be found in Table 1. Due to a malfunctioning heart rate monitor during a fitness test, one participant did not have data for follow-up CRF assessment. For MRI analyses, five participants were excluded due to excessive head motion, which resulted in inaccurate manual segmentation of hippocampal subfields in either of their two MR volumes. This resulted in N=32 for any longitudinal comparisons of volume and fitness, and N=37 for any longitudinal comparisons of fitness and cognition.

## 2.2 | Fitness Assessments

### 2.2.1 | CRF Assessment
At baseline and following the twelve-week intervention, participants completed a submaximal-graded treadmill exercise test using a modified Balke protocol (ACSM, 2014; Cooper & Storer, 2001; Hagberg, 1994). All participants were instructed to refrain from strenuous physical activity for at least 24 hours prior to the test and to eliminate consumption of caffeine for three hours prior to the fitness assessment visits. Throughout the fitness assessment visits and each exercise session during the intervention, participants wore heart rate (HR) monitors (Polar, model H1) connected to a HR watch (Polar, model FT60). CRF was determined by estimating maximal oxygen uptake ($\dot{V}O_2{}^{MAX}$) based on the known linear relationship between heart rate and oxygen uptake ($\dot{V}O_2$) (Wasserman et al., 2012). Oxygen uptake was predicted from treadmill speed and grade at each minute of the test using standard, published equations (ACSM, 2014). Following a 3-minute warm-up, participants walked at a constant speed with incremental increases in grade. Tests were terminated when each subject reached 85% of age-predicted maximal heart rate ($HR_{max}$; Tanaka et al., 2001). This termination criterion was conservatively selected to be consistent with safety measures enacted for other participant cohorts (older adults ages 55–85 years; data not shown) in the clinical trial. Heart rate and blood pressure were monitored throughout the test and at termination. Participants provided their rating of perceived exertion every three minutes using the Borg 6–20 scale (Borg, 1982). Oxygen uptake for each minute of exercise was calculated using the American College of Sports Medicine (ACSM) metabolic equation for gross $\dot{V}O_2$ for walking (Equation 1):

$$\dot{V}O_2 = 0.1 \text{mL/kg/min} \times S + 1.8 \text{ mL/kg/min} \times S \times G + 3.5 \text{ mL/kg/min} \qquad 1$$

in which $\dot{V}O_2$, S, and G, and 3.5 represent gross oxygen uptake (ml/(kg •min)), treadmill speed (m/min), the percent of the treadmill grade expressed as a fraction, and the resting oxygen uptake, respectively (ACSM, 2014). The calculated $\dot{V}O_2$ was then plotted with its

corresponding HR at the end of each minute. From this, we calculated a linear regression using individual subject's predicted $HR_{MAX}$ (Tanaka et al., 2001), allowing us to estimate maximal aerobic capacity. At baseline, the average estimated $\dot{V}O_2{}^{MAX}$ was 36.18 ($\pm$ 6.62), suggesting low CRF, and did not significantly differ between participants subsequently randomized into the two exercise intervention groups (Table S1). The CRF assessment served as a basis for designing individualized exercise prescriptions for the ET group.

**2.2.2 |   Muscular Strength Assessment—**As a secondary fitness measure, muscular strength was assessed using the one repetition maximum (1-RM) method (Baechle et al., 2000). Briefly, this method requires performing sets of decreasing repetitions starting with eight, increasing loads, and increasing rest between sets until the maximal amount of weight that can be lifted through the full range of motion with good form one time only is determined. Four selectorized weight machines (Paramount, Los Angeles, CA) provided resistance for the leg press, leg curl, chest press and latissimus dorsi (lat) pull exercises. Baseline values for the 1-RM measurements were used to design the individual progressive resistance training (RT) programs (ACSM, 2014; Logan et al., 2000) and for comparison with follow-up measures to determine the training effect of the exercise intervention. Lower body strength ratio was considered as the amount of weight pushed during the leg press divided by body weight. Upper body strength ratio was calculated from the amount of weight pushed during chest press divided by body weight. All weight values were expressed in pounds. At baseline, mean 1-RM values for the two exercises were not different between exercise groups (Table S1).

**2.3 |   Exercise Interventions**

Participants were randomly assigned to participate in either an ET program designed to increase CRF or a RT program intended as an active control. The active control group was included to control for nonspecific aspects of the intervention, such as expectation bias, time commitment, social support, intervention duration, and attention from study staff (Kinser & Robins, 2013; Lindquist et al., 2007). Both groups met three times per week for a total duration of twelve weeks at Boston University Fitness and Recreation Center. Certified personal trainers supervised all training sessions. Initial fitness testing data from each participant was reviewed by an exercise physiologist (TWS), who created individualized exercise prescriptions for the trainers to administer. Fitness data, including heart rate (HR) and perceived exertion (Borg, 1982), were monitored for participant responses during each session. These responses along with specified plans for progression, set *a priori*, were used to update the individual training programs after week four and week eight. Sessions for both groups began and ended with a warm-up and cool down. Participants wore a chest strap (Polar, H1) and HR watch (Polar, FT7) during each session. Changes in CRF and muscular strength are summarized in Figure 2 and Table S1.

**2.3.1 |   Endurance Training—**Endurance participants completed twelve weeks of training in three progressive phases of four weeks each. In phase one, subjects began walking on a treadmill for 20 minutes with the goal of increasing walking duration to 30 minutes by the completion of week four. Intensity was set as a target heart rate zone of 60–70% age-predicted $HR_{MAX}$ (Tanaka et al., 2001). The phase two objective was to achieve 40

minutes of continuous walking at target heart rates of 70–80% $HR_{MAX}$ by the end of week eight. Phase three prescriptions further increased the intensity of exercise to target heart rates of 80% $HR_{MAX}$ while maintaining the 40-minute duration used in phase two. Treadmill speed and incline were adjusted for each participant to maintain the target heart rate zone appropriate for each phase. This typically resulted in increases of the speed and/or grade as participants increased their CRF. Trainers recorded treadmill speed and incline for each participant, as well as HR and ratings of perceived exertion using the Borg Scale (Borg, 1982).

**2.3.2 | Resistance Training**—Participants in the resistance group exercised three days per week for twelve weeks with progressive increases in target loads and number of sets for eight different exercises targeting major muscle groups. Four of the resistance training exercises included those used during fitness testing (chest press, lat pulldown, leg curl, and leg press) and were supplemented with an additional four exercises of the trainer's choice to target additional muscle groups as recommended in current guidelines (Garber et al., 2011). Three four-week phases were established *a priori*. During phase one (weeks 1–4) of the program, participants performed one set of 12–15 repetitions at approximately 60–70% of the 1-RM achieved at baseline. This increased during phase two (weeks 5–8) to two sets of 12–15 repetitions using resistance that was determined by participant progression, and was again increased during phase three (weeks 9–12) to three sets of 8–12 repetitions.

## 2.4 | Mnemonic Discrimination Task

**2.4.1 | Task Materials and Stimuli**—Subjects performed a delayed-matching-to-sample task using novel face stimuli while undergoing fMRI scanning. The current study reports behavioral results only. Stimulus presentation and participant response data were displayed and collected using EPrime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The stimulus set consisted of grayscale images of central facial features of non-famous faces by eliminating peripheral features (e.g. hair, clothes). Ten runs containing twelve trials per run were created and split into two sets of five runs (Set A and Set B) constituting 60 trials per set. Each set of stimuli were counterbalanced across participants so half the participants performed Set A during baseline testing and Set B during follow-up testing, and the other half of participants performed Set B followed by Set A.

Critically, stimuli within a trial were rendered more similar to one another by objective manipulation using face morphing software (Fantamorph 5). All original stimuli within a trial were morphed to a trial-unique template at 10% similarity, 30% similarity, or 50% similarity to generate the stimuli used in the task and create three conditions (Figure 4A). Therefore, all stimuli within a trial shared some common features. Trial condition was pseudo-randomized for each run and participant so that an even number of each condition appeared in each run. Immediately prior to beginning the task, subjects were given detailed instructions and performed one practice run with a distinct set of stimuli. Response feedback was given for each trial during the practice. During test, accuracy was summarized and displayed at the end of each run.

**2.4.2 |   Task Procedure**—Participants performed the delayed-matching to sample task that relied on discrimination of similar stimuli for success while in the MRI scanner (see Figure 4B for example trial). Each trial consisted of two face stimuli presented sequentially for two seconds (Sample 1: 2 s, Sample 2: 2 s) followed by a black background and grey silhouette of a face for a ten-second delay period (Delay), and finally a third face stimulus presented for a total of two seconds (Test). All stimuli within a trial were morphed to a trial-unique template at either 10% similarity, 30% similarity, or 50% similarity. Thus, for example, on a non-match 30% similarity trial (Figure 4B), both Sample 1 and Sample 2 as well as the Test face shared 30% of the stimulus features derived from their trial-unique template. Participants were instructed to make one of three button-press responses during the test face stimulus presentation: 1) the test face displayed was the same face stimulus as Sample 1 (match), 2) the test face stimulus displayed was the same face stimulus as Sample 2 (match), or 3) the test face stimulus displayed a new face stimulus that was not previously seen (non-match). A variable length (6–12 s) inter-trial interval separated each trial. Trials were evenly split into match and non-match; match trials were split evenly across Sample 1 and Sample 2.

## 2.5 |   MRI Parameters and Subfield Segmentation

Participants were scanned within one month prior to the onset of the exercise intervention and again within three weeks after completion. Imaging data were collected on a 3 Tesla Phillips Achieva MRI scanner equipped with an 8-channel SENSE head coil at Boston University Center for Biomedical Imaging. A high-resolution whole-brain structural T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE; SENSitivity Encoding P reduction: 1.5, S reduction: 2; TR = 6.7 ms, TE = 3.1 ms, flip angle = 9°, Field of View = 25 cm, Matrix Size = 256 × 254, 150 slices, resolution = 0.98 mm × 0.98 mm × 1.22 mm) volume was acquired for each participant. Additionally, coronal sections were obtained perpendicular to the long axis of the hippocampus in a structural T2-weighted image volume with higher in-plane resolution (SENSitivity Encoding P reduction: 2, TR = 3000 ms, TE = 80 ms, flip angle = 90°, Field of View = 25 cm, Matrix Size = 576 × 450, 30 slices, resolution = 0.4 mm × 0.4 mm × 2.0 mm with a slice gap of 0.6 mm).

Regions of interest (ROIs) included the hippocampal subfields (left and right CA1, DG/CA3, subiculum). ROIs were anatomically determined and delineated using ITK-SNAP (version 3.4.0; www.itksnap.org; (Yushkevich et al., 2006)) in native space on coronal T2-weighted MR images by one researcher (MFD) as described in previous work (Nauer et al., 2015) and following previously published guidelines (Head et al., 2008; Pruessner et al., 2002, 2000). Hippocampal subfield boundaries were defined using the Duvernoy atlas (Duvernoy, 2005) and were separated into head, body and tail sections. Separation of head and body was decided by the disappearance of the uncal apex (gyrus intralimbicus) (Daugherty et al., 2015; Pruessner et al., 2000). Separation between body and tail were delineated by the presence of the wing of the ambient cistern, with slices posterior to its presence labeled as tail. Slices deemed to be in the tail section of the hippocampus were not separated by subfields due to uncertainty of subfield boundaries in the most posterior aspects of the region given the resolution. Figure 3A shows representative subfield delineations in the head and body. Intracranial volume (ICV) was measured manually using a 1-in-12 sampling

strategy (Eritaia et al., 2000) and used as a correction for each regional volume prior to statistical analyses. ICV adjustment was performed for each ROI in each hemisphere via a formula based on the analysis of covariance approach well established in the literature (Erickson et al., 2009; Head et al., 2008; Raz et al., 2004), adjusted volume = raw volume − $b \times$ (ICV − mean ICV), where $b$ is the slope of regression of an ROI volume on ICV. Intraclass correlation coefficient for ICV was determined between two raters (MFD, RKN), and was extremely high (ICC = 0.989, $F_{(71,72)} = 176$, $p < .0001$).

## 2.6 | Statistical Methods

Statistical analyses were conducted using R (version 3.1.1). Continuous variables were summarized by means and standard deviations. Estimated $\dot{V}O_2^{MAX}$ values, cognitive task performance, and ROI volumes were tested for normality using Shapiro-Wilk normality test prior to analyses. All variables at baseline, follow-up as well as the difference score between the two time points all assumed normality. Within-group analyses were completed using paired t-tests from time 1 (baseline) to time 2 (follow-up). Simple correlations were performed using Pearson's $r$. Repeated-measures ANOVA or multiple linear regression models were used to compare group (ET, RT) over time. Models for change in fitness measurements included sex, age, and exercise group as covariates. Multiple linear regression models were used to elucidate the relationship between change in CRF and change in dependent measures and included sex, age, education as covariates, and corrected for multiple comparisons using the Bonferroni method. We used the the *p.adjust* function in R to correct observed p-values for multiple comparisons and report adjusted p-values. The Bonferroni-corrected (i.e. adjusted) p-values are calculated in *p-adjust* by multiplying the observed p-values by the number of comparisons.

**2.6.1 | Fitness Analyses—**CRF was operationalized using estimated $\dot{V}O_2^{max}$ (mL/kg/min). Baseline comparisons between groups for both CRF and muscular strength were investigated using t-tests. Interventional assessment of the exercise groups on CRF was accomplished using repeated measures Analysis of Variance (rm-ANOVA) with estimated $\dot{V}O_2^{MAX}$ from baseline and follow-up as the repeated measure, training group (ET, RT) as a between-group factor, time (baseline, follow-up) as a within-subject factor, and sex, and age as covariates. Changes in CRF were determined as the difference in estimated $\dot{V}O_2^{MAX}$ between baseline and follow-up, and are reported in Table S1 and Figure 2B. Changes in leg press, leg curl, chest press, and lat pulldown strength were calculated as percent change from baseline for each exercise.

**2.6.2 | Volumetric Analyses—**Raw volumes were extracted from ITK-SNAP for each hippocampal subfield ROI in $mm^3$, converted to $cm^3$ and corrected for differences in ICV as detailed above. Each subfield (DG/CA3, CA1, and subiculum) was separated by hemisphere and subsectioned into head and body, resulting in twelve subfield ROIs. First, we investigated the potential effect of group on ROI volume across time using rm-ANOVA, including education, age, and sex as covariates. Next, we collapsed across groups and examined how change in CRF related to change in ROI volume. The delta change scores

reflected the percent change in volume from baseline and the change in CRF (in ml/kg/min for estimated $\dot{V}O_2{}^{MAX}$) from baseline. Simple correlations were examined first, using Pearson's correlations to examine the relationship between change in CRF and change in ROI volume, with Bonferroni corrections for laterality. Next, any significant or marginally significant correlations were assessed when controlling for the effect of sex, age, and education using multiple regression models with Bonferroni corrections for multiple comparisons. We also sought to confirm that improvements in strength following the exercise intervention did not have effects on hippocampal volume or cognition (see Supplemental Materials for methods and results).

**2.6.3 | Behavioral Analyses**—The primary outcome variable for the mnemonic discrimination task was corrected accuracy. This was defined as the difference between the rates of correct responses (hits) during a match trial (e.g. response of "1" when the test face stimulus matched the first face stimulus in the trial) minus incorrect responses (false alarms) on non-match trials (e.g. response of "1" or "2" when the test face stimulus was new). This removes any potential effect of subject response bias. For the majority of analyses, we separated out trials by similarity level (10%, 30%, 50%) to investigate the effect of stimulus similarity. To test the prediction that initially lower-fit individuals may reap greater cognitive benefit by improving CRF compared to initially higher-fit individuals, we also collapsed data across both exercise groups and subdivided participants by pre-intervention CRF level, using a median split of baseline estimated $\dot{V}O_2{}^{MAX}$ similar to methods used in other studies (Suwabe, Hyodo, Byun, Ochi, Fukuie, et al., 2017). We utilized rm-ANOVA methods for analyses of group, initial fitness level, and time effects on cognition and multiple regression to determine relationships between change in CRF and change in task performance, including education, age, and sex as covariates and applied Bonferroni corrections to account for similarity level.

# 3 | Results

## 3.1 | CRF improves following the exercise intervention.

Participant demographic information is summarized in Table 1 and fitness data is summarized in Table S1. Prior to analyzing effects of improving CRF on ROI volumes or cognition, we analyzed baseline differences in CRF or strength between groups as well as the effectiveness of the exercise interventions to improve CRF.

**3.1.1 | Measures of fitness at baseline**—To determine if there were baseline differences in fitness measures between the two exercise groups, we performed a Welch two-sample t-test for each variable (see Table S1 for baseline group comparison). Importantly, there were no statistically significant group differences at baseline across any measures (all $p > .1$), including CRF, defined as estimated $\dot{V}O_2{}^{MAX}$ (endurance training (ET) group mean: $35.26 \pm 4.6$, resistance training (RT) group mean: $36.93 \pm 7.9$, $t(32.9) = -0.82$, $p = 0.42$). Both groups had variability in fitness levels (see Figure 2A). There were no baseline differences between the intervention groups for any of the measures of strength, measured as the 1-RM weight in pounds (see Table S1 for average 1-RM scores; all $p > .05$), and on other

measures of fitness, including resting HR (ET: $82 \pm 9.5$; RT: $82.5 \pm 7.2$; $t(29.4) = 0.16$, $p = 0.87$) or baseline body mass index (ET: $23.6 \pm 4.18$; RT: $25.7 \pm 5.2$, $t(36) = 1.35$, $p = 0.19$). Additionally, there was no difference in attendance at the exercise sessions between groups, with the ET group having attended an average of 80% of sessions and the RT group attending 81% of sessions ($t(32.6) = 0.26$, $p = 0.79$).

**3.1.2 | Changes in cardiovascular fitness**—Within the ET group, there was a modest effect comparing baseline and follow-up CRF (baseline mean $35.26 \pm 4.6$, follow-up mean $37.67 \pm 4.4$; $t(16) = 3.25$, $p = 0.005$), suggesting the ET intervention was effective in improving CRF. We assessed the interaction of group X time on CRF controlling for age and sex using a repeated-measures ANOVA and found a significant main effect of time ($F_{1,35} = 11.6$, $p = 0.002$), but no main effect of group ($F_{1,34} = 0.31$, $p = 0.58$) and no interaction between group X time ($F_{1,35} = 1.2$, $p = 0.28$). This indicated that participants improved CRF over time, but there was not a statistically significant effect of exercise group on improving CRF. A planned t-test revealed a trend towards the ET group showing greater improvement in CRF than the RT group (ET group: mean CRF increase of $2.41 \pm 3.1$, indicating an average of a 7.4% increase from baseline; RT group: mean CRF increase of $1.26 \pm 3.3$, indicating an average of a 4.4% increase from baseline). However, this result did not reach statistical significance ($t(34.7) = 1.1$, $p = 0.28$; Figure 2B), likely due to the large variability in change in CRF following exercise training (Lundby et al., 2017). The trend for improvement in CRF from baseline for the RT group is perhaps not surprising given the study design of using an active exercise group, and is consistent with a recent meta-analysis that demonstrated RT can significantly improve $\dot{V}O_2^{MAX}$ compared to non-active controls (Yamamoto et al., 2016). Participants were not prohibited from engaging in other physical activity outside of the exercise intervention program. This may have contributed to the lack of group effect. Therefore, similar to previous exercise intervention studies (Jonasson et al., 2016; Kleemeyer et al., 2016), for the remainder of analyses we collapsed across ET and RT groups and operationalized change in CRF as a predictor of changes in regional brain volume and cognition.

**3.1.3 | Changes in Strength**—Average percent change from baseline by group for the strength measures is shown in Figure 2C. As expected, the RT group significantly improved their 1-RM from baseline on all four of the resistance exercises assessed with paired t-tests (chest press: $t(20) = 6.39$, $p < 0.001$; lat pulldown: $t(20) = 6.43$, $p < 0.001$; leg press: $t(19) = 5.41$, $p < 0.001$; leg curl: $t(19) = 6.45$, $p < 0.001$). The ET group did not show significant changes in strength on three of the four strength measures (chest press: $t(16) = 0.71$, $p = 0.49$, lat pulldown: $t(15) = 0.32$, $p = 0.75$, leg curl: $t(16) = 1.86$, $p = 0.08$, see Table S1). The ET group did significantly improve strength on leg press ($t(16) = 2.23$, $p = 0.04$). We also ran multiple linear regression models for the percent change from baseline separately on all strength tests, including exercise group (ET, RT) as the predictor and sex as a covariate. Results from this analysis demonstrate a main effect of exercise group for all measures of strength (chest press ($F_{1,35} = 38.77$, $p < 0.001$); lat pull ($F_{1,33} = 29.5$, $p < 0.001$), leg press ($F_{1,34} = 9.67$, $p = 0.004$) and leg curl ($F_{1,34} = 15.9$, $p < 0.001$), such that the RT group improved strength from baseline significantly more than the ET group for all four 1-RM tests (see Figure 2C and Table S1).

### 3.2 | Positive change in CRF predicts an increase in DG/CA3 volume.

First, we examined the relationship between estimated $\dot{V}O_2{}^{MAX}$ and volume in each ROI at baseline. We found a significant positive relationship between estimated $\dot{V}O_2{}^{MAX}$ and volume only for the head of left DG/CA3 (r = 0.356, $t(31)$ = 2.12, $p$ = 0.04). No other ROI volumes correlated with estimated $\dot{V}O_2{}^{MAX}$ at baseline. Additionally, we investigated whether group membership (ET, RT) predicted volume change over time. We entered ROI volume data into a rm-ANOVA with time (baseline, follow-up) as the within-subject factor and group (ET, RT) as the between-subject factor, and entered age, sex and education as covariates. We found no interactive effect between group and time for any ROI analysis, indicating group membership did not relate to change in ROI volume. This was not surprising, given that the two groups did not differentially increase CRF. Therefore, we collapsed across groups to elucidate the relationship between exercise-induced changes in CRF and potential volumetric changes in the hippocampus, specifically in the head of the DG/CA3 region. To examine this, we subdivided the hippocampus by head, body and tail, and partitioned out the hippocampal subfields within the head and body only. Then, we compared the percent change in volume from baseline in each ROI to the change in CRF again using simple Pearson correlations, with Bonferroni corrections for hemispheric laterality as we had no a-priori hypotheses about hemisphere. There was a positive correlation between change in CRF and change in the ICV-corrected volume in the left head of DG/CA3 ($r$ = 0.413, $t(30)$ = 2.49, $p_{adjusted}$ = 0.037; Figure 3B, *left*). Additionally, we found marginally significant negative correlations between change in CRF and change in ICV-corrected volume in the left head of subiculum ($r$ = –0.365, $t(30)$ = 2.15, $p_{adjusted}$ = 0.079; Figure 3B, *middle*) and a marginally significant negative correlation between change in CRF and change in ICV-corrected volume in the body of right CA1 (r = –0.393, $t(30)$ = 2.34, $p_{adjusted}$ = 0.052; Figure 3B, *right*). No other regions showed any relationship between change in CRF and change in volume. Next, we assessed these relationships using multiple regression models that controlled for sex, age, and education and applied Bonferroni corrections for multiple comparisons based on the number of statistical models. Addition of covariates mitigated the marginally significant relationships between change in CRF and ROI volume for the right body of CA1 ($F_{4,27}$ = 1.37, $t(27)$ = 1.41, $p_{adjusted}$ = 0.51) and the left head of the subiculum ($F_{(4,27)}$ = 2.14, $t(27)$ = 2.11, $p_{adjusted}$ = 0.135). In contrast, the relationship between change in estimated $\dot{V}O_2{}^{MAX}$ and change in volume in the left head of DG/CA3 remained robust ($F_{4,27}$ = 1.74, $t(27)$ = 2.6, $p_{adjusted}$ = 0.048, Figure 3B, *left*). This result is consistent with our *a-priori* hypothesis, and suggests that a positive change in fitness predicts a positive change in volume unique to the head of the left DG/CA3.

### 3.3 | Change in CRF predicts mnemonic discrimination when stimulus similarity is high.

Another main objective of the current study was to determine if improvement in CRF following the exercise intervention can predict enhanced performance on a mnemonic discrimination task with varying levels of stimulus similarity. First, for each stimulus similarity condition (10% similarity, 30% similarity, or 50% similarity, see Figure 4A), we ran multiple linear regression models to test if there were main effects of baseline CRF or baseline group differences (ET, RT) on task accuracy, and included age, education and sex as

covariates. Results from the multiple linear regression models showed that baseline levels of CRF did not predict corrected accuracy on the task for any condition (10% similarity: $F_{1,32}$ = .97, $p = 0.33$; 30% similarity: $F_{1,32} = .05$, $p = 0.82$; 50% similarity: $F_{1,32} = 2.12$, $p = 0.16$). Importantly, results confirmed there was no main effect of group at baseline on task accuracy for any condition (10% similarity: $F_{1,32} = .04$, $p = 0.84$; 30% similarity: $F_{1,32} =$ 1.88, $p = 0.18$; 50% similarity: $F_{1,32} = 1.48$, $p = 0.23$), demonstrating that ET and RT groups were equivalent at baseline. Performance on the task at baseline was not significantly different for match (average match accuracy: $0.79 \pm 0.16$ for 10% similarity; $0.75 \pm 0.18$ for 30% similarity) compared to non-match (average accuracy: $0.84 \pm 0.15$ for 10% similarity; $0.7 \pm 0.19$ for 30% similarity) trials at the 10% and 30% similarity levels. As expected, in the 50% condition, accuracy was significantly higher for match (mean accuracy: 0.69 $\pm 0.18$) compared non-match (mean accuracy: $0.53 \pm 0.32$; $t(37) = 3.69$ $p < 0.001$) trials, reflecting greater difficulty on trials with higher interference (e.g. increased load of pattern separation for high overlap conditions). To correct for response bias, the primary outcome variable was corrected accuracy (correct match trial "hits" – incorrect non-match "false alarms"). This is a standard approach to remove any potential effects of subject response bias.

To investigate the effect of the exercise intervention and parametric manipulation of the stimulus similarity level over time on corrected accuracy, we entered data into a rm-ANOVA with stimulus similarity (10%, 30%, 50% trials) and time (baseline, follow-up) as within-subject factors and group (ET, RT) as the between-subject factor, and entered age, sex and education as covariates. Results indicated a main effect of stimulus similarity ($F_{2,180} =$ 71.82, $p < 0.0001$; Figure 4C), but no main effect of time ($F_{1,180} = 3.2$, $p = 0.08$) or exercise group ($F_{1,32}$) = 2.36, p < 0.14). No interaction effects were significant (similarity X time: $F_{2,180} = .46$, $p = 0.63$; similarity X group: $F_{2,180} = .51$, $p = 0.60$); group X similarity X time; $F_{2,180} = .98$, $p = 0.38$), suggesting that exercise group did not have differential effects on corrected accuracy. Post-hoc testing with Bonferroni correction identified that all 3 conditions (10% similarity, 30% similarity, 50% similarity) were highly significantly different from one another (Figure 4C). We found participants performed best on the lowest interference trials (10% trials; mean corrected accuracy = $0.60 \pm 0.21$) compared to both the 30% trials (mean corrected accuracy = $0.45 +/- 0.23$; $p < 0.001$ compared to 10%) and 50% trials (mean corrected accuracy = 0.19; $p < 0.001$ compared to 10% trials). Post-hoc comparison between 30% and 50% demonstrates significantly lower performance on 50% trials ($p < 0.001$). This indicates that subjects performed more poorly on the task as stimulus similarity, and thus interference, increased.

Since participants in both exercise groups improved their CRF, we reasoned that change in CRF following the intervention should be more predictive of change in task accuracy than group membership per se. We found that there was no significant correlation between changes in CRF and changes in task performance across all participants (10% similarity: r = −0.07, $p = 0.67$; 30% similarity: r = −0.11, $p = 0.53$; 50% similarity: r = 0.13, $p = 0.46$). We therefore tested the prediction that individuals who were lower-fit at baseline may reap a greater cognitive benefit by improving their CRF. We collapsed data across both exercise groups and subdivided participants by pre-intervention CRF level, using a median split of

baseline estimated $\dot{V}O_2^{MAX}$ (median estimated $\dot{V}O_2^{MAX}$: 36.1) (Suwabe, Hyodo, Byun, Ochi, Fukuie, et al., 2017). This allowed us to separate out aerobically higher-fit individuals from lower-fit individuals in our sample at the start of the intervention. First, we tested whether initial fitness category predicted performance on the mnemonic discrimination task again using a rm-ANOVA. Importantly, results confirmed a lack of a main effect of initial fitness level ($F_{1,33} = 0.45$, p = 0.5), indicating there were no differences in cognitive performance between participants that were initially higher-fit and those that were initially-lower fit. This indicates that initial CRF alone did not relate to cognitive performance on this task. We hypothesized that successful mnemonic discrimination would be most positively affected by greater improvement in CRF in initially lower-fit individuals. Using multiple regression, we analyzed the effects of baseline fitness level (higher-fit, lower-fit) and change in CRF on change in corrected accuracy, including age, sex, and education as covariates, and applied Bonferroni corrections for the three similarity levels. Additionally, based on previous literature (Creer et al., 2010), we hypothesized that this interaction may be most apparent when task demand for discrimination is high (e.g. 50% similarity). As hypothesized, the interaction between change in CRF and initial fitness level successfully predicted change in corrected accuracy only in the 50% similarity condition ($F_{1,30} = 13.6$, $p = 0.002$, $p_{adjusted} = 0.006$; see Figure 4D). There was no interactive effect between change in CRF and initial fitness level for either the 30% ($F_{1,30} = 2.8$, $p = 0.12$, $p_{adjusted} = 0.36$) or the 10% ($F_{1,30} = 0.46$, $p = 0.51$, $p_{adjusted} = 1$) similarity conditions. This suggests that improving CRF positively impacts performance on a task that requires very fine discrimination, and that lower-fit individuals may gain the greatest cognitive benefit by improving their CRF.

## 4 | Discussion

In the present study, we examined the effects of improving CRF on hippocampal subfield volumes and behavioral performance on a mnemonic discrimination task that putatively relies on pattern separation. We tested the hypothesis that increasing CRF following a 12-week exercise training program would correlate with an increase in hippocampal volume specific to the DG/CA3 subregion. Critically, we demonstrate the hypothesized positive relationship between improved CRF and increased DG/CA3 volume to be present specifically in the head, or anterior most portion, of the hippocampal long axis. We found no relationship between changes in CRF and volume change in the body of DG/CA3, head or body of CA1, the body of subiculum or the hippocampal tail. Change in CRF negatively correlated with volume in the head of the subiculum, although the addition of covariates mitigated this relationship. Behaviorally, consistent with and predicted by our volumetric finding, we found that initially lower-fit individuals who improved their CRF following a twelve-week exercise intervention demonstrated significantly enhanced behavioral performance on the mnemonic discrimination task when task demands for fine discrimination were high.

### 4.1 | Exercise-induced volumetric change in the hippocampus is unique to the DG/CA3 subregion

The presence of AHN in the DG is well accepted in rodents (Altman & Das, 1965) and non-human primates (Gould et al., 1999; Kornack & Rakic, 1999). Studies in rodents have

reliably demonstrated that exercise both increases AHN in young animals (van Praag et al., 1999) and attenuates age-related decline of AHN (van Praag et al., 2005). Conversely, postmortem studies in humans have shown conflicting results regarding the presence of AHN. The first evidence of AHN in postmortem human tissue was found in a study by Eriksson and colleagues, demonstrating that neurons continue to generate in the human DG into late adulthood (Eriksson et al., 1998). Recently, controversy on this topic has been reignited by two studies showing data with contradictory results. Sorrells and colleagues examined samples of postmortem hippocampal tissue from 59 individuals across the lifespan ranging from fetal development to late adulthood and found high rates of cell proliferation in the DG during development that dropped precipitously postnatally and were not present in most specimens collected from individuals past the age of adolescence (Sorrells et al., 2018). In contrast, Boldrini and colleagues found a substantial presence of intermediate neural progenitors and immature neurons in the DG of humans posthumously (Boldrini et al., 2018). The authors of this study also found stability in neurogenesis levels that persisted into the eighth decade of life.

We cannot currently track AHN *in vivo* in humans, however, support from neuroimaging studies, including the results presented here, has contributed important indirect evidence toward the presence of exercise-induced hippocampal plasticity in humans. A cross-species study demonstrated increased cerebral blood volume specific to the DG following aerobic exercise training in both rodents and humans and correlated with AHN in the same region in rodents (Pereira et al., 2007). Increased regional cerebral blood flow and volume to the hippocampus in response to an aerobic exercise intervention has been extended to older adult populations (Maass et al., 2015), although this upregulation may be reduced with advanced age. In rodents, AHN and angiogenesis are known to be tightly co-regulated (Palmer et al., 2000), and the same has been shown in human studies (Boldrini et al., 2012). Volumetry may provide useful as a proxy for plasticity, and previous reports have demonstrated a positive relationship between aerobic exercise and markers of hippocampal integrity (Burzynska et al., 2014; Erickson et al., 2011; Kleemeyer et al., 2016; Maass et al., 2015; Pereira et al., 2007; Schwarb et al., 2017; Thomas et al., 2016). In older adult humans, greater CRF and improving CRF have been linked to increased gray matter in the hippocampus (Erickson et al., 2009, 2011) and greater functional connectivity (Voss, Prakash, et al., 2010). Clinical trials and cross-sectional studies in humans show increased hippocampal volume or attenuated hippocampal atrophy following aerobic exercise (Erickson et al., 2009, 2011; Honea et al., 2009; Pereira et al., 2007; Smith et al., 2014). In addition to aerobic exercise, measures of physical activity such as daily walking also correlate with hippocampal volume in older women (Varma et al., 2015). This suggests that baseline levels of physical activity and CRF in older adults have a relationship to hippocampal volume. However, at the subfield level, addressing the relationship between physical activity and hippocampal integrity have shown mixed results in older adults. Cross-sectional data from older women show a positive relationship between amount of daily walking and surface area of the left subiculum (Varma et al., 2016). Data from another study in older adults show that volume in the left hippocampus, particularly the CA fields, had a positive relationship with physical activity over a 24-month period, although the subfield specific results were attenuated upon adjustment for baseline differences (Rosano et al.,

2017). The vast majority of these previous studies have investigated these questions within the context of aging, when there are known changes to brain structure. Here, we investigated the effect of CRF on brain plasticity in young adulthood, during a time of relative stability in brain structure (Voss et al., 2011).

One of the strengths of our study is the longitudinal design which allows for detection of subtle within-subject changes as a result of improving CRF. Our data show that young adults who show greater improvement in CRF show greater volume change in the DG/CA3 subregion, specifically in the most anterior region. This result is consistent with a previous exercise intervention study in young adults that showed detectable volumetric changes in the anterior hippocampus in as little as six weeks of high-intensity interval exercise training (Thomas et al., 2016). Our results extend this work by specifying the anterior portion of the DG/CA3 shows a positive relationship between volume and CRF improvement following our twelve-week exercise protocol. Interestingly, in the study by Thomas and colleagues, the anterior volume change was shown to be transient once exercise sessions ceased, as determined by a follow-up MR scan six weeks post-intervention (Thomas et al., 2016). Our study did not include additional follow-up MR scans extending past three weeks subsequent to the termination of the exercise program. More work should be done to assess the amount of exercise needed as well as length of intervention (i.e. dose-dependency of exercise) required to induce a sustained volume change. While techniques like MRI cannot definitively prove the existence of AHN, human neuroimaging data, including data presented in the current study, lend compelling indirect and converging evidence in support of heightened plasticity in this specific subregion in healthy young adults who respond to moderate-intensity aerobic exercise.

## 4.2 | Observed relationship between volume and CRF is specific to the anterior hippocampus

We show that the anterior-most region of the hippocampus positively responds to change in CRF, a finding that is complementary to literature in older (Erickson et al., 2011; Maass et al., 2015) and younger to middle aged adults (Herting & Nagel, 2013; Thomas et al., 2016; Wagner et al., 2015, 2017). We present evidence that the anterior hippocampus is not homogenously responsive to exercise-induced changes. Our data show that in young adults, this exercise-induced effect is restricted to the DG/CA3 subregion. Conversely, change in CRF had little effect on any subfields within the body and no effect on the hippocampal tail. This specificity is consistent with previous reports showing positive relationships between CRF, exercise, and changes to the most anterior division of the hippocampus in humans (Erickson et al., 2011; Maass et al., 2015; Pereira et al., 2007; Thomas et al., 2016; Varma et al., 2016; Wagner et al., 2015, 2017), with no effect of aerobic exercise and CRF on the body and tail of the hippocampus.

Evidence from rodent studies demonstrate differential effects along the long axis of the hippocampus for exercise-induced AHN. Work in rodents has identified a greater amount of basal levels of newborn neurons in the dorsal/septal DG (analogous to the posterior DG in humans) and a larger pool of mature neurons in the ventral portion (analogous to the anterior DG) (Piatti et al., 2011; Snyder et al., 2009; Tanti et al., 2013). Antidepressant treatment

seems to increase AHN specifically in the rodent ventral DG (Tanti et al., 2013), and voluntary exercise enhances neuronal development in the rodent ventral hippocampus (Piatti et al., 2011), suggesting that in humans this anterior subregion shows amenable plasticity and ability to increase the pool of newborn neurons. In animal models, the ventral hippocampus in particular may be more vulnerable to factors such as stress and depression, and these effects can have substantial consequences on local neuronal development during the stages of neurogenesis (Tanti & Belzung, 2013). It is possible that the hippocampus has a gradient of susceptibility to plasticity along its long-axis. Our data, together with the existing literature suggest that the anterior portion of the hippocampus has greater susceptibility to plasticity.

## 4.3 | Improving CRF enhances performance on memory tasks requiring fine discrimination

Our longitudinal analyses of the behavioral data revealed that improving CRF, particularly in initially lower-fit individuals, enhanced mnemonic discrimination for the most difficult trials containing high overlap and thus elevated levels of interference. The most difficult condition (i.e., 50% similarity) is likely placing a greater demand on the DG to successfully pattern separate highly overlapping stimuli. Our result importantly connects improving CRF with enhanced behavioral performance on tasks requiring fine discrimination, and is consistent with previous reports in rodents demonstrating that increased AHN favorably enhances performance when the task demands very fine discrimination (Creer et al., 2010; Sahay et al., 2011). In addition, we found no relationship between CRF at baseline and performance on the WAIS-IV digit span task. There was also no relationship between change in CRF and change in performance on any aspect of the digit span task (see Supplemental Information), suggesting that improving CRF does not enact global changes in memory performance.

Exercise intervention studies in children (Chaddock et al., 2010) and older adults (Erickson et al., 2011) have demonstrated a positive relationship between improving CRF and enhanced cognition, though the cognitive processes tested have varied and have not necessarily been dependent on the hippocampus or AHN. One previous study in healthy younger adults has also shown that there is a relationship between improved CRF and behavioral pattern separation following six weeks of high-intensity interval training (Déry et al., 2013). Our results demonstrate a similar relationship between improving CRF following a twelve-week moderate intensity exercise program and enhanced performance on a mnemonic discrimination task with healthy young adults. Recent evidence also suggests that mnemonic discrimination may be enhanced in healthy young adults with just a single acute bout of moderate intensity (Suwabe, Hyodo, Byun, Ochi, Yassa, et al., 2017) and light intensity aerobic exercise, likely by rapidly increasing functional connectivity between the DG/CA3 subregions and the neocortex (Suwabe et al., 2018) rather than by upregulating AHN. Future work in humans should focus on how different factors, including dose-dependency, length of exercise intervention, and age across the adult lifespan may affect brain structure and cognitive performance, both transiently and longer-term. While there is evidence in rodents that high intensity aerobic exercise, but not resistance training, promotes AHN, this effect may also be modulated by a genetic predisposition to respond to exercise (Nokia et al., 2016). Currently, it is unclear how this translates to humans. Our work shows

that some individuals, particularly those who are initially lower-fit, may reap a greater cognitive benefit than those who have initially high CRF.

### 4.4 | Limitations and considerations

It is important to discuss the limitations of the current study. Our subject pool had more women than men (81.5% female), and previous work suggests that women may have more susceptibility to the effects of CRF on cognition (Baker & Frank, 2010; Baym et al., 2014). We note that our larger female sample may be a factor in our results, although we found no sex differences across any of our analyses. To partially account for this imbalance, we included sex as a covariate in all statistical analyses. Additionally, our participant pool had moderate variability in terms of baseline fitness levels (see Figure 2A), despite enrollment criteria which required subjects to be sedentary based on self-report as defined by the American College of Sports Medicine. However, it is of note that in our sample, baseline levels of CRF did not predict behavioral accuracy on the task at any level of similarity. We also note the use of submaximal exercise testing rather than a maximal exercise test as a limitation. Submaximal testing relies on estimation of $\dot{V}O_2^{MAX}$ and of the two components used for that estimation ($\dot{V}O_2$ for speed and grade, and predicted $HR_{MAX}$ based on age) and likely introduces some variability in our CRF measurement. However, we chose the method of a submaximal exercise test to be consistent with safety measures and protocols enacted for other participant cohorts (older adults ages 55–85; data not shown) in the clinical trial. Additionally, submaximal exercise tests are a reliable, established way to estimate CRF without the need for high-intensity maximal testing that may be inappropriate for certain populations (ACSM, 2014). It is also important to note that differences in daily physical activity beyond structured exercise is an important component in overall health, and has been linked to hippocampal volume in older adults (Varma et al., 2015). Therefore, we acknowledge the lack of objective monitoring of daily physical activity of our participants during the intervention as a limitation of the current study. Future exercise intervention studies should aim to concurrently collect objective measures of physical activity beyond self-report, such as tracking movement using an accelerometer, to control for daily activity outside experimental exercise sessions.

In the current study, we employed structural neuroimaging approaches allowing us to examine hippocampal subregional volumes. A limitation of this approach, as noted earlier, is that this approach does not allow us to track AHN *in vivo* in humans. Although several studies have relied on volumetry as a proxy for plasticity (Erickson et al., 2011; Honea et al., 2009; Maass et al., 2015; Pereira et al., 2007; Smith et al., 2014; Thomas et al., 2016), potentially related to upregulated AHN (Pereira et al., 2007), changes in gray matter volume may also include other molecular and cellular phenomena including synaptogenesis, dendritic branching or pruning, spine formation, angiogenesis, and gliogenesis (Tardif et al., 2016). Therefore, we must be cautious about our interpretations of this method as it offers a macroscopic lens to understanding brain-body-behavior interactions that are complementary to more invasive methods used in animal studies that provide more detailed analysis of the underlying cellular and molecular mechanisms. In addition to volumetry, future studies in humans should also employ additional multimodal imaging techniques to study the

relationship between CRF, neuroplasticity, and cognition in an effort to provide a link between macrolevel methods and more biologically specific methods (Tardif et al., 2016). In addition, future studies in rodents should seek to characterize exercise-induced effects across the septotemporal axis of the hippocampus and determine the differential involvement of newborn neurons across this axis in cognitive tasks, including those requiring pattern separation (Tanti & Belzung, 2013).

We did not have specific predictions about lateralization of volumetric findings. Our results demonstrated a positive relationship between change in CRF and change in volume in the head of the DG/CA3 within the left hemisphere. This left lateralization is consistent with several previous studies that have also found heightened exercise-related plasticity effects in the left hippocampus (Erickson et al., 2009; Rosano et al., 2017; Wagner et al., 2017), although none of these previous studies explicitly tested for laterality of the exercise/CRF effect in the hippocampus. It is possible that a longer exercise intervention is necessary for detection of bilateral effects on hippocampal volume. For example, exercise interventions over the course of one year have demonstrated bilateral attenuation of anterior volume loss within the hippocampus in older adults (Erickson et al., 2011). However, another research group demonstrated that engaging in physical activity over the course of two years was positively related to left hippocampal volume only (Rosano et al., 2017). Results from studies of aging individuals have suggested the possibility of more severe age-related hemispheric atrophy specific to the left hemisphere compared to the right (Taki et al., 2011). Rodent studies do not typically examine laterality, despite the likelihood that corresponding regions across the two hemispheres do not necessarily functionally mirror one another and possibly provide complementary roles in processing related tasks (Lister et al., 2006). Future experiments in humans and rodents should investigate laterality effects of exercise-induced changes in regional brain structure and function.

While the current experiment focused on the effects of exercise on the hippocampal formation, it is important to note that the effects of CRF on the brain extend beyond the hippocampus. We have previously demonstrated a relationship between higher CRF and volume in the entorhinal cortex (Whiteman et al., 2016). Hayes and colleagues recently reported a relationship between higher CRF and greater fMRI activation in bilateral prefrontal cortex, medial frontal cortex, bilateral thalamus and left hippocampus during an associative encoding task (Hayes et al., 2017). They additionally showed that task activation in higher-fit older adults more similarly reflected activation in younger adults than lower-fit older adults. Higher CRF, both in cross-sectional and interventional studies, has been linked to greater activation (Herting & Nagel, 2013) and connectivity within large-scale networks such as the default mode network and frontal executive network (Voss, Erickson, et al., 2010; Voss, Prakash, et al., 2010). Lastly, higher CRF in young adults has been associated with greater resting state functional connectivity specifically between the prefrontal cortex and the anterior, but not posterior hippocampus (Stillman et al., 2018). Future studies are needed to investigate differential effects of CRF and improving CRF on connectivity between hippocampal subregions and other regions of the brain.

Despite these limitations, the research described in our manuscript makes a clear, novel contribution to the literature in several important ways. First, it extends previous exercise

intervention studies (e.g., Erickson et al., 2011; Jonasson et al., 2016; Kleemeyer et al., 2016, among others) not only to cognitively healthy young adults, but also to the hippocampal subfield level, and second, in this study, we uniquely combined an assessment of behavioral pattern separation with volumetric measurement of hippocampal subfields, extending previous work on exercise effects on behavioral pattern separation in young adults (Déry et al., 2013) and, critically, providing the most direct comparison with prominent animal models of exercise-induced hippocampal plasticity mechanisms to date. A 2019 review by Voss et al. has highlighted a clear need for research on the effects of aerobic exercise and other exercise modalities on behavioral pattern separation in humans (Voss et al., 2019). The results presented here make an important step toward filling this critical knowledge gap. Moreover, our results highlight several new avenues of future research in humans: 1) exercise training-induced functional changes during behavioral pattern separation in hippocampal subfields using high-resolution fMRI in human subjects, and 2) exercise-induced structural changes in hippocampal subfields in the aging human brain. Furthermore, this work suggests that change in CRF may be a proxy for adult hippocampal neurogenesis in humans, an idea that is consistent with previous work (Pereira et al., 2007) and that is testable by combining human subjects research with animal models that include an assessment of CRF.

### 4.5 | Conclusions

The current study provides novel evidence linking aerobic exercise with plasticity in the anterior DG/CA3 subregion and hippocampal-dependent memory. Specifically, the present study expands upon previous literature across rodents, humans, and computational models of newborn neurons and demonstrates that improving CRF with exercise training specifically enhances volume in the anterior portion of the DG/CA3, the neurogenic zone of the hippocampus. We also demonstrate a positive relationship between change in CRF and corrected accuracy for trials requiring the highest level of discrimination on a putative behavioral pattern separation task. This relationship was observed in individuals who were initially lower-fit, suggesting that individuals who show greater improvement in their CRF may receive greater cognitive benefit. Future longitudinal studies in humans are needed to focus on determining how different exercise parameters, including dose-dependency, length and type of exercise intervention, affect brain structure, function, and connectivity, as well as how they may translate to changes in cognitive performance across the adult lifespan.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

# References

ACSM. (2014). Guidelines for Exercise Testing and Prescription. (Pescatello LS, Arena R, Riebe D, & Thompson P, Eds.) (9th ed.). Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health.

Aimone JB, Deng W, & Gage FH (2011). Resolving New Memories: A Critical Look at the Dentate Gyrus, Adult Neurogenesis, and Pattern Separation. Neuron, 70(4), 589–596. 10.1016/j.neuron. 2011.05.010 [PubMed: 21609818]

Aimone JB, Wiles J, & Gage FH (2009). Computational Influence of Adult Neurogenesis on Memory Encoding. Neuron, 61(2), 187–202. 10.1016/j.neuron.2008.11.026 [PubMed: 19186162]

Altman J, & Das GD (1965). Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. The Journal of Comparative Neurology, 124(3), 319–335. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5861717 [PubMed: 5861717]

Amrein I, Nosswitz M, Slomianka L, van Dijk RM, Engler S, Klaus F, … Azim K (2015). Septo-temporal distribution and lineage progression of hippocampal neurogenesis in a primate (Callithrix jacchus) in comparison to mice. Frontiers in Neuroanatomy, 9(June), 1–15. 10.3389/fnana. 2015.00085 [PubMed: 25657619]

Baechle T, Earle RW, & Wathen D (2000). Resistance Training In Bechle T & Earle RW (Eds.), National Strength and Conditioning Association Essentials of Strength Training and Conditioning (2nd ed., pp. 408–409). Champaign, IL: Human Kinetics.

Baker LDL, & Frank L (2010). Effects of Aerobic Exercise on Mild Cognitive Impairment: a controlled trial. Archives of Neurology, 67(1), 71–79. 10.1001/archneurol.2009.307.Effects [PubMed: 20065132]

Barnes DE, Yaffe K, Satariano W. a, & Tager IB (2003). A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. Journal of the American Geriatrics Society, 51(4), 459–465. 10.1046/j.1532-5415.2003.51153.x [PubMed: 12657064]

Baym CL, Khan NA, Pence A, Raine LB, Hillman CH, & Cohen NJ (2014). Aerobic fitness predicts relational memory but not item memory performance in healthy young adults. Journal of Cognitive Neuroscience, 26(11), 2645–2652. 10.1162/jocn_a_00667 [PubMed: 24893739]

Boldrini M, Fulmore CA, Tartt AN, Simeon LR, Pavlova I, Poposka V, … Mann JJ (2018). Human Hippocampal Neurogenesis Persists throughout Aging. Cell Stem Cell, 22(4), 589–599. 10.1016/j.stem.2018.03.015 [PubMed: 29625071]

Boldrini M, Hen R, Underwood MD, Rosoklija GB, Dwork AJ, Mann JJ, & Arango V (2012). Hippocampal angiogenesis and progenitor cell proliferation are increased with antidepressant use in major depression. Biological Psychiatry, 72(7), 562–571. 10.1016/j.biopsych.2012.04.024 [PubMed: 22652019]

Borg G (1982). Ratings of perceived exertion and heart rates during short-term cycle exercise and their use in a new cycling strength test. International Journal of Sports Medicine, 3(3), 153–158. 10.1055/s-2008-1026080 [PubMed: 7129724]

Burzynska AZ, Chaddock-Heyman L, Voss MW, Wong CN, Gothe NP, Olson EA, … Kramer AF (2014). Physical activity and cardiorespiratory fitness are beneficial for white matter in low-fit older adults. PLoS ONE, 9(9). 10.1371/journal.pone.0107413

Chaddock L, Erickson KI, Shaurya R, Kim JS, Voss MW, Vanpatter M, … Kramer AF (2010). A neuroimaging investigation of the association between aerobic fitness, hippocampal volume, and memory performance in preadolescent children. Brain Research, 1358, 172–183. 10.1016/j.brainres.2010.08.049 [PubMed: 20735996]

Christie BR, Eadie BD, Kannangara TS, Robillard JM, Shin J, & Titterness AK (2008). Exercising our brains: How physical activity impacts synaptic plasticity in the dentate gyrus. NeuroMolecular Medicine, 10(2), 47–58. 10.1007/s12017-008-8033-2 [PubMed: 18535925]

Colcombe S, & Kramer AF (2003). Fitness effects on the Cognitive Function of Older Adults: A Meta-Analytic Study. Psychological Science, 14(2), 125–130. 10.1111/1467-9280.t01-1-01430 [PubMed: 12661673]

Cooper CB, & Storer TW (2001). Exercise Testing and Interpretation. A Practical Approach. New York, NY, NY: Cambridge University Press.

Creer DJ, Romberg C, Saksida LM, van Praag H, & Bussey TJ (2010). Running enhances spatial pattern separation in mice. Proceedings of the National Academy of Sciences of the United States of America, 107(5), 2367–2372. 10.1073/pnas.0911725107 [PubMed: 20133882]

Daugherty AM, Yu Q, Flinn R, & Ofen N (2015). A reliable and valid method for manual demarcation of hippocampal head, body, and tail. International Journal of Developmental Neuroscience, 41, 115–122. 10.1016/j.ijdevneu.2015.02.001 [PubMed: 25660945]

Déry N, Pilgrim M, Gibala M, Gillen J, Wojtowicz JM, Macqueen G, & Becker S (2013). Adult hippocampal neurogenesis reduces memory interference in humans: opposing effects of aerobic exercise and depression. Frontiers in Neuroscience, 7(7 APR), 66 10.3389/fnins.2013.00066 [PubMed: 23641193]

Duvernoy HM (2005). The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI (3rd ed.). Springer.

Eadie BD, Redila VA, & Christie BR (2005). Voluntary exercise alters the cytoarchitecture of the adult dentate gyrus by increasing cellular proliferation, dendritic complexity, and spine density. Journal of Comparative Neurology, 486(1), 39–47. 10.1002/cne.20493 [PubMed: 15834963]

Erickson KI, Prakash RS, Voss MW, Chaddock L, Hu L, Morris KS, … Kramer AF (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. Hippocampus, 19(10), 1030–1039. 10.1002/hipo.20547 [PubMed: 19123237]

Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, … Kramer AF (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. 10.1073/pnas.1015950108 [PubMed: 21282661]

Eriksson PS, Perfilieva E, Björk-Eriksson T, Alborn AM, Nordborg C, Peterson DA, & Gage FH (1998). Neurogenesis in the adult human hippocampus. Nature Medicine, 4(11), 1313–1317. 10.1038/3305

Eritaia J, Wood SJ, Stuart GW, Bridle N, Dudgeon P, Maruff P, … Pantelis C (2000). An optimized method for estimating intracranial volume from magnetic resonance images. Magnetic Resonance in Medicine, 44(6), 973–977. 10.1002/1522-2594(200012)44:6&lt;973::AID-MRM21&gt;3.0.CO;2-H [PubMed: 11108637]

Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I-M, … American College of Sports Medicine. (2011). American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Medicine and Science in Sports and Exercise, 43(7), 1334–1359. 10.1249/MSS.0b013e318213fefb [PubMed: 21694556]

Gomez-Pinilla F, & Hillman C (2013). The influence of exercise on cognitive abilities. Comprehensive Physiology, 3(1), 403–428. 10.1002/cphy.c110063 [PubMed: 23720292]

Gould E, Reeves AJ, Fallah M, Tanapat P, Gross CG, & Fuchs E (1999). Hippocampal neurogenesis in adult Old World primates. Proceedings of the National Academy of Sciences, 96(9), 5263–5267. 10.1073/pnas.96.9.5263

Griffin ÉW, Mullally S, Foley C, Warmington SA, O'Mara SM, & Kelly ÁM (2011). Aerobic exercise improves hippocampal function and increases BDNF in the serum of young adult males. Physiology and Behavior, 104(5), 934–941. 10.1016/j.physbeh.2011.06.005 [PubMed: 21722657]

Hagberg JM (1994). Exercise assessment of arthritic and elderly individuals. Bailliere's Clinical Rheumatology, 8(1), 29–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8149448 [PubMed: 8149448]

Hasselmo ME, & Wyble BP (1997). Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. Behavioural Brain Research, 89(1–2), 1–34. 10.1016/S0166-4328(97)00048-X [PubMed: 9475612]

Hayes SM, Hayes JP, Williams VJ, Liu H, & Verfaellie M (2017). FMRI activity during associative encoding is correlated with cardiorespiratory fitness and source memory performance in older adults. Cortex, 91(10), 208–220. 10.1016/j.cortex.2017.01.002 [PubMed: 28161031]

Head D, Rodrigue KM, Kennedy KM, & Raz N (2008). Neuroanatomical and cognitive mediators of age-related differences in episodic memory. Neuropsychology, 22(4), 491–507. 10.1037/0894-4105.22.4.491 [PubMed: 18590361]

Herting MM, & Nagel BJ (2013). Differences in brain activity during a verbal associative memory encoding task in high- and low-fit adolescents. Journal of Cognitive Neuroscience, 25(4), 595–612. 10.1162/jocn_a_00344 [PubMed: 23249350]

Honea RA, Thomas GP, Harsha A, Anderson HS, Donnelly JE, Brooks WM, & Burns JM (2009). Cardiorespiratory fitness and preserved medial temporal lobe volume in Alzheimer disease. Alzheimer Disease and Associated Disorders, 23(3), 188–197. 10.1097/WAD.0b013e31819cb8a2 [PubMed: 19812458]

Jonasson LS, Nyberg L, Kramer AF, Lundquist A, Riklund K, & Boraxbekk C-J (2016). Aerobic Exercise Intervention, Cognitive Performance, and Brain Structure: Results from the Physical Influences on Brain in Aging (PHIBRA) Study. Frontiers in Aging Neuroscience, 8(3), 336 10.3389/fnagi.2016.00336 [PubMed: 28149277]

Kee N, Teixeira CM, Wang AH, & Frankland PW (2007). Preferential incorporation of adult-generated granule cells into spatial memory networks in the dentate gyrus. Nature Neuroscience, 10(3), 355–362. 10.1038/nn1847 [PubMed: 17277773]

Kinser PA, & Robins JL (2013). Control group design: enhancing rigor in research of mind-body therapies for depression. Evidence-Based Complementary and Alternative Medicine: ECAM, 2013, 140467 10.1155/2013/140467 [PubMed: 23662111]

Kleemeyer MM, Kühn S, Prindle J, Bodammer NC, Brechtel L, Garthe A, … Lindenberger U (2016). Changes in fitness are associated with changes in hippocampal microstructure and hippocampal volume among older adults. NeuroImage, 131, 155–161. 10.1016/j.neuroimage.2015.11.026 [PubMed: 26584869]

Kornack DR, & Rakic P (1999). Continuation of neurogenesis in the hippocampus of the adult macaque monkey. Proceedings of the National Academy of Sciences of the United States of America, 96(10), 5768–5773. 10.1073/PNAS.96.10.5768 [PubMed: 10318959]

Lindquist R, Wyman JF, Talley KMC, Findorff MJ, & Gross CR (2007). Design of control-group conditions in clinical trials of behavioral interventions. Journal of Nursing Scholarship: An Official Publication of Sigma Theta Tau International Honor Society of Nursing, 39(3), 214–221. 10.1111/j.1547-5069.2007.00171.x

Lister JP, Tonkiss J, Blatt GJ, Kemper TL, DeBassio WA, Galler JR, & Rosene DL (2006). Asymmetry of neuron numbers in the hippocampal formation of prenatally malnourished and normally nourished rats: a stereological investigation. Hippocampus, 16(11), 946–958. 10.1002/hipo.20221 [PubMed: 16983649]

Logan P, Fornasiero D, Abernethy P, & Lynch K (2000). Protocols for the assessment of isoinertial strength In Gore CJ (Ed.), Physiological Tests for Elite Athletes. Champaign, IL: Human Kinetics.

Lundby C, Montero D, & Joyner M (2017). Biology of VO2max: looking under the physiology lamp. Acta Physiologica, 220(2), 218–228. 10.1111/apha.12827 [PubMed: 27888580]

Maass A, Düzel S, Goerke M, Becke A, Sobieray U, Neumann K, … Düzel E (2015). Vascular hippocampal plasticity after aerobic exercise in older adults. Molecular Psychiatry, 20(5), 585–593. 10.1038/mp.2014.114 [PubMed: 25311366]

Nauer RK, Whiteman AS, Dunne MF, Stern CE, & Schon K (2015). Hippocampal subfield and medial temporal cortical persistent activity during working memory reflects ongoing encoding. Frontiers in Systems Neuroscience, 9(March), 30 10.3389/fnsys.2015.00030 [PubMed: 25859188]

Neeper SA, Góauctemez-Pinilla F, Choi J, & Cotman C (1995). Exercise and brain neurotrophins. Nature, 373(6510), 109–109. 10.1038/373109a0 [PubMed: 7816089]

Nokia MS, Lensu S, Ahtiainen JP, Johansson PP, Koch LG, Britton SL, & Kainulainen H (2016). Physical exercise increases adult hippocampal neurogenesis in male rats provided it is aerobic and sustained. The Journal of Physiology, 00, 1–19. 10.1113/JP271552

O'Reilly RC, & McClelland JL (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. Hippocampus, 4(6), 661–682. 10.1002/hipo.450040605 [PubMed: 7704110]

Palmer TD, Willhoite AR, & Gage FH (2000). Vascular niche for adult hippocampal neurogenesis. The Journal of Comparative Neurology, 425(4), 479–494. 10.1002/1096-9861(20001002)425:4&lt;479::AID-CNE2&gt;3.0.CO;2-3 [PubMed: 10975875]

Park H, & Poo M (2013). Neurotrophin regulation of neural circuit development and function. Nature Reviews, Neuroscience, 14(1), 7–23. 10.1038/nrn3379 [PubMed: 23254191]

Pereira AC, Huddleston DE, Brickman AM, Sosunov AA, Hen R, McKhann GM, … Small SA (2007). An in vivo correlate of exercise-induced neurogenesis in the adult dentate gyrus. Proceedings of the National Academy of Sciences of the United States of America, 104(13), 5638–5643. 10.1073/pnas.0611721104 [PubMed: 17374720]

Piatti VC, Davies-Sala MG, Espósito MS, Mongiat LA, Trinchero MF, & Schinder AF (2011). The timing for neuronal maturation in the adult hippocampus is modulated by local network activity. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 31(21), 7715–7728. 10.1523/JNEUROSCI.1380-11.2011 [PubMed: 21613484]

Poppenk J, Evensmoen HR, Moscovitch M, & Nadel L (2013). Long-axis specialization of the human hippocampus. Trends in Cognitive Sciences, 17(5), 230–240. 10.1016/j.tics.2013.03.005 [PubMed: 23597720]

Pruessner JC, Köhler S, Crane J, Lord C, Byrne A, Kabani N, … Evans AC (2002). Entorhinal and Parahippocampal Cortex from High-resolution MR Images: Considering the Variability of the Collateral Sulcus, 1342–1353.

Pruessner JC, Li LM, Serles W, Pruessner M, Collins DL, Kabani N, … Evans AC (2000). Volumetry of Hippocampus and Amygdala with High-resolution MRI and Three-dimensional Analysis Software: Minimizing the Discrepancies between Laboratories, 433–442.

Ramirez-Amaya V, Marrone DF, Gage FH, Worley PF, & Barnes CA (2006). Integration of new neurons into functional neural networks. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 26(47), 12237–12241. 10.1523/JNEUROSCI.2195-06.2006 [PubMed: 17122048]

Raz N, Rodrigue KM, Head D, Kennedy KM, & Acker JD (2004). Differential aging of the medial temporal lobe: a study of a five-year change. Neurology, 62(3), 433–438. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14872026 [PubMed: 14872026]

Rosano C, Guralnik J, Pahor M, Glynn NW, Newman AB, Ibrahim TS, … Aizenstein HJ (2017). Hippocampal Response to a 24-Month Physical Activity Intervention in Sedentary Older Adults. The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 25(3), 209–217. 10.1016/j.jagp.2016.11.007 [PubMed: 27986412]

Sahay A, Scobie KN, Hill AS, O'Carroll CM, Kheirbek MA, Burghardt NS, … Hen R (2011). Increasing adult hippocampal neurogenesis is sufficient to improve pattern separation. Nature, 472(7344), 466–470. 10.1038/nature09817 [PubMed: 21460835]

Schwarb H, Johnson CL, Daugherty AM, Hillman CH, Kramer AF, Cohen NJ, & Barbey AK (2017). Aerobic fitness, hippocampal viscoelasticity, and relational memory performance. NeuroImage, 153(3), 179–188. 10.1016/j.neuroimage.2017.03.061 [PubMed: 28366763]

Smith JC, Nielson KA, Woodard JL, Seidenberg M, Durgerian S, Hazlett KE, … Rao SM (2014). Physical activity reduces hippocampal atrophy in elders at genetic risk for Alzheimer's disease. Frontiers in Aging Neuroscience, 6(APR), 1–7. 10.3389/fnagi.2014.00061 [PubMed: 24478697]

Snyder JS, Radik R, Wojtowicz JM, & Cameron HA (2009). Anatomical gradients of adult neurogenesis and activity: young neurons in the ventral dentate gyrus are activated by water maze training. Hippocampus, 19(4), 360–370. 10.1002/hipo.20525 [PubMed: 19004012]

Sorrells SF, Paredes MF, Cebrian-Silla A, Sandoval K, Qi D, Kelley KW, … Alvarez-Buylla A (2018). Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. Nature, 555(7696), 377–381. 10.1038/nature25975 [PubMed: 29513649]

Stillman CM, Uyar F, Huang H, Grove GA, Watt JC, Wollam ME, & Erickson KI (2018). Cardiorespiratory fitness is associated with enhanced hippocampal functional connectivity in healthy young adults. Hippocampus, 28(3), 239–247. 10.1002/hipo.22827 [PubMed: 29341303]

Stranahan AM, Khalil D, & Gould E (2007). Running induces widespread structural alterations in the hippocampus and entorhinal cortex. Hippocampus, 17(11), 1017–1022. 10.1002/hipo.20348 [PubMed: 17636549]

Strange BA, Witter MP, Lein ES, & Moser EI (2014). Functional organization of the hippocampal longitudinal axis. Nature Reviews. Neuroscience, 15(10), 655–669. 10.1038/nrn3785 [PubMed: 25234264]

Strauss E, Sherman EMS, & Spreen O (2006). A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary (3rd ed.). New York, NY, NY: Oxford University Press.

Suwabe K, Byun K, Hyodo K, Reagh ZM, Roberts JM, Matsushita A, … Soya H (2018). Rapid stimulation of human dentate gyrus function with acute mild exercise. Proceedings of the National Academy of Sciences of the United States of America, 1–6. 10.1073/pnas.1805668115

Suwabe K, Hyodo K, Byun K, Ochi G, Fukuie T, Shimizu T, … Soya H (2017). Aerobic fitness associates with mnemonic discrimination as a mediator of physical activity effects: evidence for memory flexibility in young adults. Scientific Reports, 7(1), 5140 10.1038/s41598-017-04850-y [PubMed: 28698596]

Suwabe K, Hyodo K, Byun K, Ochi G, Yassa MA, & Soya H (2017). Acute moderate exercise improves mnemonic discrimination in young adults. Hippocampus, 27(3), 229–234. 10.1002/hipo. 22695 [PubMed: 27997992]

Taki Y, Thyreau B, Kinomura S, Sato K, Goto R, Kawashima R, & Fukuda H (2011). Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals. PloS One, 6(7), e22734 10.1371/journal.pone.0022734 [PubMed: 21818377]

Tanaka H, Monahan KD, & Seals DR (2001). Age-predicted maximal heart rate revisited. Journal of the American College of Cardiology, 37(1), 153–156. 10.1016/S0735-1097(00)01054-8 [PubMed: 11153730]

Tanti A, & Belzung C (2013). Neurogenesis along the septo-temporal axis of the hippocampus: Are depression and the action of antidepressants region-specific? Neuroscience, 252, 234–252. 10.1016/j.neuroscience.2013.08.017 [PubMed: 23973415]

Tanti A, Westphal WP, Girault V, Brizard B, Devers S, Leguisquet AM, … Belzung C (2013). Region-dependent and stage-specific effects of stress, environmental enrichment, and antidepressant treatment on hippocampal neurogenesis. Hippocampus, 23(9), 797–811. 10.1002/hipo.22134 [PubMed: 23592526]

Tardif CL, Gauthier CJ, Steele CJ, Bazin P-L, Schäfer A, Schaefer A, … Villringer A (2016). Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. NeuroImage, 131, 55–72. 10.1016/j.neuroimage.2015.08.047 [PubMed: 26318050]

Thomas AG, Dennis A, Rawlings NB, Stagg CJ, Matthews L, Morris M, … Johansen-Berg H (2016). Multi-modal characterization of rapid anterior hippocampal volume increase associated with aerobic exercise. NeuroImage, 131, 162–170. 10.1016/j.neuroimage.2015.10.090 [PubMed: 26654786]

Thompson CL, Pathak SD, Jeromin A, Ng LL, MacPherson CR, Mortrud MT, … Lein ES (2008). Genomic Anatomy of the Hippocampus. Neuron, 60(6), 1010–1021. 10.1016/j.neuron. 2008.12.008 [PubMed: 19109908]

van Praag H, Kempermann G, & Gage FH (1999). Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. Nature Neuroscience, 2(3), 266–270. 10.1038/6368 [PubMed: 10195220]

van Praag H, Shubert T, Zhao C, & Gage FH (2005). Exercise enhances learning and hippocampal neurogenesis in aged mice. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 25(38), 8680–8685. 10.1523/JNEUROSCI.1731-05.2005 [PubMed: 16177036]

Varma VR, Chuang YF, Harris GC, Tan EJ, & Carlson MC (2015). Low-intensity daily walking activity is associated with hippocampal volume in older adults. Hippocampus, 25(5), 605–615. 10.1002/hipo.22397 [PubMed: 25483019]

Varma VR, Tang X, & Carlson MC (2016). Hippocampal sub-regional shape and physical activity in older adults. Hippocampus, 26(8), 1051–1060. 10.1002/hipo.22586 [PubMed: 27009597]

Vivar C, Peterson BD, & van Praag H (2016). Running rewires the neuronal network of adult-born dentate granule cells. NeuroImage, 131, 29–41. 10.1016/j.neuroimage.2015.11.031 [PubMed: 26589333]

Voss MW, Erickson KI, Prakash RS, Chaddock L, Malkowski E, Alves H, … Kramer AF (2010). Functional connectivity: a source of variance in the association between cardiorespiratory fitness

and cognition? Neuropsychologia, 48(5), 1394–1406. 10.1016/j.neuropsychologia.2010.01.005 [PubMed: 20079755]

Voss MW, Nagamatsu LS, Liu-Ambrose T, & Kramer AF (2011). Exercise, brain, and cognition across the life span. Journal of Applied Physiology (Bethesda, Md.: 1985), 111(5), 1505–1513. 10.1152/japplphysiol.00210.2011

Voss MW, Prakash RS, Erickson KI, Basak C, Chaddock L, Kim JS, … Kramer AF (2010). Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. Frontiers in Aging Neuroscience, 2(8), 1–17. 10.3389/fnagi.2010.00032 [PubMed: 20552041]

Voss MW, Soto C, Yoo S, Sodoma M, Vivar C, & van Praag H (2019). Exercise and Hippocampal Memory Systems. Trends in Cognitive Sciences. 10.1016/j.tics.2019.01.006

Wagner G, Herbsleb M, de la Cruz F, Schumann A, Brünner F, Schachtzabel C, … Bär K-J (2015). Hippocampal structure, metabolism, and inflammatory response after a 6-week intense aerobic exercise in healthy young adults: a controlled trial. Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism, 35(10), 1570–1578. 10.1038/jcbfm.2015.125

Wagner G, Herbsleb M, de la Cruz F, Schumann A, Köhler S, Puta C, … Bär KJ (2017). Changes in fMRI activation in anterior hippocampus and motor cortex during memory retrieval after an intense exercise intervention. Biological Psychology, 124, 65–78. 10.1016/j.biopsycho. 2017.01.003 [PubMed: 28119067]

Wasserman K, Hansen J, Sue DY, Stringer W, Sietsema K, Sun X, & Whipp B (2012). Exercise Physiology. In: Principles of Exercise Testing and Interpretation (5th ed.). Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health.

Whiteman AS, Young DE, Budson AE, Stern CE, & Schon K (2016). Entorhinal volume, aerobic fitness, and recognition memory in healthy young adults: A voxel-based morphometry study. NeuroImage, 126, 229–238. 10.1016/j.neuroimage.2015.11.049 [PubMed: 26631814]

Yamamoto S, Hotta K, Ota E, Mori R, & Matsunaga A (2016). Effects of resistance training on muscle strength, exercise capacity, and mobility in middle-aged and elderly patients with coronary artery disease: A meta-analysis. Journal of Cardiology, 68(2), 125–134. 10.1016/j.jjcc.2015.09.005 [PubMed: 26690738]

Yushkevich P. a, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, & Gerig G (2006). User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. NeuroImage, 31(3), 1116–1128. 10.1016/j.neuroimage.2006.01.015 [PubMed: 16545965]
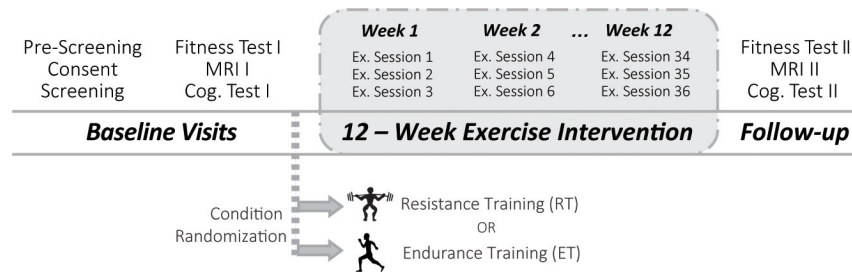
**Figure 1. Diagram for the study design and randomization for both exercise groups.**
Resistance Training (RT), N = 21; Endurance Training (ET), N = 17.

**A**

Estimated VO$_{2MAX}$



**B**



**C**
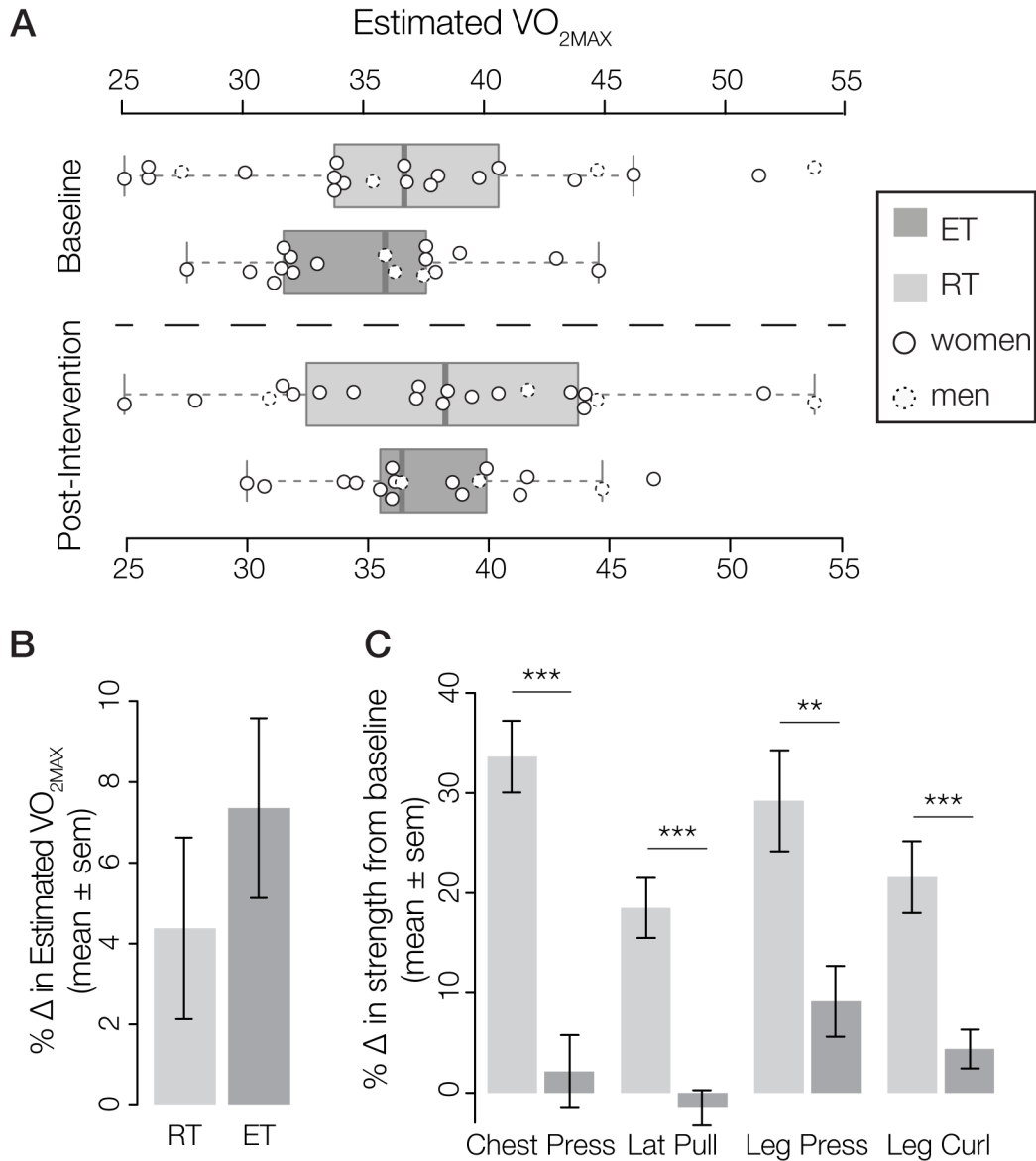


**Figure 2. Fitness Results.**

(A) Boxplots showing estimated $\dot{V}O_2^{MAX}$, the measure of cardiorespiratory fitness (CRF) for each exercise group (ET and RT) at baseline fitness test (top) and post-intervention testing (bottom). Individual data points are overlain (closed circle – women; dotted circle – men). (B) Change in CRF for the RT and ET exercise groups. Within-group t-tests show that only the ET group had a significant increase in CRF; group comparison in repeated measures-ANOVA do not show significantly different changes at the group level. (C) Increases in strength from baseline, measured by the percent change from baseline (lbs) for each exercise across the groups. The RT group showed significantly greater improvements compared to the ET group for each exercise (***p < 0.0001; ** p < 0.001). *N (ET) = 17; N (RT) = 21.*
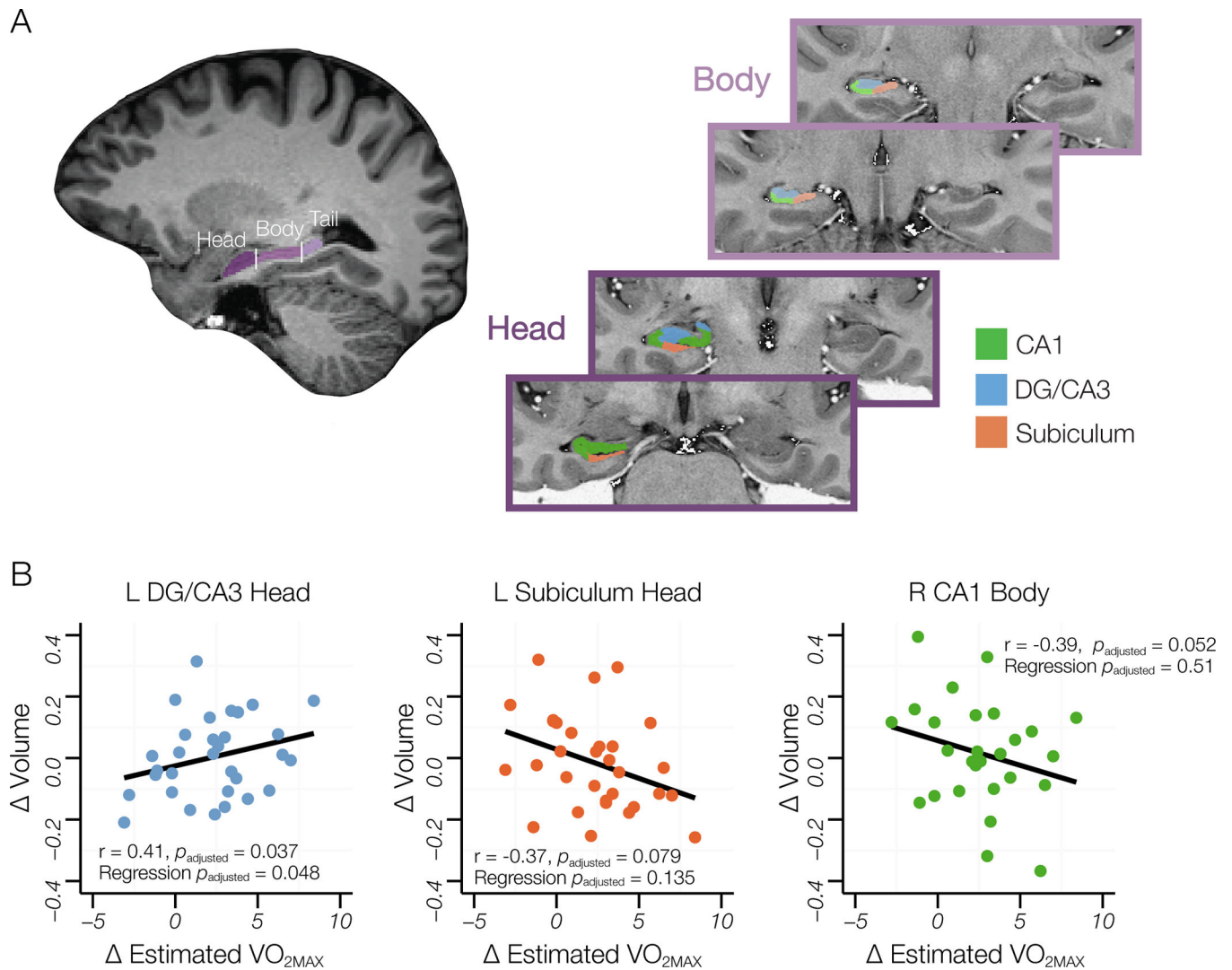
**Figure 3. Delineations of Regions of Interest (ROI) and relationships between changes in CRF and hippocampal subfield volume.**

(A) ROIs for hippocampal subfield segmentation were manually traced and included CA1, DG/CA3 and subiculum in the head and body. Head, body, and tail separations were decided by anatomical landmarks (see text). Hippocampal tail was not segmented into subfields. (B) Scatterplots of the association between change in CRF and subfield volumes. There was a significant positive relationship between change in CRF and change in adjusted volume in the head of the DG/CA3 (B, left), and a negative relationship between change in CRF and adjusted volume in left head of the subiculum (B, middle) and adjusted volume in the body of right CA1. Subsequent multiple regression models including covariates mitigated the relationships between change in CRF and volume in the left head of the subiculum and right body of CA1, but the positive relationship between change in CRF and volume in the left head of DG/CA3 remained (B, left).
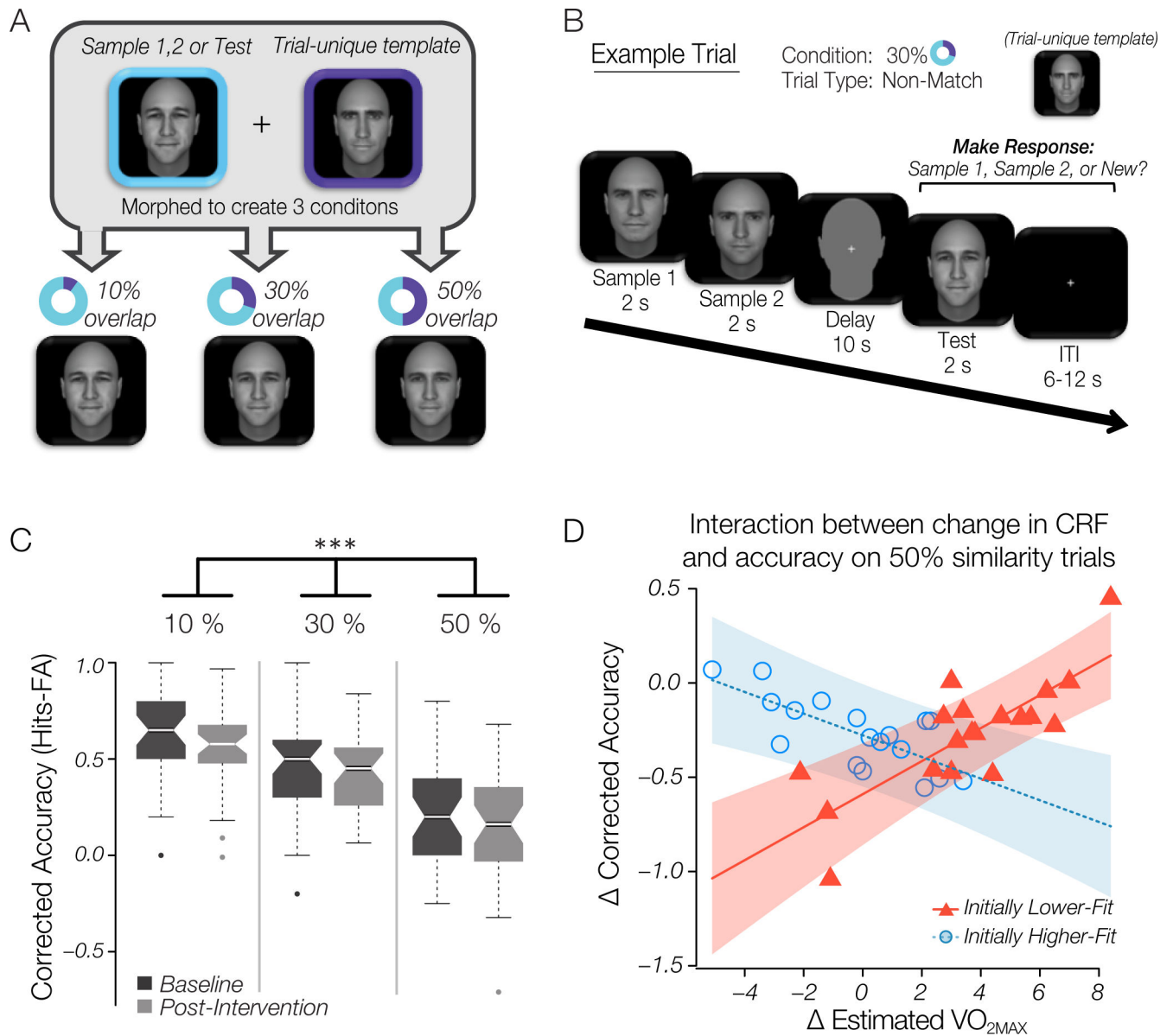
**Figure 4. Mnemonic discrimination task and behavioral results.**
(A) Stimulus similarity was objectively created by morphing all stimuli within a trial to a trial-unique template at low (10%), moderate (30%) or high (50%) similarity. (B) An example trial of the delayed-matching-to sample task. Half the trials were non-match (as in the example shown) and half were match trials whereby the Test face stimulus was the same as Sample 1 or Sample 2. Trials were evenly split across the three similarity levels. Inter-trial interval (ITI) varied between 6–12 s. (C) Boxplots displaying corrected accuracy collapsed across all subjects at baseline and follow-up by similarity condition. Results indicate a main effect of stimulus similarity (*** $p < .001$). (D) Scatterplot showing the relationship between change in fitness percentile and change in corrected accuracy across time. Data is separated by a median split into higher-fit (blue circles) and lower-fit (red triangles) individuals based on initial fitness level at baseline. Regression results show that initially lower-fit individuals

reap a cognitive benefit from improving their CRF, as evidenced by the positive relationship between change in CRF and change in corrected accuracy for lower-fit individuals.

**Table 1.**

Participant Demographics.

| | Mean (SD) | | |
|---|---|---|---|
| **Variable** | **Endurance Group (ET)** | **Resistance Group (RT)** | **p-value** |
| N | 17 | 21 | --- |
| Age | 24.94 (2.5) | 26.33 (3.7) | 0.18 |
| Sex | 3 M, 14 F | 4 M, 17 F | --- |
| Education | 16.70 (1.0) | 17.09 (1.5) | 0.35 |
| Attendance (%) | 80 (12) % | 81 (11) % | 0.79 |

Means (s.d.) shown for the participants.

**Table 2.**

Pearson correlations between change in fitness percentile and change in ROI volume.

|      | ROI | Correlation (r) with fitness | Standardized beta (± St. Err) for fitness |
|------|-----|------------------------------|-------------------------------------------|
| **HEAD** | L. Subiculum | −0.353[*] | −0.322 (0.002) |
|      | R. Subiculum | −0.193 | −0.178 (0.003) |
|      | L. CA1 | −0.140 | −0.122 (0.002) |
|      | R. CA1 | −0.052 | −0.034 (0.001) |
|      | L. CA3/DG | 0.377[*] | 0.361 (0.002)[*] |
|      | R. CA3/DG | −0.099 | −0.103 (0.002) |
| **BODY** | L. Subiculum | 0.203 | 0.199 (0.002) |
|      | R. Subiculum | 0.046 | 0.042 (0.003) |
|      | L. CA1 | −0.076 | −0.058 (0.003) |
|      | R. CA1 | −0.321 | −0.301 (0.002) |
|      | L. CA3/DG | −0.041 | −0.039 (0.001) |
|      | R. CA3/DG | −0.199 | −0.165 (0.001) |
| **TAIL** | L. Hippocampus | −0.089 | −0.11 (0.002) |
|      | R. Hippocampus | −0.057 | −0.093 (0.003) |

[*] $p < .05$