RESEARCH ARTICLE

WILEY

Cardiorespiratory fitness, hippocampal subfield volumes, and mnemonic discrimination task performance in aging

Kathryn L. Kern[1,2] 🆔 | Thomas W. Storer[3] | Karin Schon[1,2,4,5]

[1]Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, Massachusetts

[2]Center for Systems Neuroscience, Boston University, Boston, Massachusetts

[3]Men's Health, Aging, and Metabolism Unit, Brigham and Women's Hospital, Boston, Massachusetts

[4]Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts

[5]Center for Memory and Brain, Boston University, Boston, Massachusetts

Correspondence
Kathryn L. Kern, Department of Anatomy and Neurobiology, Boston University School of Medicine, 72 East Concord Street, L-1004, Boston, MA 02118.
Email: klkern@bu.edu

Abstract

Aging and exercise have opposing effects on mnemonic discrimination task performance, which putatively taxes pattern separation mechanisms reliant on the dentate gyrus (DG) subfield of the hippocampus. In young adults, increasing cardiorespiratory fitness (CRF) has been shown to improve mnemonic discrimination task performance and increase left anterior DG/CA3 volume. It is unknown how these variables interact in cognitive aging, yet this knowledge is critical, given the established effects of aging on hippocampal plasticity. To investigate these relationships, 65 older adults (aged 55–85 years) completed a submaximal treadmill test to estimate CRF, a mnemonic discrimination task, and a high-resolution MRI scan to determine hippocampal subfield volumes. Our older adult sample demonstrated the lowest task accuracy in the condition with the greatest stimuli similarity and left DG/CA3 body volume significantly predicted accuracy in this condition. Our results did not provide support for relationships between CRF and task accuracy or CRF and DG/CA3 volume as evidenced in studies of young adults. Instead, CRF predicted bilateral subiculum volume in older adult women, not men. Altogether, these findings provide further support for a role of the DG in behavioral pattern separation in humans and suggest that CRF may have differential effects on hippocampal subfield integrity in older adult men and women. ClinicalTrials.gov identifiers: (a) Neuroimaging Study of Exercise and Memory Function, NCT02057354; (b) The Entorhinal Cortex and Aerobic Exercise in Aging, NCT02775760; (c) Physical Activity and Cognition Study, NCT02773121.

KEYWORDS
aging, cardiorespiratory fitness, hippocampus, mnemonic discrimination, MRI, pattern separation

1 | INTRODUCTION

The hippocampus is a heterogeneous structure composed of interconnected, yet anatomically separate subfields that together play an essential role in episodic memory formation. The hippocampus demonstrates striking plasticity throughout the adult lifespan, and as such,

exhibits a structural and functional change in response to aging and lifestyle factors, such as exercise. Cognitive aging has negative effects on hippocampal integrity. In particular, aging is associated with decreased hippocampal volume (Jack et al., 1997; Raz, Rodrigue, Head, Kennedy, & Acker, 2004; Raz et al., 2005) and impaired performance on episodic memory tasks (Schaie, 2013; Small, Stern, Tang, & Mayeux, 1999). Accurate episodic memory formation relies on the neurocomputational process of pattern separation, which is the disambiguation of similar stimulus input into nonoverlapping neural representations (Yassa & Stark, 2011). This process is thought to rely on the dentate gyrus (DG), the hippocampal subfield that demonstrates neurogenesis throughout the adult lifespan (Boldrini et al., 2018; Eriksson et al., 1998; Spalding et al., 2013; for opposing perspective, see Sorrells et al., 2018). Cognitive aging is also associated with decreased accuracy on mnemonic discrimination tasks (Nauer, Schon, & Stern, 2020; Stark & Stark, 2017; Stark, Yassa, Lacy, & Stark, 2013; Yassa et al., 2011; Yassa & Stark, 2011), which are behavioral paradigms designed to tax pattern separation. Mechanistically, it is likely that this behavioral impairment is related to the vulnerability of the DG subfield in aging. Rodent, nonhuman primate, and human studies indicate that although the number of DG neurons is largely unaffected with age (Erickson, 2003), deviations in DG neuronal physiology (Gazzaley, Siegel, Kordower, Mufson, & Morrison, 1996; Penner et al., 2011; Small, Chawla, Buonocore, Rapp, & Barnes, 2004; Small, Tsai, DeLaPaz, Mayeux, & Stern, 2002) and decreases in the rate of DG neurogenesis (Gould et al., 1999; Kuhn, Dickinson-Anson, & Gage, 1996; Ngwenya, Heyworth, Shwe, Moore, & Rosene, 2015) may underlie age-related changes in hippocampal functional and structural integrity.

Whereas aging is associated with impaired hippocampal plasticity, aerobic exercise is associated with enhanced hippocampal plasticity. In both young and older adults, aerobic exercise that increases cardiorespiratory fitness has been associated with positive effects on hippocampal structure and function. Cardiorespiratory fitness is a measure representative of the body's capacity to integrate the cardiovascular, respiratory, and skeletal muscle systems to generate dynamic, moderate-to-vigorous intensity exercise for elongated periods of time (Pescatello & American College of Sports Medicine, 2014). In young adults specifically, aerobic exercise training that increases cardiorespiratory fitness is associated with increased anterior hippocampal volume (Thomas et al., 2016). This phenomenon may be driven by plasticity of the DG in particular, as neuroimaging studies with subfield-level precision indicate that increasing cardiorespiratory fitness through aerobic exercise training results in increased DG cerebral blood volume (Pereira et al., 2007) and left anterior DG/CA3 volume (Nauer, Dunne, Stern, Storer, & Schon, 2020). Behaviorally in young adults, studies suggest that increasing cardiorespiratory fitness may enhance mnemonic discrimination task performance (Déry et al., 2013; Nauer, Dunne, et al., 2020). It is of interest that these hippocampal findings are specific to the DG given that the rodent literature demonstrates similar, DG-specific exercise-induced plasticity (Clark, Brzezinska, Puchalski, Krone, & Rhodes, 2009; Eadie, Redila, & Christie, 2005; Marlatt, Potter, Lucassen, & van Praag, 2012; Nokia et al., 2016; Pereira et al., 2007;

Redila & Christie, 2006; van Praag, Christie, Sejnowski, & Gage, 1999; van Praag, Kempermann, & Gage, 1999; van Praag, Shubert, Zhao, & Gage, 2005). Behaviorally in rodents, this DG-specific plasticity results in enhanced performance on spatial pattern separation tasks (Creer, Romberg, Saksida, van Praag, & Bussey, 2010). Altogether, these studies in rodents and young adult humans indicate that exercise has profound effects on hippocampal plasticity and may have differential effects on the hippocampal subfields.

Studies in older adults also consistently demonstrate positive effects of aerobic exercise that increases cardiorespiratory fitness on hippocampal integrity. One way to explore hypotheses regarding the effects of aerobic exercise on the brain cross-sectionally is to examine associations between cardiorespiratory fitness and brain structure and function. Critically, aerobic exercise training increases cardiorespiratory fitness in older adults (Hagberg et al., 1989; Kohrt et al., 1991), and as such, cardiorespiratory fitness is thought to be a driving mechanism underlying the effects of aerobic exercise on the brain. Notably, cross-sectional studies and randomized controlled trials demonstrate relationships between cardiorespiratory fitness, aerobic exercise, and hippocampal volume. Erickson et al. (2009) demonstrated that greater cardiorespiratory fitness is associated with greater bilateral hippocampal volume in older adults. These results have since been extended to multiple randomized controlled trials, the findings of which suggest that increased cardiorespiratory fitness achieved through aerobic exercise training is associated with increased hippocampal volume (Erickson et al., 2011; Jonasson et al., 2017). The relationship between aerobic exercise training and greater hippocampal volume in older adults is thought to result from the attenuation of hippocampal atrophy associated with aging (Erickson et al., 2011; Firth et al., 2018), rather than increased volume compared to baseline measurements, as seen in young adults (Nauer, Dunne, et al., 2020; Thomas et al., 2016).

Although relationships between aerobic exercise, cardiorespiratory fitness, and whole hippocampal volume in older adults are well established, studies in this population do not converge upon a clear explanation for the differential effects of aerobic exercise or cardiorespiratory fitness on the distinct hippocampal subfields, as they do in young adults. For example, Rosano et al. (2017) demonstrated that exposure to a multi-modal exercise training program resulted in increased volume of the left Cornu Ammonis (CA1–CA3) subfields, but not the right CA subfields, nor bilateral DG. Comparatively, Varma, Tang, and Carlson (2016) showed that greater walking activity was associated with larger surface area of the bilateral subiculum in women, not men. Together, this literature suggests that aerobic exercise training and physical activity have robust effects on the hippocampus; however, a clear picture for how aerobic exercise or cardiorespiratory fitness affect the hippocampal subfields differentially in older adults remains elusive. This knowledge is critical, given that the hippocampal subfields have differential vulnerability to aging (Small et al., 2004; Small et al., 2002; for review see Small, Schobel, Buxton, Witter, & Barnes, 2011) and are associated with unique cognitive and computational operations implicated in learning and memory (for reviews see Kesner, Lee, & Gilbert, 2004; Small et al., 2011).

Thus, the objective of the current study was to investigate the relationships between cardiorespiratory fitness, hippocampal subfield volumes, and mnemonic discrimination task performance in cognitive aging. We sought to test the hypothesis that cardiorespiratory fitness will be associated with enhanced function and structure of the hippocampal subfields, which will be evidenced by (a) a positive relationship between cardiorespiratory fitness and mnemonic discrimination task performance, and (b) a positive relationship between cardiorespiratory fitness and hippocampal subfield volumes. First, we investigated if the DG-specific effects of increased cardiorespiratory fitness observed in young adults, including enhanced mnemonic discrimination task performance and increased left anterior DG/CA3 volume (Nauer, Dunne, et al., 2020), extend to older adults. Next, we aimed to examine the relationships between cardiorespiratory fitness and other hippocampal subfield volumes given that some of the older adult human literature (Rosano et al., 2017; Varma et al., 2016) suggests that the CA1 and subiculum may also be modulated by cardiorespiratory fitness. To evaluate hippocampal function and structure at the subfield level in older adults, we utilized a hippocampally dependent mnemonic discrimination task (MDT) designed to putatively measure behavioral pattern separation and high-resolution structural MRI with an automatic segmentation protocol optimized for hippocampal subfield-level precision. Our results indicate that MDT accuracy was lowest in the task condition with the greatest stimuli similarity, and left DG/CA3 body volume significantly predicted performance in this condition in our sample. Furthermore, cardiorespiratory fitness was not associated with DG/CA3 volume as predicted, but was associated with bilateral subiculum volume in women, not men. Altogether, these findings provide further support for a role of the DG in behavioral pattern separation in humans and suggest that cardiorespiratory fitness may have differential effects on hippocampal subfield integrity in older adult men and women.

2 | METHODS

2.1 | Participants

Participant data for this study were acquired from three larger studies in our laboratory examining the effects of aerobic exercise training on the hippocampal memory system (MRI Study 1: Neuroimaging Study of Exercise and Memory Function, ClinicalTrials.gov Identifier: NCT02057354; MRI Study 2: The Entorhinal Cortex and Aerobic Exercise in Aging, ClinicalTrials.gov Identifier: NCT02775760; Behavioral Study: Physical Activity and Cognition Study, ClinicalTrials.gov Identifier: NCT02773121). For these studies, we recruited healthy older adults between the ages of 55–85 years from the greater Boston area. Inclusion criteria based on self-report included fluency in English (must have attended elementary school and higher in English), nonsmoking within the past 6 months, and not participating in at least 30 min of moderate-intensity physical activity on at least 3 days per week for at least 3 months. Exclusion criteria based on self-report included poor vision that cannot be corrected; past or present

conditions that affect cognitive function or may be associated with altered brain function and structure (e.g., learning disabilities, neurological disorders/conditions, psychiatric disorders/conditions, and/or severe stress); use of prescription medication or other drugs that are psychoactive; and/or drug abuse or alcohol misuse, given that this may affect performance on cognitive outcome measures. Exclusion criteria that were examined at the first study visit to determine eligibility included evidence of cognitive impairment, as determined by performance on the Dementia Rating Scale-2 (DRS-2) (Mattis, 1976), the Trail-Making Test (TMT) Versions A and B (Strauss, Sherman, & Spreen, 2006), and the Stroop Test, Victoria version (VST; Strauss et al., 2006); evidence of impaired brain function and/or structure, indicated by a difference greater than five kilograms on the grip strength test using the Smedley Dynamometer (Strauss et al., 2006); and evidence of psychiatric disorders or conditions and/or severe stress that affect cognitive function or may be associated with altered brain function and structure, examined with the Beck Anxiety Inventory (Beck & Steer, 1993), the Geriatric Depression Scale Long Form (30-item; Yesavage et al., 1982), and the Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983). Participants were also excluded if they self-reported contra indicators for participation in fitness testing or exercise training. This included the presence of an acute infection, diagnosis of metabolic conditions (e.g., diabetes mellitus type 1 or type 2, kidney failure, liver disease, thyroid disorders such as thyrotoxicosis/hyperthyroidism), cancer, and/or severe anemia; heart, circulatory, or respiratory conditions, current musculoskeletal impairment (without primary care physician clearance), or diagnosis of electrolyte disorder; and/or use of prescription medications or other drugs that are cardioactive. Additionally, participants were excluded if their anthropometric measures, as measured at the first study visit, were indicative of high-risk obesity (Pescatello & American College of Sports Medicine, 2014): participants with a body mass index (BMI) greater than or equal to 35 kg/m^2, *and* a waist circumference greater than 102 cm (40 in.) for men or 88 cm (35 in.) for women, *and* a hip-to-waist ratio greater than 1.03 for men or 0.90 for women were excluded. Participants were only excluded if all three criteria were met. These exclusion criteria were used because of the established relationship between obesity and increased risk of cardiovascular disease and musculoskeletal impairment, which affects risk stratification for participation in fitness testing and exercise training (Pescatello & American College of Sports Medicine, 2014). Finally, two of the three studies from which we retrieved data included MRI scans (MRI Study 1, MRI Study 2). These participants were screened for additional exclusion criteria, including claustrophobia; the presence of ferromagnetic metal in or on the body that cannot be removed (e.g., pacemaker, defibrillator, metal implant that is ferrous/not MRI compatible); and/or extreme motion-sickness (which may be aggravated by tasks completed in the scanner; not reported here). All participants provided signed, informed consent prior to participation in the study.

All protocols and study procedures adhered to the Code of Ethics of the World Medical Association and were approved by the Boston University Medical Campus Institutional Review Board. Participants performed either two or three study visits at Boston University

School of Medicine and/or Boston University Charles River Campus depending on their study of enrollment. Participants enrolled in the behavioral study without MRI testing completed two study visits: the first included informed consent and eligibility screening; the second included behavioral testing followed by fitness testing. Participants enrolled in the two MRI studies completed three study visits: the first included informed consent and eligibility screening; the second included fitness testing; and the third included behavioral testing and a structural MRI scan. Because of the known effects of acute exercise on cognition (Suwabe et al., 2018, 2017; reviewed in Tomporowski, 2003), all MRI visits were scheduled at least 24 hr after cardiorespiratory fitness testing or before cardiorespiratory fitness testing.

Altogether, 65 participants completed the required study visits, with 52 participants from the MRI studies and 13 participants from the behavioral study without MRI testing. Participant characteristics (including demographics and physiology) for included participants are summarized by study in Table 1 and by sex in Table 2. Out of the 52 participants with MRI data, eight were excluded due to motion that interfered with preprocessing and/or segmentation procedures, and two were excluded due to incidental findings in the regions of interest on their structural scan. This resulted in 42 participants with usable MRI data. Out of these 42 participants, 35 successfully completed cardiorespiratory fitness testing, whereas the other seven participants failed to reach a criterion of 85% of their age-predicted maximum heart rate during testing (see below for details). This resulted in 35 datasets for the analyses that examined the relationship between cardiorespiratory fitness and hippocampal subfield structure. Secondly, out of the 42 participants with usable MRI data, 21 also

completed the MDT. This resulted in 21 datasets for the analyses that examined the relationship between hippocampal subfield structure and MDT performance. Out of the 13 participants with behavioral data only, 10 successfully completed cardiorespiratory fitness testing, whereas the other three participants failed to reach the 85% criterion of their age-predicted maximum heart rate during testing, as above. Thus, when compiled with the 19 participants who had both behavioral and cardiorespiratory fitness data from the MRI studies, this resulted in 29 datasets for the analyses that examined the relationship between cardiorespiratory fitness and MDT performance.

2.2 | Cardiorespiratory fitness assessment

In the current study, cardiorespiratory fitness was operationally defined as estimated maximal oxygen uptake ($\dot{V}O_{2max}$), the gold-standard measure of cardiorespiratory fitness in exercise physiology (Mitchell, Sproule, & Chapman, 1958; Wagner, 1996). To estimate $\dot{V}O_{2max}$, each participant completed a submaximal incremental treadmill test adhering to a modified Balke protocol (Pescatello & American College of Sports Medicine, 2014). Before testing at the Boston University Fitness and Recreation Center, study staff instructed participants to refrain from aerobic exercise and/or resistance training for 24 hr prior to testing and refrain from any caffeine consumption for 3 hr prior to testing. The incremental treadmill test protocol implemented by our laboratory (Kronman et al., 2020; Nauer, Dunne, et al., 2020; Nauer, Schon, & Stern, 2020) includes three phases: 3 min of warm-up, approximately 8–12 min of incremental exercise, and 3 min of cool-down. Throughout

TABLE 1 Participant characteristics for a sample of 65 older adults (aged 55–85 years) separated by study

	All participants	MRI study 1	MRI study 2	Behavioral study
Demographics—mean ± SD				
N	65 (39 women)	25 (12 women)	27 (19 women)	13 (8 women)
Age (years)	65.66 ± 6.97*	62.56 ± 7.05	68.00 ± 5.47	66.77 ± 7.83
Education (years)	16.42 ± 2.27*	15.88 ± 1.86	17.26 ± 2.46	15.69 ± 2.18
Physiology—mean ± SD				
N (reached 85% HR$_{max}$)	53 (34 women)	19	24	10
Estimated $\dot{V}O_{2max}$ (ml kg^{-1} min^{-1})	29.31 ± 6.08†	30.07 ± 5.31	30.25 ± 6.17	25.60 ± 6.43
Resting heart rate (beats per minute)	69.21 ± 12.03	67.37 ± 14.58	72.42 ± 10.35	65.00 ± 9.09
BMI	27.44 ± 4.85	27.98 ± 3.60	27.82 ± 5.78	25.52 ± 4.41
Physiology—median (IQR)				
Estimated $\dot{V}O_{2max}$ percentile (range 1–99)	84 (32.50)*	79 (43.25)	97 (20.50)	76 (20.00)
Neuropsychological measures—mean ± SD				
TMT B/A ratio	2.162 ± 0.640	2.096 ± 0.618	2.330 ± 0.689	1.941 ± 0.516
VST ratio	2.015 ± 0.421	2.042 ± 0.491	1.953 ± 0.360	2.094 ± 0.408
Neuropsychological measures—median (IQR)				
DRS-2 Total raw score	142 (2.0)	142 (2.0)	141 (3.5)	142 (3.0)
DRS-2 memory raw score	25 (2.0)	25 (1.0)	25 (1.5)	24 (2.0)

Abbreviations: DRS-2, Dementia Rating Scale-2; HR$_{max}$, maximum heart rate; TMT, Trail-Making Test; $\dot{V}O_{2max}$, maximal oxygen uptake; VST, Stroop Test, Victoria version.
Denotes significant difference (p < .05) between studies.
†Denotes a trend toward significance in the difference (p < .10†) between studies.

TABLE 2 Participant characteristics for a sample of 65 older adults (aged 55–85 years) separated by sex

	All participants	Men	Women
Demographics—mean ± SD			
N	65 (39 women)	26	39
Age(years)	65.66 ± 6.97	66.62 ± 8.10	65.03 ± 6.14
Education (years)	16.42 ± 2.27	16.50 ± 2.20	16.36 ± 2.35
Physiology—mean ± SD			
N (reached 85% HR_{max})	53 (34 women)	19	34
Estimated $\dot{V}O_{2max}$ (ml kg^{-1} min^{-1})	29.31 ± 6.08*	31.84 ± 6.08	27.89 ± 5.69
Resting heart rate (beats per minute)	69.21 ± 12.03	70.58 ± 13.31	68.44 ± 11.39
BMI	27.44 ± 4.85	27.64 ± 3.77	27.13 ± 5.28
Physiology—median (IQR)			
Estimated $\dot{V}O_{2max}$ percentile (range 1–99)	84.0 (32.50)**	68.5 (37.25)	93.0 (22.00)
Neuropsychological measures—mean ± SD			
TMT B/A ratio	2.162 ± 0.640	2.215 ± 0.530	2.127 ± 0.708
VST ratio	2.015 ± 0.421	1.942 ± 0.416	2.064 ± 0.423
Neuropsychological measures—median (IQR)			
DRS-2 Total raw score	142 (2.0)	142 (2.0)	141 (2.5)
DRS-2 memory raw score	25 (2)	25 (1)	24 (2)

Abbreviations: DRS-2, Dementia Rating Scale-2; HR_{max}, maximum heart rate; TMT, Trail-Making Test; $\dot{V}O_{2max}$, maximal oxygen uptake; VST, Stroop Test, Victoria version.
Denotes significant difference ($p < .05$) between sexes.
**Denotes very significant ($p < .01$) difference between sexes.

the duration of the test, a study staff member monitored the participant's heart rate. In order to accomplish this, each participant wore a heart rate watch (MRI Study 1, Behavioral Study: Polar, model FT7; MRI Study 2: Polar, model A300) that wirelessly paired to a heart rate monitor fixed to a chest strap (Polar, model H1). Additionally, a study staff member measured and recorded blood pressure every 3 min throughout testing in order to ensure participant safety, as recommended (Pescatello & American College of Sports Medicine, 2014). During the incremental exercise phase of the test, each participant walked on a treadmill at their pre-determined fastest comfortable walking pace while the study staff increased treadmill grade every minute. A study staff member recorded heart rate in the last 5 s of each minute of exercise. Finally, the study staff terminated the test when the participant reached 85% of their age-predicted maximum heart rate (Tanaka, Monahan, & Seals, 2001). This submaximal protocol was chosen for the safety of our older participants.

We estimated oxygen uptake ($\dot{V}O_2$) for each treadmill stage based upon a standard metabolic equation for gross $\dot{V}O_2$ for walking (Equation (1)):

$$\dot{V}O_2 = \left(0.1\,ml\,kg^{-1}\,min^{-1}\cdot S\right) + \left(1.8\,ml\,kg^{-1}\,min^{-1}\cdot S\cdot G\right) + 3.5\,ml\,kg^{-1}\,min^{-1} \tag{1}$$

in which $\dot{V}O_2$, S, and G represent gross oxygen uptake (ml kg^{-1} min^{-1}), treadmill speed (m/min), and the treadmill percent grade/100, respectively (Pescatello & American College of Sports Medicine, 2014). The last component of Equation 1 (3.5 ml kg^{-1} min^{-1}) corresponds to estimated resting $\dot{V}O_2$. We estimated $\dot{V}O_{2max}$ for each participant

using the calculated $\dot{V}O_2$ and its corresponding heart rate, and by exploiting the known linear relationship between $\dot{V}O_2$ and heart rate (Wasserman, 2012). Thus, using the data collected from the submaximal incremental treadmill test, we were able to estimate $\dot{V}O_{2max}$, which operationally defined cardiorespiratory fitness in the current study.

Finally, to appreciate the distribution of cardiorespiratory fitness values collected from our participants within the context of national reference values, we interpolated all estimated $\dot{V}O_{2max}$ scores to age- and sex-specific percentiles as determined by the Fitness Registry and the Importance of Exercise National Database (FRIEND; Kaminsky, Arena, & Myers, 2015). Although these percentiles were not used for further analyses, they provide important insight toward the fitness level of our participants compared to references by age and sex for the US population.

2.3 | Mnemonic discrimination task

2.3.1 | Design of task and stimuli

Our laboratory developed an adapted delayed matching-to-sample task designed to tax mnemonic discrimination ability at parametrically varying levels (Nauer, Dunne, et al., 2020). We created this mnemonic discrimination task (MDT) and collected the data using E-Prime 2.0 (https://pstnet.com/products/e-prime/; RRID:SCR_009567). The stimulus set comprises a series of grayscale images of nonfamous faces edited to consist of only central facial features (i.e., peripheral features such as hair, clothes, etc. were eliminated). We created 10 runs of the task,

which were split into two separate sets (Set A and Set B) to allow for counterbalancing across participants in our larger research studies including an aerobic exercise training program. Within the current study, which includes only baseline data from the larger studies, half of the participants performed Set A and the other half performed Set B.

Importantly, to examine mnemonic discrimination at parametrically varying levels of stimulus similarity, we designed the task using FantaMorph 5 (https://www.fantamorph.com) to morph all face stimuli within a trial to have either low (10%), moderate (30%), or high (50%) overlap with a trial-unique template (Figure 3a). Consequently, all face stimuli in a trial have a parametrically varied percentage of facial similarity. The stimuli in the 50% condition objectively have the greatest representational similarity of all three conditions; thus, this condition in particular theoretically places the highest taxation on pattern separation to successfully resolve interference between stimulus input during encoding (Nauer, Dunne, et al., 2020). We pseudorandomized trial condition (10, 30, and 50%) within each run such that each run had an even number of all three conditions.

2.3.2 | Task procedure

In a single trial of the task (Figure 3b), the participant is subsequently shown two face stimuli (Sample 1, 4 s; Sample 2, 4 s), a black screen with a gray silhouette of a face containing a central crosshair (Delay, 10 s), and a test face stimulus (Test, 4 s). During the Test face presentation, the participant is asked to make a button-press response to indicate if they believe that the Test face is the same as Sample 1 (match trial), the same as Sample 2 (match trial), or is a new face (nonmatch trial). As aforementioned, each face shown in a trial is morphed to have either low (10%), moderate (30%), or high (50%) similarity with a trial-unique template that the participant does not see during the trial. Thus, all three faces in the trial share either 10, 30, or 50% overlap in facial features. Each trial is followed by a variable inter-trial interval (ITI; 6–12 s). Furthermore, within each run, trials were split evenly into nonmatch and match trials, the latter of which were separated evenly into matches to Sample 1 and matches to Sample 2. In the current study, participants completed five runs total with each run comprising 12 trials for a total of 60 completed trials per participant.

To familiarize the participants with the task, each participant completed a practice version of the task which consisted of one run of a distinct set of stimuli. In the practice task, participants received response feedback after each trial. In the experimental version of the task, participants did not receive response feedback after each trial; however, overall accuracy was provided at the end of each run.

2.4 | MRI data acquisition and hippocampal subfield segmentation

2.4.1 | MRI data acquisition—Dataset 1

Twenty-five (12 women) participants enrolled in MRI Study 1 were scanned using a 3 Tesla Phillips Achieva MRI scanner equipped with an eight-channel Sensitivity Encoding (SENSE) head coil at Boston University Center for Biomedical Imaging. We collected two anatomical scans for each participant: first, a high-resolution whole-brain structural T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE) volume (TR (ms) = 6.8; TE (ms) = 3.1; acquisition matrix $M \times P = 256 \times 254$; slices = 150; acquisition voxel MPS (mm^3) = 0.98 × 0.98 × 1.20; field of view FH × AP (mm^2) = 250 × 250; SENSitivity Encoding P reduction (AP): 1.5, P os factor: 1, S reduction (RL): 2; flip angle (degrees) = 9), and second, a structural T2-weighted volume with high in-plane resolution (TR (ms) = 3,000; TE (ms) = 80; acquisition matrix $M \times P = 576 \times 450$; slices = 30; slice gap (mm) = 0.6; acquisition voxel MPS (mm) = 0.40 × 0.41 × 2.00; field of view FH × RL (mm^2) = 230 × 185; SENSitivity Encoding P reduction (RL): 2, P os factor: 1; flip angle (degrees) = 90).

2.4.2 | MRI data acquisition—Dataset 2

Twenty-seven (19 women) participants enrolled in MRI Study 2 were scanned using a 3 Tesla Siemens MAGNETOM Prisma MRI scanner equipped with a stock 64-channel head coil at Boston University Cognitive Neuroimaging Center. We collected two anatomical scans for each participant: a high-resolution whole-brain structural T1-weighted magnetization-prepared rapid acquisition gradient multi-echo (multi-echo MPRAGE [van der Kouwe, Benner, Salat, & Fischl, 2008]) volume (slices (sagittal) = 176; TR (ms) = 2,200; TE (ms) = 1.67; TI (ms) = 1,100; acquisition matrix = 230 × 230; voxel resolution (mm^3) = 1.0 × 1.0 × 1.0; field of view = 230 × 230; GRAPPA acceleration = 4; flip angle (degrees) = 7) and second, a structural T2-weighted volume with high in-plane resolution (TR (ms) = 8,020.0; TE (ms) = 80.0; acquisition matrix = 384 × 384; slices = 30; spacing between slices (mm) = 2; voxel resolution (mm^3) = 0.39 × 0.39 × 2.0; field of view = 384 × 384; flip angle (degrees) = 150).

2.4.3 | Segmentation of hippocampal subfields

To segment the hippocampal subfields, we used the well-established Automatic Segmentation of Hippocampal Subfields (ASHS) software package (https://www.nitrc.org/projects/ashs; Yushkevich et al., 2015). This software requires a T1-weighted volume and an oblique coronal T2-weighted volume with high in-plane resolution (0.4 × 0.4 mm^2). For the current study, we used ASHS to automatically label the main hippocampal subfields on each participant's T2-weighted volume based on the Penn Memory Center 3 T ASHS atlas (Yushkevich et al., 2015). Regions of interest included the following hippocampal subfields, examined bilaterally: CA1, CA2, CA3, DG, and subiculum. Given the lower reliability of the boundaries between DG and CA3, these two regions were combined into one region of interest (DG/CA3) for all statistical analyses. As in our laboratory's previous work (Nauer, Dunne, et al., 2020), we next used established anatomic guidelines (Pruessner, 2000) to manually segment the hippocampal subfields into head, body, and tail. Briefly, slices anterior to the uncal apex disappearance were labeled as head, slices posterior to the uncal apex
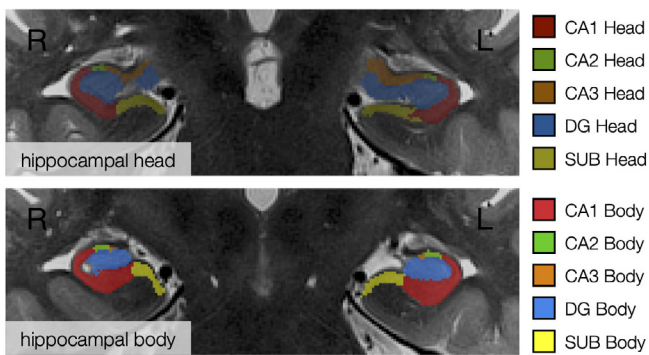
FIGURE 1 Example of hippocampal head and body subfield segmentations. The hippocampus was automatically segmented into the following subfields: Cornu Ammonis (CA) 1–3 (CA1, CA2, CA3), dentate gyrus (DG), and subiculum (SUB). ROIs for analyses included CA1, DG/CA3, and SUB. Next, the hippocampal subfields were manually segmented into the head, body, and tail using established anatomic guidelines. Subfield delineations in the tail were not used for analyses due to low reliability within and between participants

disappearance but anterior to the wing of the ambient cistern were labeled as body, and slices posterior to the wing of the ambient cistern were labeled as tail (see example slices of hippocampal head and body in Figure 1). One researcher (K. L. K.) performed the head, body, and tail delineations for each participant on all coronal slices of the T2-weighted volumes in native space using ITK-SNAP Version 3.8.0 (www.itksnap.org; RRID:SCR_002010; Yushkevich et al., 2006). All segmentations were inspected for accuracy. Notably, head, body, and tail segmentations for the DG/CA3 only were used for statistical analyses, as the DG/CA3 subfields are to our knowledge the only hippocampal subfields for which there is evidence for differential effects of cardiore-spiratory fitness along the long axis of the hippocampus. Furthermore, DG/CA3 segmentations suitable for analyses were limited to the head and body of DG/CA3, given that subfield delineations in the tail were not reliable within or between participants.

Finally, we corrected all region of interest volumes for differences in intracranial volume (ICV) prior to statistical analyses. ICV was supplied by the ASHS output for each participant. We used an adjustment that is well established in the literature (Erickson et al., 2009; Erickson et al., 2011; Head, Rodrigue, Kennedy, & Raz, 2008; Nauer, Dunne, et al., 2020; Raz et al., 2004; Raz et al., 2005) and takes an analysis of covariance approach (Equation 2):

$$\text{Adjusted volume} = \text{Raw volume} - b \times (\text{ICV} - \text{mean ICV}) \quad (2)$$

where b is the slope of the regression of a region of interest volume on ICV. Adjusted volumes were used in all applicable analyses described in the current study.

2.5 | Statistical methods

All statistical analyses were performed using RStudio Version 1.0.153 (https://rstudio.com/; RRID:SCR_000432). First, we examined all continuous variables, including physiological measures, neuropsychological

measures, MDT performance, and hippocampal subfield volumes, for normality using a Shapiro–Wilk test of normality. Next, we summarized all continuous variables with normal distributions using mean and standard deviation. We examined the demographic data (Age, Education), physiological data (estimated $\dot{V}O_{2max}$, resting heart rate, BMI), and neuropsychological measures (TMT B/A ratio, VST ratio) for significant differences between studies using a one-way analysis of variance (ANOVA) for independent samples and for significant differences between sexes using Welch's Two-Sample t-tests. Given that estimated $\dot{V}O_{2max}$ percentile, DRS-2 total raw score, and DRS-2 memory raw score did not assume normality, we summarized these variables using median and interquartile range. Furthermore, we checked these variables for study differences using Kruskal–Wallis Rank Sum Tests and for sex differences using Mann–Whitney U tests.

We next sought to examine relationships between neuropsychological measures, participant demographics, and MDT performance. The primary outcome measure for the MDT, corrected accuracy, was calculated by subtracting false alarms (incorrect nonmatch trials) from hits (correct match trials). To examine relationships between MDT corrected accuracy and the normally distributed variables of TMT B/A ratio and VST ratio, we used simple Pearson correlations. To examine relationships between MDT corrected accuracy and the nonnormally distributed variables of DRS-2 total raw score and DRS-2 memory raw score, we used Spearman's rank-order correlations. We corrected all Pearson and Spearman's correlations for multiple comparisons by applying Benjamini & Hochberg (false discovery rate, FDR) corrections (Benjamini & Hochberg, 1995) for similarity condition (three). These corrections were implemented using the *p.adjust* function in R. Finally, we used Welch's Two-Sample t-tests to examine differences between the sexes in MDT performance in all three similarity conditions. As with the above correlations, these tests were all corrected for multiple comparisons using FDR corrections.

To examine relationships between estimated $\dot{V}O_{2max}$, MDT performance, and hippocampal subfield volumes, we employed a statistical approach previously used in our laboratory (Nauer, Dunne, et al., 2020). To perform this approach, we first investigated relationships between variables by running simple Pearson correlations. Next, we examined significant correlations ($p < .05*$) and correlations that were trending toward significance ($p < .10†$) using linear regression models to control for covariates such as age (years, continuous), sex (men, women), education (degree years), and study (MRI Study 1, MRI Study 2, Behavioral Study). When appropriate, we also included DRS-2 memory raw score as a covariate to determine if an observed relationship was mediated by memory-related capability. Both the correlational step and the linear regression step of this approach involved correction for multiple comparisons by applying FDR corrections.

2.5.1 | MDT performance and its relationship to DG/CA3 volume

First, we examined participant performance between the different MDT similarity conditions (10, 30, and 50%). To test if performance significantly varied between similarity conditions, we used a one-way

ANOVA for paired samples with subsequent post hoc testing consisting of a Tukey test. We predicted that older adults would demonstrate the lowest accuracy in the MDT condition with the highest stimuli similarity.

Next, we sought to examine DG/CA3 subfield volume as a predictor of MDT performance. First, we examined simple Pearson correlations between ICV-adjusted bilateral DG/CA3 head and body volumes and task performance in the 50% MDT condition. We examined the DG/CA3 head and DG/CA3 body separately given the well-established literature suggesting different roles of the anterior hippocampus in stress and emotion and the posterior hippocampus in cognitive function (Bannerman et al., 2004; Fanselow & Dong, 2010; Strange, Witter, Lein, & Moser, 2014). Furthermore, we focused on behavioral performance in the 50% MDT condition given that this condition was designed to place the greatest taxation on mnemonic discrimination ability. After running Pearson correlations, we adjusted these values by applying FDR corrections for number of comparisons (four, i.e., laterality and head/body split). To follow, we evaluated significant correlations ($p < .05$*) and correlations that were trending toward significance ($p < .10$[†]) when controlling for age, sex, education, and study using linear regression analyses. Additionally, we ran a second set of linear regression analyses with DRS-2 memory raw score included as a covariate to determine if any observed relationship may be mediated by memory-related capability. For the aforementioned analyses, any participant whose ICV-adjusted ROI volume was more than two standard deviations greater than or less than the sample mean was considered an outlier and removed from the analysis (Seo, 2006).

2.5.2 | Cardiorespiratory fitness as a predictor of MDT performance and hippocampal subfield volumes

Our next goal was to examine cardiorespiratory fitness as a predictor of MDT performance and hippocampal subfield volumes. To begin, we examined simple Pearson correlations between estimated $\dot{V}O_{2max}$, which operationally defined cardiorespiratory fitness, and task performance in all MDT conditions. This aimed to test the prediction that cardiorespiratory fitness would predict task accuracy in the 50% MDT condition. We corrected these correlations for multiple comparisons (three, i.e., number of MDT conditions) using FDR corrections. Subsequently, we examined significant correlations ($p < .05$*) and correlations that were trending toward significance ($p < .10$[†]) when controlling for age, sex, education, and study using linear regression analyses. We also ran a second set of linear regression analyses that included DRS-2 memory raw score as a covariate to determine if any observed relationships may be mediated by memory-related capability.

Next, we sought to examine cardiorespiratory fitness as a predictor of adjusted hippocampal subfield volumes. As in our previously published work (Nauer, Dunne, et al., 2020) and the previous analyses in the current study, we began by examining simple Pearson correlations between estimated $\dot{V}O_{2max}$ and ICV-adjusted hippocampal subfield volumes. Given the sex-specific findings in other exercise studies (Brown, Peiffer, & Martins, 2013; Colcombe & Kramer, 2003; Varma

et al., 2016; Varma, Chuang, Harris, Tan, & Carlson, 2015), we also evaluated these correlations in sub-groups by sex. Thus, we applied FDR corrections for laterality and sex to our Pearson correlations when appropriate. To follow, we examined significant correlations ($p < .05$*) and correlations that were trending toward significance ($p < .10$[†]) while controlling for age, sex, education, and study using linear regression analyses. We also ran a second set of linear regression analyses that included DRS-2 memory raw score as a covariate to determine if any observed relationships may be mediated by memory-related capability. Any participant whose ICV-adjusted ROI volume was more than two standard deviations greater than or less than the sample mean was considered an outlier and removed from the analysis (Seo, 2006). We predicted that cardiorespiratory fitness would be associated with DG/CA3 volume, as evidenced by our laboratory in young adults (Nauer, Dunne, et al., 2020), and may also demonstrate a positive relationship with other hippocampal subfields, including the CA1 and subiculum, as evidenced in other publications (Rosano et al., 2017; Varma et al., 2016).

3 | RESULTS

3.1 | Participant demographics, physiology, and neuropsychological measures

Participant characteristics (including demographics and neuropsychological measures) for 65 participants are summarized by study in Table 1 and by sex in Table 2. Fifty-three of these participants successfully completed the submaximal treadmill test to estimate $\dot{V}O_{2max}$, which operationally defined cardiorespiratory fitness. Physiological measures from these 53 participants are also summarized by study in Table 1 and by sex in Table 2.

3.1.1 | Differences between studies in demographics

For participant demographics, we examined differences between study samples (MRI Study 1:1, MRI Study 2:2, Behavioral Study: 3) using analysis of variance models. We completed this step as a control to examine equivalence between cohorts. There was a significant difference in age between studies ($F_{2,62} = 4.625$, $p = .013$*). A Tukey multiple comparisons of means test indicated that this was due to a significant difference between MRI Study 1 (1) and MRI Study 2 (2) (3 vs. 1: $p = .158$; 2 vs. 1: $p = .012$*; 3 vs. 2: $p = .846$), demonstrating MRI Study 2 participants were significantly older than MRI Study 1 participants. There was also a significant difference in education between studies ($F_{2,62} = 3.469$, $p = .037$*). A Tukey multiple comparisons of means test indicated that this was due to a trend toward significance in the difference between MRI Study 1 (1) and MRI Study 2 (2) and a trend toward significance in the difference between MRI Study 2 (2) and the Behavioral Study (3) (3 vs. 1: $p = .966$; 2 vs. 1: $p = .067$[†]; 3 vs. 2: $p = .094$[†]). This suggests that MRI Study

2 participants tended to be more highly educated than those of the other two groups, although this was not statistically significant.

3.1.2 | Differences between sexes in demographics

For participant demographics, we examined differences between sexes within all participants using Welch's Two-Sample t tests. There was no significant difference in age ($t[43.617] = 0.851$, $p = .399$) or education ($t[56.130] = 0.247$, $p = .806$) between men and women across the three cohorts.

3.1.3 | Differences between studies in physiological variables

Mean estimated $\dot{V}O_{2max}$ over all participants was 29.31 ± 6.08 ml kg^{-1} min^{-1} (mean $\pm$ SD; Table 1). As with participant demographics, we investigated differences between study samples as a control to examine equivalence using analysis of variance models. There was a trend toward significance in the difference in estimated $\dot{V}O_{2max}$ between studies ($F_{2,50} = 2.420$, $p = .099^{\dagger}$); however, a Tukey multiple comparisons of means test did not reach significance when examining differences between cohorts (MRI Study 1:1, MRI Study 2:2, Behavioral Study: 3) (3 vs. 1: $p = .140$; 2 vs. 1: $p = .995$; 3 vs. 2: $p = .103$). There were no significant differences between studies in resting heart rate ($F_{2,50} = 1.737$, $p = .187$) or BMI ($F_{2,50} = 0.975$, $p = .384$). Finally, to appreciate the distribution of cardiorespiratory fitness values collected from our participants within the context of national normative values, we interpolated all estimated $\dot{V}O_{2max}$ values to age- and sex-specific percentiles, as described in the Methods section above. Estimated $\dot{V}O_{2max}$ percentiles were not normally distributed, as determined by a Shapiro–Wilk test of normality ($W = 0.854$, $p < .001^{***}$). The median estimated $\dot{V}O_{2max}$ percentile over all participants was 84 (32.5) (median [interquartile range]). There was a significant difference in the estimated $\dot{V}O_{2max}$ percentile between studies, as determined by a Kruskal-Wallis Rank Sum Test ($H(2) = 8.523$, $p = .014^{*}$). Post hoc testing of pairwise comparisons demonstrated this was due to a significant difference between MRI Study 1 (1) and MRI Study 2 (2) (3 vs. 1: $p = .738$; 2 vs. 1: $p = .012^{*}$; 3 vs. 2: $p = .170$), showing that MRI Study 2 participants were fitter than participants of MRI Study 1.

3.1.4 | Differences between sexes in physiological variables

We examined differences between sexes within all participants using Welch's Two Sample t tests. Estimated $\dot{V}O_{2max}$ varied significantly by sex, such that men demonstrated significantly higher estimated $\dot{V}O_{2max}$ with a mean $\pm$ SD of 31.84 ± 6.08 ml kg^{-1} min^{-1} compared to women with a mean $\pm$ SD of 27.89 ± 5.69 ml kg^{-1} min^{-1} ($t[35.276] = 2.322$, $p = .026^{*}$). Figure 2 demonstrates the variability within

estimated $\dot{V}O_{2max}$ values of participants. Neither resting heart rate ($t[32.755] = 0.590$, $p = .559$) nor BMI ($t[50.069] = 0.770$, $p = .445$) differed significantly between sexes. Finally, as estimated $\dot{V}O_{2max}$ percentiles were not normally distributed, we used a Mann–Whitney U test to examine the data for sex differences. This analysis suggested that there was a significant difference in estimated $\dot{V}O_{2max}$ percentiles between men and women, such that women had significantly higher estimated $\dot{V}O_{2max}$ percentiles with a median (IQR) of 93.0 (22.00) compared to men with a median (IQR) of 68.5 (37.25) ($W = 121$, $p = .002^{**}$).

3.1.5 | Differences between studies in neuropsychological variables

As with participant demographics and physiological variables, we investigated differences between study samples in neuropsychological
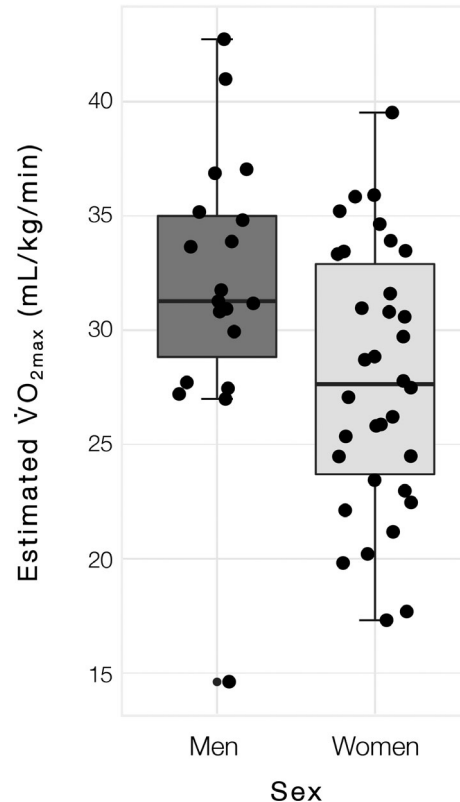

FIGURE 2 Distribution of cardiorespiratory fitness, as measured by estimated maximal oxygen uptake ($\dot{V}O_{2max}$), for a sample of 53 older adults separated by sex. Fifty-three participants successfully completed the submaximal treadmill test to estimate $\dot{V}O_{2max}$, which operationally defined cardiorespiratory fitness. Mean estimated $\dot{V}O_{2max}$ over all participants was 29.31 ± 6.08 ml kg^{-1} min^{-1} (mean $\pm$ SD). Estimated $\dot{V}O_{2max}$ varied significantly by sex, such that men demonstrated significantly higher estimated $\dot{V}O_{2max}$ with a mean of 31.84 ± 6.08 ml kg^{-1} min^{-1} compared to women with a mean of 27.89 ± 5.69 ml kg^{-1} min^{-1} ($t[35.276] = 2.322$, $p = .026^{*}$). This figure demonstrates the variability within estimated $\dot{V}O_{2max}$ values of participants

measures using analysis of variance models and Kruskal-Wallis tests. Both TMT B/A ratio ($W = 0.967$, $p = .076^{\dagger}$) and VST ratio ($W = 0.972$, $p = .146$) were normally distributed; thus, we examined differences between study samples in these variables using analysis of variance models. There was no significant difference between studies in TMT B/A ratio ($F_{2,62} = 1.887$, $p = .160$) or in VST ratio ($F_{2,62} = 0.571$, $p = .568$). Because DRS-2 total raw scores ($W = 0.933$, $p = .002^{**}$) and DRS-2 memory raw scores ($W = 0.772$, $p < .001^{***}$) were not normally distributed, we examined differences between studies in these variables using a Kruskal–Wallis Rank Sum test. There was no significant difference between studies in DRS-2 total raw score ($H(2) = 0.001$, $p = 1$) or in DRS-2 memory raw score ($H(2) = 0.323$, $p = .851$).

3.1.6 | Differences between sexes in neuropsychological variables

Next, we examined differences between sexes within all participants using Welch's Two-Sample t tests and Mann–Whitney U tests. There was no significant difference between men and women in TMT B/A ratio ($t[62.084] = 0.571$, $p = .570$) or in VST ratio ($t[54.347] = -1.146$, $p = .257$). Furthermore, there was no significant difference between men and women in DRS-2 total raw scores ($W = 465.5$, $p = .577$) or in DRS-2 memory raw scores ($W = 535$, $p = .690$).

3.2 | Participant demographics, neuropsychological measures, and MDT performance

We first sought to examine the relationships between the neuropsychological variables and MDT corrected accuracy. Given that data for the TMT B/A ratio and VST ratio were normally distributed, we examined these relationships using simple Pearson correlations (Table S1). There was a significant correlation between TMT B/A ratio and corrected accuracy only in the 10% similarity condition ($r = -.416$, $t[36] = -2.745$, $p = .009^{**}$, p_{adjust} for conditions = .028*). The correlation between TMT B/A ratio and corrected accuracy in the 30% similarity condition was trending toward significance ($r = -.303$, $t[36] = -1.911$, $p = .064^{\dagger}$, p_{adjust} for similarity condition = .096†), and there was no significant correlation between TMT B/A ratio and corrected accuracy in the 50% similarity condition ($r = -.211$, $t[36] = -1.293$, $p = .204$, p_{adjust} for similarity condition = .204). Additionally, there was no significant correlation between VST ratio and corrected accuracy in any of the similarity conditions (10%: $r = -.157$, $t[36] = -0.953$, $p = .347$, p_{adjust} for similarity condition = .520; 30%: $r = -.056$, $t[36] = -0.337$, $p = .738$, p_{adjust} for similarity condition = .738; 50%: $r = -.233$, $t[36] = -1.438$, $p = .159$, p_{adjust} for similarity condition = .477). Because both DRS-2 total raw score and DRS-2 memory raw score were not normally distributed, we examined these relationships using Spearman's rank-order correlations (Table S1). There was no significant correlation between DRS-2 total raw score and corrected accuracy in any similarity condition (10%: $r_s = .234$, $p = .158$, p_{adjust} for similarity condition = .158; 30%: $r_s = .327$,

$p = .045^{*}$, p_{adjust} for similarity condition = .119; 50%: $r_s = .288$, $p = .079^{\dagger}$, p_{adjust} for similarity condition = .119). Comparatively, there was a significant relationship between DRS-2 memory raw score and corrected accuracy in the 10% similarity condition ($r_s = .411$, $p = .010^{*}$, p_{adjust} for similarity condition = .031*) and 50% similarity condition ($r_s = .359$, $p = .027^{*}$, p_{adjust} for similarity condition = .040*). The correlation between DRS-2 memory raw score and corrected accuracy in the 30% similarity condition was trending toward significance ($r_s = .280$, $p = .089^{\dagger}$, p_{adjust} for similarity condition = .089†). Altogether, these results show that corrected accuracy in the 10% similarity condition significantly correlated with TMT B/A ratio and that corrected accuracy in the 10 and 50% similarity conditions significantly correlated with DRS-2 memory raw score.

Next, we used Welch's Two-Sample t tests to examine differences between the sexes in task performance. There were no significant differences between men and women in any of the task similarity conditions (10%: $t[34.525] = .146$, $p = .885$, p_{adjust} for similarity condition = .909; 30%: $t[35.843] = 0.116$, $p = .909$, p_{adjust} for similarity condition = .909; 50%: $t[35.889] = 1.169$, $p = .250$, p_{adjust} for similarity condition = .750).

3.3 | Older adults demonstrate lowest MDT accuracy in condition with highest stimuli similarity

To characterize task performance for our study sample, we performed a one-way ANOVA for paired samples. This analysis demonstrated a significant effect of similarity on task performance ($F = 39.75$, $p < .001^{***}$; Figure 3c). Post hoc testing using a Tukey multiple comparisons of means test indicated that average corrected accuracy in the 50% MDT condition (mean ± SD corrected accuracy = -0.011 ± 0.252) was significantly lower than in the 30% MDT condition (mean ± SD corrected accuracy = 0.190 ± 0.237; $p < .01^{**}$ compared to 50%) and lower than in the 10% MDT condition (mean ± standard deviation corrected accuracy = 0.313 ± 0.294; $p < .001^{***}$ compared to 50%). This result supports our original prediction that older adults will demonstrate the lowest task accuracy in the MDT condition with the greatest stimuli similarity.

3.4 | Left DG/CA3 body volume predicts MDT performance when stimuli similarity is high

We next evaluated bilateral DG/CA3 volumes as independent predictors of MDT performance. We examined DG/CA3 head and DG/CA3 body separately given the well-established literature suggesting different roles of the anterior and posterior hippocampus in stress/ emotion and cognitive function, respectively (Bannerman et al., 2004; Fanselow & Dong, 2010; Strange et al., 2014). Additionally, we focused on behavioral performance in the 50% MDT condition given that this condition was designed to place the greatest taxation on mnemonic discrimination ability. Notably, two participants (both MRI Study 1, sex: men) were removed from the right and left DG/CA3
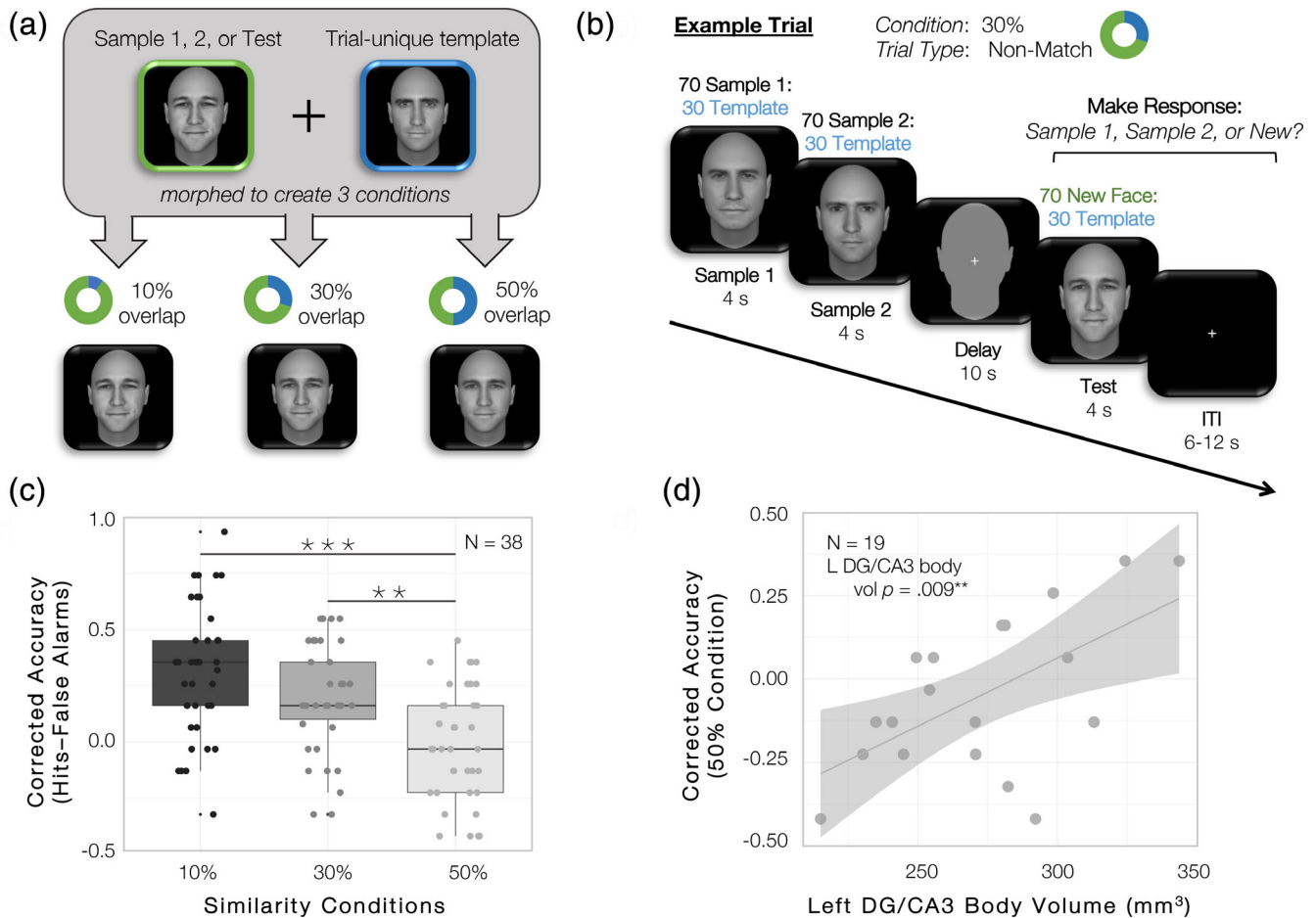
FIGURE 3 Mnemonic discrimination task, behavioral results, and the relationship between behavior and DG/CA3 body volume. (a) Stimuli similarity was parametrically manipulated by morphing each face to a trial unique template at low (10%), moderate (30%), or high (50%) similarity. (b) An example trial of the mnemonic discrimination task. Half of the trials were nonmatch trials, in which the Test face was not the same as Sample 1 or Sample 2 (as demonstrated above), and half of the trials were match trials, in which the Test face was the same as either Sample 1 or Sample 2 (ITI, inter-trial interval) (c) Older adults demonstrated significantly decreased accuracy in the 50% MDT condition, the condition with the greatest mnemonic discrimination requirement. (Analysis of variance demonstrated a significant effect of similarity on task performance ($F = 39.75$, $p < .001***$); post hoc testing using a Tukey multiple comparisons of means indicated that average corrected accuracy in the 50% MDT condition was significantly lower than in the 30% MDT condition ($p < .01**$) and the 10% MDT condition ($p < .001***$).) (d) Left DG/CA3 body volume significantly predicted corrected accuracy in the 50% MDT condition ($N = 19$, L DG/CA3 body vol $t[14] = 3.051$, $p = .009**$, delta-$R^2 = 0.303$; model $F_{4,14} = 4.167$, adjusted $R^2 = 0.413$)

head analyses and two participants (both MRI Study 1, sex: one man, one woman) were removed from the left DG/CA3 body analyses, as their ICV-adjusted bilateral ROI volumes were more than two standard deviations greater than or less than the sample mean. Pearson correlation revealed a significant correlation between left DG/CA3 body volume and corrected accuracy in the 50% MDT condition, which remained significant after FDR correction for laterality and head/body split ($N = 19$, $r = .589$, $t[17] = 3.004$, $p = .008**$ (p_{adjust} for laterality and head/body split = .032*)). There was no significant correlation between right DG/CA3 head volume ($N = 19$, $r = .060$, $t[17] = .249$, $p = .806$ (p_{adjust} for laterality and head/body split = .806)), right DG/CA3 body volume ($N = 21$, $r = .414$, $t[19] = 1.985$, $p = .062^{\dagger}$ (p_{adjust} for laterality and head/body split = .124)), or left DG/CA3

head volume ($N = 19$, $r = -.295$, $t[17] = -1.275$, $p = .219$ (p_{adjust} for laterality and head/body split = .293)) and corrected accuracy in the 50% MDT condition.

Subsequently, we examined significant correlations when controlling for age, sex, and education using linear regression analyses. (The study covariate was not included, given that all data for these analyses came from MRI Study 1). Given that the only statistically significant correlation was between left DG/CA3 body volume and corrected accuracy in the 50% MDT condition, we exclusively examined this relationship using linear regression analysis. Notably, the same two participants (both MRI Study 1, sex: one man, one woman) were removed from the left DG/CA3 body linear regression analysis as was performed in the correlational step, above. This analysis demonstrated

that left DG/CA3 body volume significantly predicted corrected accuracy in the 50% MDT condition (L DG/CA3 body vol $t[14]$ = 3.051, p = .009**, delta-R^2 = .303; model $F_{4,14}$ = 4.167, adjusted-R^2 = .413; Figure 3d). What is more, this finding remained significant when DRS-2 memory raw score was included as a covariate (L DG/CA3 body vol $t[13]$ = 3.249, p = .006**, delta-R^2 = 0.323; DRS-2 memory raw score $t[13]$ = 1.380, p = .191, delta-R^2 = 0.058; model $F_{5,13}$ = 3.930, adjusted R^2 = .449). This result supports our original prediction that greater DG/CA3 volume is associated with better performance in the MDT condition with the highest stimuli similarity.

3.5 | The relationship between cardiorespiratory fitness and MDT performance is not statistically significant

Second, we examined estimated $\dot{V}O_{2max}$ as a predictor of MDT performance. Using simple Pearson correlations, we found a significant correlation between estimated $\dot{V}O_{2max}$ and corrected accuracy only in the 50% MDT condition, which was trending toward significance after FDR correction for number of task conditions (three) (N = 29, r = .428, $t[27]$ = 2.461, p = .021* (p_{adjust} for conditions = .062†)). There was no significant correlation between estimated $\dot{V}O_{2max}$ and corrected accuracy in the 30% MDT condition (N = 29, r = −.014, $t[27]$ = −.070, p = .945 (p_{adjust} for conditions = .945)) or the 10% MDT condition (N = 29, r = .185, $t[27]$ = 0.979, p = .336 (p_{adjust} for conditions = .504)). We then examined these relationships when controlling for covariates of age, sex, education, and study using linear regression. Given that the only correlation trending toward significance was between estimated $\dot{V}O_{2max}$ and the 50% MDT condition, we exclusively examined this relationship using linear regression analysis. Estimated $\dot{V}O_{2max}$ did not significantly predict corrected accuracy in the 50% condition (estimated $\dot{V}O_{2max}$ $t[23]$ = 1.578, p = .128, delta-R^2 = 0.063; model $F_{5,23}$ = 3.323, adjusted R^2 = 0.293). Furthermore, this relationship remained nonsignificant when DRS-2 memory raw score was included

as a covariate (estimated $\dot{V}O_{2max}$ $t[22]$ = 1.686, p = .106, delta-R^2 = 0.064; DRS-2 memory raw score $t[22]$ = 1.921, p = .068†, delta-R^2 = 0.083; model $F_{6,22}$ = 3.708, adjusted R^2 = 0.367). These findings do not support our prediction that greater cardiorespiratory fitness is associated with better mnemonic discrimination task performance in the 50% similarity condition.

3.6 | Cardiorespiratory fitness significantly predicts bilateral subiculum volume in women

Next, we investigated the relationships between cardiorespiratory fitness and hippocampal subfield volumes. We began with simple Pearson correlations between estimated $\dot{V}O_{2max}$ and ICV-adjusted volumes, applying FDR corrections for laterality when appropriate. Pearson's r values for all correlations are reported in Table 3. The first point of investigation for this component of the current study was to determine if our findings in young adults, namely, the correlation between cardiorespiratory fitness and left DG/CA3 head volume (Nauer, Dunne, et al., 2020), extends to older adults. Our older adult data did not demonstrate a significant correlation between estimated $\dot{V}O_{2max}$ and left DG/CA3 head volume (N = 35, r = .008, $t[33]$ = .048, p = .962 (p_{adjust} for laterality = .962)). The correlation between estimated $\dot{V}O_{2max}$ and bilateral DG/CA3 head volume was also nonsignificant (N = 35, r = −.025, $t[33]$ = −0.145, p = .886). Next, we examined the relationship between cardiorespiratory fitness and other whole hippocampal subfields (CA1, subiculum), as in other studies (Rosano et al., 2017; Varma et al., 2016). We examined these relationships collapsed across hemispheres, as we had a hypothesis regarding lateralized effects of estimated $\dot{V}O_{2max}$ only for the DG/CA3 subfield (Nauer, Dunne, et al., 2020). Notably, one participant (MRI Study 2, sex: man) was removed from the CA1 analyses and one participant (MRI Study 2, sex: woman) was removed from the subiculum analyses, as their ICV-adjusted bilateral ROI volumes were more than two standard deviations greater than or less than the sample mean. There was

TABLE 3 Pearson correlation between $\dot{V}O_{2max}$ and hippocampal subfield volumes

| Volumes (mm³) | Pearson's r; p value (p_{adjust}) | | |
| | All participants | All participants—Separated by sex | |
		Men	Women
DG/CA3 (R/L, head only)			
L DG/CA3 Head	r = .008, p = .962 (p_{adjust} = .962)	r = −.391; p = .187 (p_{adjust} = .231)	r = .266; p = .231 (p_{adjust} = .231)
R DG/CA3 Head	r = −.056, p = .752 (p_{adjust} = .962)	r = −.434; p = .138 (p_{adjust} = .231)	r = .270; p = .225 (p_{adjust} = .231)
Bilateral DG/CA3 Head	r = −.025, p = .886	r = −.473; p = .103 (p_{adjust} = .194)	r = .288; p = .194 (p_{adjust} = .194)
Whole hippocampal subfields (bilateral)			
Bilateral CA1	r = .126, p = .478	r = .018; p = .957 (p_{adjust} = .957)	r = .220; p = .324 (p_{adjust} = .648)
Bilateral DG/CA3	r = .126, p = .478	r = −.233; p = .444 (p_{adjust} = .444)	r = .384; p = .086† (p_{adjust} = .171)
Bilateral subiculum	r = .366, p = .033*	r = −.069; p = .823 (p_{adjust} = .823)	r = .608; p = .003** (p_{adjust} = .007**)

Denotes significant correlation (p < .05).
Denotes very significant correlation (p < .01).
†Denotes a trend toward significance (p < .10†).

no significant correlation between estimated $\dot{V}O_{2max}$ and bilateral CA1 volume ($N = 34$, $r = .126$, $t[32] = 0.718$, $p = .478$). However, there was a significant positive correlation between estimated $\dot{V}O_{2max}$ and bilateral subiculum volume ($N = 34$, $r = .366$, $t[32] = 2.228$, $p = .033^*$).

Subsequently, we evaluated these correlations in sub-groups of men and women, given the sex-specific findings in other studies that examined the relationship between cardiorespiratory fitness and brain structure (Brown et al., 2013; Varma et al., 2015, 2016). As above, we began with simple Pearson correlations between estimated $\dot{V}O_{2max}$ and ICV-adjusted volumes, applying FDR corrections for multiple comparisons for sex and laterality when appropriate (Table 3). Our results did not demonstrate a significant correlation or trend toward significance in the correlation between estimated $\dot{V}O_{2max}$ and left DG/CA3 head volume in men or in women (men: $N = 13$, $r = -.391$, $t[11] = -1.408$, $p = .187$ (p_{adjust} for sex and laterality = .231); women: $N = 22$, $r = .266$, $t[20] = 1.236$, $p = .231$ (p_{adjust} for sex and laterality = .231)). The correlation between estimated $\dot{V}O_{2max}$ and bilateral DG/CA3 head volume was also nonsignificant for both sexes (men: $N = 13$, $r = -.473$, $t[11] = -1.780$, $p = .103$ (p_{adjust} for sex = .194); women: $N = 22$, $r = .288$, $t[20] = 1.345$, $p = .194$ (p_{adjust} for sex = .194)). To follow, we tested for sex-specific associations between cardiorespiratory fitness and the CA1 and subiculum subfields, using the approach above. Additionally, the same outliers were removed as in the CA1 and subiculum correlational analyses aforementioned. There was no significant correlation between estimated $\dot{V}O_{2max}$ and bilateral CA1 volume in men or in women (men: $N = 12$, $r = .018$, $t[10] = 0.056$, $p = .957$ (p_{adjust} for sex = .957); women: $N = 22$, $r = .220$, $t[20] = 1.011$, $p = .324$ (p_{adjust} for sex = .648)). However, there was a significant correlation between estimated $\dot{V}O_{2max}$ and bilateral subiculum volume in women only (men: $N = 13$, $r = -.069$, $t[11] = -0.229$, $p = .823$ (p_{adjust} for

sex = .823); women: $N = 21$, $r = .608$, $t[19] = 3.340$, $p = .003^{**}$ (p_{adjust} for sex = .007**)).

Next, we examined significant correlations when controlling for age, education, and study using linear regression analyses. Given that the only statistically significant correlation was between estimated $\dot{V}O_{2max}$ and bilateral subiculum volume in women, we exclusively examined this relationship using linear regression analysis. Estimated $\dot{V}O_{2max}$ significantly predicted ICV-adjusted bilateral subiculum volume in women when controlling for age, education, and study ($N = 21$, estimated $\dot{V}O_{2max}$ $t[16] = 2.645$, $p = .018^*$, delta-$R^2 = 0.232$; model $F_{4,16} = 3.548$, adjusted $R^2 = 0.338$; Figure 4a). Moreover, this finding remained significant when DRS-2 memory raw score was included as a covariate ($N = 21$, estimated $\dot{V}O_{2max}$ $t[15] = 2.371$, $p = .032^*$, delta-$R^2 = 0.194$; DRS-2 memory raw score $t[15] = -0.583$, $p = .569$, delta-$R^2 = 0.012$; model $F_{5,15} = 2.789$, adjusted $R^2 = 0.309$).

To investigate if the observed association between estimated $\dot{V}O_{2max}$ and ICV-adjusted subiculum volume was statistically driven by an interaction between estimated $\dot{V}O_{2max}$ and sex, we performed a linear regression analysis with all participants (both men and women, minus the single subiculum outlier aforementioned) to investigate this interaction when controlling for age, education, and study. The results from this analysis suggest that the association between subiculum volume and the interaction between estimated $\dot{V}O_{2max}$ and sex was trending toward significance ($N = 34$, estimated $\dot{V}O_{2max}$:Sex $t[27] = 1.910$, $p = .067^\dagger$, delta-$R^2 = 0.089$; model $F_{6,27} = 2.367$, adjusted $R^2 = 0.199$; Figure 4b). In this linear regression, estimated $\dot{V}O_{2max}$ did not predict subiculum volume (estimated $\dot{V}O_{2max}$ $t[27] = -0.604$, $p = .551$). These results suggest that the relationship between cardiorespiratory fitness and subiculum volume may be sex-dependent.
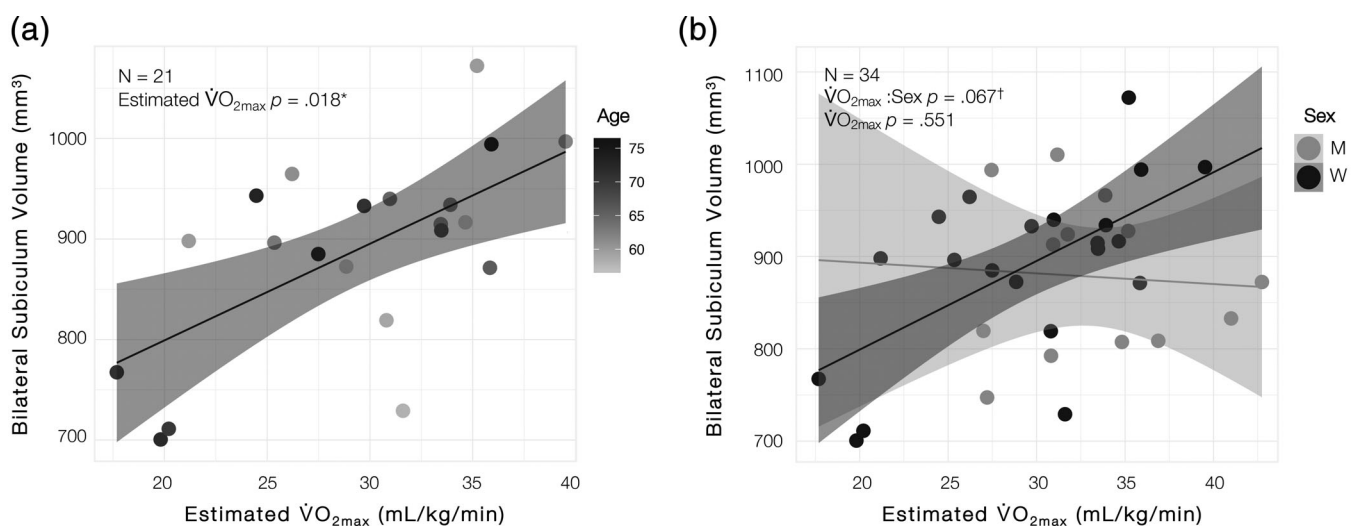

FIGURE 4 Cardiorespiratory fitness, as measured by $\dot{V}O_{2max}$, significantly predicted bilateral ICV-corrected subiculum volume in women. (a) $\dot{V}O_{2max}$ significantly predicted bilateral subiculum volume in women ($N = 21$, estimated $\dot{V}O_{2max}$ $t[16] = 2.645$, $p = .018^*$, delta-$R^2 = 0.232$; model $F_{4,16} = 3.548$, adjusted $R^2 = 0.338$). (b) The association between subiculum volume and the interaction between estimated $\dot{V}O_{2max}$ and sex was trending toward significance ($N = 34$, $\dot{V}O_{2max}$: Sex $t[27] = 1.910$, $p = .067^\dagger$, delta-$R^2 = 0.089$; model $F_{6,27} = 2.367$, adjusted $R^2 = 0.199$). Estimated $\dot{V}O_{2max}$ did not significantly predict subiculum volume in this model (estimated $\dot{V}O_{2max}$ $t[27] = -0.604$, $p = .551$)

4 | DISCUSSION

This study sought to examine cardiorespiratory fitness as a modulator of hippocampal subfield structure and function in cognitive aging. We investigated cardiorespiratory fitness as a predictor of markers of hippocampal subfield integrity, including hippocampal subfield volumes and performance on a hippocampally dependent MDT. This MDT was designed as a putative measure of behavioral pattern separation and as such, was designed to engage the DG/CA3 circuit. Our results show that, in line with the previously published literature and our initial prediction, our older adult sample demonstrated the lowest MDT accuracy in the condition with the highest stimuli similarity. Left DG/CA3 body volume significantly predicted performance in the highest similarity condition. However, contrary to our initial prediction based on our findings in young adults (Nauer, Dunne, et al., 2020), there was not a significant relationship between cardiorespiratory fitness and DG/CA3 volume. Instead, cardiorespiratory fitness significantly predicted bilateral subiculum volume in women, not men. These data provide continued support for the role of the DG/CA3 subfield in mnemonic discrimination in cognitive aging, and furthermore, suggest that lifestyle factors such as fitness may have differential effects on the hippocampal subfields in men and women, and considered alongside previous work, possibly also in young and older adults.

4.1 | Older adults demonstrate lowest accuracy in the MDT condition with the highest stimuli similarity; performance in this condition is predicted by left DG/CA3 body volume

A consistently growing body of literature indicates that aging is associated with significant impairments on tasks designed to putatively measure behavioral pattern separation (Holden, Toner, Pirogovsky, Kirwan, & Gilbert, 2013; Nauer, Schon, & Stern, 2020; Stark et al., 2013; Stark & Stark, 2017; Toner, Pirogovsky, Kirwan, & Gilbert, 2009; Yassa et al., 2011; Yassa & Stark, 2011). To build upon these studies, our older adult participants completed an MDT designed to behaviorally tax pattern separation at parametrically varying levels of stimuli similarity. The task was designed such that mnemonic interference and thus, pattern separation requirement increased as stimuli similarity increased. This design is similar to that of other pattern separation tasks in rodents (Clelland et al., 2009; Creer et al., 2010) and putative behavioral pattern separation tasks in humans (Lacy, Yassa, Stark, Muftuler, & Stark, 2010; Motley & Kirwan, 2012; Yassa et al., 2011; Yassa, Mattfeld, Stark, & Stark, 2011). Our sample of cognitively intact older adults demonstrated significantly lower accuracy in the condition with the greatest mnemonic interference when compared to the two other conditions with lesser levels of mnemonic interference. Identifying an increasing deficit in mnemonic discrimination performance as interference levels are elevated, but better performance as interference levels are minimized, provides tentative

indication that the deficit in mnemonic discrimination may arise from underlying impairment in hippocampal computation, namely, pattern separation (Leal & Yassa, 2018). Thus, this finding builds upon the extant literature to further suggest that the impaired mnemonic discrimination observed in older adult humans may be due to an effect of aging on the neural substrate responsible for pattern separation.

It is worth noting that there were a handful of participants who performed below chance in the highest similarity condition of the MDT. This would indicate that the rate at which the participant false alarmed (nonmatch trial incorrectly called a "match") was greater than the rate at which the participant hit (match trial correctly called a "match"). This is notable, as the overlapping characteristics of the nonmatch trial stimuli were designed to putatively tax pattern separation-related mechanisms. Thus, false alarming may be indicative of erroneous pattern completion (Kirwan & Stark, 2007; Stark et al., 2013). As previously noted, the observed trends in participant performance on the MDT suggest that the observed behavioral deficit may arise from underlying impairments in pattern separation-related mechanisms. Conceivably, corrected accuracy rates below 0% may be representative of functional impairments in the pattern separation system that result in erroneous pattern completion, and consequently, high false alarm rates. Future studies should consider incorporating functional neuroimaging of the hippocampal subfields to examine if behavioral impairments are associated with DG/CA3 functional changes, which would provide further evidence for an age-related effect on the neural substrate responsible for pattern separation.

Second, our results build upon the current literature that suggests the DG/CA3 circuit is involved in pattern separation in humans. Although it was computational models that originally proposed DG/CA3 circuitry as the underlying neural correlate of pattern separation (Marr, 1971; O'Reilly & McClelland, 1994; Treves & Rolls, 1994), human studies have since provided functional and structural evidence to support this hypothesis. A seminal fMRI study by Bakker, Kirwan, Miller, and Stark (2008) provided evidence that the DG/CA3 subregion uniquely demonstrates BOLD signal activity consistent with pattern separation. This finding has since been corroborated in multiple fMRI studies designed to examine hippocampal subfield activation during behavioral pattern separation paradigms (Berron et al., 2016; Kyle, Stokes, Lieberman, Hassan, & Ekstrom, 2015; Lacy et al., 2010). Alongside the functional evidence suggesting a role for the DG/CA3 circuit in pattern separation in humans, our results indicate that the structural integrity of these subregions may modulate mnemonic discrimination task performance. In the current study, left DG/CA3 body volume significantly predicted MDT corrected accuracy in the condition with the highest stimuli similarity. No other hippocampal subfield volume significantly predicted corrected accuracy in any similarity condition of the task. These results fit well with a previous study in older adults that reported greater volume of the left DG/CA3 was uniquely associated with better ability to discriminate similar mnemonic representations, whereas volumes of other hippocampal subfields were not related to mnemonic discrimination ability (Doxey & Kirwan, 2015). Thus, the results of the current study build upon the

extant literature to provide further structural evidence for the role of the DG/CA3 circuit in pattern separation in humans.

4.2 | Observed relationship between left DG/CA3 body volume and mnemonic discrimination performance

In the current study, the relationship between DG/CA3 volume and mnemonic discrimination performance was specific to the DG/CA3 body. The specificity of the observed relationship to the posterior hippocampus builds upon the existing literature which suggests that the longitudinal axis of the hippocampus is organized along a gradient, with the posterior component (dorsal in rodents) supporting cognitive functions such as spatial memory and the anterior component (ventral in rodents) supporting emotional responses (Bannerman et al., 2004; Fanselow & Dong, 2010; Strange et al., 2014). The literature also indicates the presence of a concurrent gradient associated with the level of detail in the neural representation of episodic memories. Studies suggest that the subfield compositions of the posterior and anterior hippocampus bias these regions toward pattern separation and pattern completion, respectively (Poppenk, Evensmoen, Moscovitch, & Nadel, 2013). Altogether, this literature suggesting a role for the posterior hippocampus in cognitive function and more specifically, pattern separation, provides a firm foundation for our result suggesting a role for the DG/CA3 body in predicting mnemonic discrimination performance when stimuli similarity is high.

Second, it is worth noting that the identified relationship between DG/CA3 body volume and MDT performance was observed only in the left DG/CA3 body. The differential recruitment of the left and right hippocampus has also been examined in the mnemonic discrimination fMRI literature. Bakker et al. (2008) and Lacy et al. (2010) provide no evidence for lateralization, with their results indicating that the bilateral DG/CA3 is involved in object mnemonic discrimination. Similarly, the ultra-high-resolution fMRI study by Berron et al. (2016) found no hemispheric differences in their results, suggesting that bilateral DG BOLD signal discriminates among highly similar scenes. Comparatively, Kyle et al. (2015) did demonstrate lateralized results, showing that left DG/CA3 BOLD signal was uniquely related to accurate discrimination of spatial distances. At a whole-hippocampal level, some of the early functional neuroimaging studies suggest that the hippocampi are lateralized based on the quality of the information (i.e., verbal or nonverbal, spatial) relevant to the memory encoding task at hand (Kelley et al., 1998; Motley & Kirwan, 2012; Powell et al., 2005). Thus, it is possible that hippocampal lateralization patterns during mnemonic discrimination are dependent upon the characteristics of the stimuli, such that tasks involving object, spatial, and/or emotional stimuli may differentially recruit the left and/or right hippocampus. Future neuroimaging studies should include multiple modalities of mnemonic discrimination tasks and multiple modalities of imaging (functional and structural) to examine how the discrimination of different types of information may be differentially related to left and/or right DG/CA3 function and structure.

4.3 | Associations between cardiorespiratory fitness and mnemonic discrimination task performance in older adults are unclear

Given the relationship between wheel running and enhanced pattern separation in rodents (Creer et al., 2010), as well as the relationship between increased cardiorespiratory fitness and improved mnemonic discrimination in young adults (Déry et al., 2013; Nauer, Dunne, et al., 2020), we predicted that greater cardiorespiratory fitness would be significantly associated with greater MDT accuracy in the condition with the highest stimuli similarity. On the contrary, our results did not support this prediction. Although it is possible that cardiorespiratory fitness does not predict performance on this task in cognitively intact older adults, these null results may be due to a number of alternative explanations. First, both Nauer, Dunne, et al. (2020) and Déry et al. (2013) manipulated cardiorespiratory fitness through exposure to aerobic exercise training. It is possible that this longitudinal experimental design, including the manipulation of cardiorespiratory fitness, is more sensitive to detecting existing relationships between fitness and mnemonic discrimination. This could in part be attributed to the fact that longitudinal study designs control for interindividual variability, which is critical given the influence of individual genetic differences on cardiorespiratory fitness (Barbato et al., 1998; Teran-Garcia, Rankinen, & Bouchard, 2008).

Second, it is possible that the link between cardiorespiratory fitness and mnemonic discrimination is domain-specific. A study from our laboratory recently demonstrated that although aging is associated with impaired mnemonic discrimination in both spatial and nonspatial domains, greater cardiorespiratory fitness was associated with an attenuated age-related decline in spatial contextual discrimination only (Nauer, Schon, & Stern, 2020). It has been established that the hippocampus is involved in two neocortical systems for memory: the posterior medial system, in which contextual information is processed through the parahippocampal cortex and medial entorhinal cortex, then converges upon the hippocampus, and the anterior temporal system, in which item information is processed through the perirhinal cortex and lateral entorhinal cortex, then converges upon the hippocampus (Hunsaker, Chen, Tran, & Kesner, 2013; for reviews see Ranganath & Ritchey, 2012; Ritchey, Libby, & Ranganath, 2015). It is possible that cardiorespiratory fitness modulates mnemonic discrimination through a mechanism unique to the structure and/or connectivity of brain regions in the posterior medial system, allowing for effects isolated to tasks requiring contextual or spatial discrimination.

Finally, it is also possible that the small sample size for this analysis ($N = 29$) and/or the potential error in predicting $\dot{V}O_{2max}$ from heart rate and work rate rather than measuring it directly from gas exchange played a role in our null results for this prediction. Considered together, future work should bear in mind the above alternative explanations and possible limitations in developing their experimental design, including consideration of longitudinal study design to control for interindividual variability, spatial and nonspatial mnemonic discrimination tasks to examine differential effects of cardiorespiratory fitness on mnemonic discrimination domains, and finally, adequate

sample size and direct measurement of $\dot{V}O_{2max}$ when it is safe for the study sample.

4.4 | Cardiorespiratory fitness is associated with subiculum volume, but not DG/CA3 volume, in a manner that may be sex-dependent

One of the main goals of the current study was to extend findings from studies of young adults that have examined relationships between aerobic exercise, cardiorespiratory fitness, and hippocampal subfield structure to older adults. We predicted that cardiorespiratory fitness would be positively associated with DG/CA3 volume, as evidenced in a study of young adults (Nauer, Dunne, et al., 2020), and may also demonstrate positive effects on other hippocampal subfields, including the CA1 and subiculum, as evidenced in prior studies of older adults (Rosano et al., 2017; Varma et al., 2016). Since the seminal rodent experiments which showed that wheel running leads to upregulation of adult hippocampal neurogenesis in the DG (van Praag, Christie, et al., 1999; van Praag, Kempermann, & Gage, 1999), a small number of human studies have demonstrated relationships between exercise, cardiorespiratory fitness, and DG plasticity in young adults (Nauer, Dunne, et al., 2020; Pereira et al., 2007). Our findings, considered alongside other studies in older adult populations, suggest that there may be independent and/or related mechanisms that underlie aerobic exercise- and fitness-related plasticity in the aging brain. In our sample of older adults, cardiorespiratory fitness was not associated with DG/CA3 volume as predicted. Instead, cardiorespiratory fitness was associated with bilateral subiculum volume in women, not men. This result is complementary to that published by Varma et al. (2016), who found that objectively measured daily walking activity was associated with larger subiculum surface area in women, not men. Considered together, these findings suggest that both cardiorespiratory fitness, which is linked to moderate-to-vigorous intensity exercise (Pescatello & American College of Sports Medicine, 2014), and physical activity, which is linked to low-to-vigorous intensity exercise (Howley, 2001), may be associated with enhanced subicular plasticity in older adult women. Given that the current study did not include objective measures of daily physical activity, we cannot make claims about relationships between cardiorespiratory fitness and daily physical activity or how these lifestyle factors may together or separately modulate subiculum volume. Future studies should consider including measures of both cardiorespiratory fitness and objective measures of daily physical activity to disentangle how these two variables may independently and cooperatively influence hippocampal integrity.

Furthermore, our finding suggests a relationship between cardiorespiratory fitness and subiculum volume in older adult women provides important insight for future studies seeking to examine relationships between cardiorespiratory fitness, brain function, and brain structure. The subiculum is relatively understudied compared to other hippocampal subfields (Ledergerber & Moser, 2017; O'Mara, Sanchez-Vives, Brotons-Mas, & O'Hare, 2009; Roy et al., 2017). However,

a growing body of literature suggests a role for the subiculum in both spatial navigation and episodic memory retrieval. Addressing the former, in rodents, subiculum lesions are associated with impaired place navigation (Morris, Schenk, Tweedie, & Jarrard, 1990). Additionally, electrophysiological recordings from subiculum neurons suggest that these cells' firing patterns represent specific positions or boundaries in the environment, which together, may work to flexibly support navigation (Brotons-Mas, Schaffelhofer, Guger, O'Mara, & Sanchez-Vives, 2017). Given this proposed role of the subiculum in spatial navigation, it stands to reason that if cardiorespiratory fitness modulates subicular integrity, cardiorespiratory fitness may also modulate performance on spatial navigation tasks. Importantly, such a relationship has been demonstrated, as a recent behavioral study indicated that greater cardiorespiratory fitness was associated with better performance on a spatial route disambiguation task in older adults (Nauer, Schon, & Stern, 2020). Thus, future studies should examine the relationships between cardiorespiratory fitness, subiculum volume, and spatial navigation to determine if subicular integrity may modulate observed relationships between cardiorespiratory fitness and spatial navigation task performance.

Second, the subiculum also plays an important role in episodic memory retrieval. This is in part evidenced by human fMRI studies suggesting that the subiculum is selectively active during the retrieval phase of episodic memory tasks (Eldridge, 2005; Gabrieli, 1997). In addition, rodent work using a combined approach of circuit tracing and optogenetic manipulation demonstrated that dorsal subiculum output is critical for memory retrieval and for the regulation of stress hormone response induced by memory retrieval (Roy et al., 2017). The latter finding highlights the role of the subiculum in regulating the body's stress response by the modulation of the hypothalamic–pituitary–adrenal (HPA) axis (Herman & Mueller, 2006; Roy et al., 2017). This relationship between the subiculum and HPA axis is critical given that the dysregulation of the HPA system in humans is associated with impairments in learning and memory (Lupien et al., 1998, 2002; Young, Sahakian, Robbins, & Cowen, 1999). Altogether, this literature suggests that the subiculum, by means of its role in memory retrieval and memory retrieval-induced stress hormone response, may provide a possible pathway for lifestyle factors and therapeutic interventions to affect learning and memory. Future studies should examine markers of subicular plasticity and stress hormone levels as mediators of the relationship between cardiorespiratory fitness and episodic memory performance.

Importantly, the association between cardiorespiratory fitness and subiculum volume was observed in women, not men. Other human studies have also found sex-specific relationships between exercise-related measures and hippocampal plasticity. Varma et al. (2015) found that greater low-intensity walking activity was associated with larger hippocampal volume in older adult women, not men. Later work by the same group then demonstrated that greater daily walking activity was associated with greater surface area of the subiculum, also only in older adult women (Varma et al., 2016). Although the underlying mechanism responsible for this sex-specificity is unclear, it is possible that sex differences in steroid hormones and/or in neurotrophin levels may play a role.

Speaking to the former, sex steroid hormones are known to have neuro-protective effects on the brain (for reviews see Pike, Carroll, Rosario, & Barron, 2009; Siddiqui et al., 2016). Women demonstrate a significant decrease in reproductive hormones after menopause; thus, it is possible that cardiorespiratory fitness may upregulate these low basal levels of sex hormones which in turn, could result in a revived neuroprotective effect on the brain in women (Brown et al., 2013). Second, it is also possible that sex differences in neurotrophin levels play a role in the sex-specific association observed between cardiorespiratory fitness and subiculum volume. In rodents, studies suggest that exposure to an enriched environment that includes a running wheel results in sex-dependent changes in levels of neurotrophins, such as brain-derived neurotrophic factor and nerve growth factor, in various regions of the brain (Bakos et al., 2009; Zhu et al., 2006). Interactive effects of sex steroid hormones and neurotrophins may lead to sex-specific changes in neuroplasticity of the implicated brain region (Chan & Ye, 2017). Further research is required to determine how sex steroid hormones and sex-dependent differences in neurotrophin levels may modulate sex-dependent associations between cardiorespiratory fitness and brain structure and function.

4.5 | Limitations and considerations

While discussing the implications of our results, it is important to acknowledge the limitations of the current study. First, the descriptive statistics for our study sample suggested the presence of a selection bias for both highly educated and highly fit individuals. Given that higher educational attainment has protective effects on cognition and brain health during aging (Rentz et al., 2009; for reviews see Farah, 2017; Fratiglioni & Wang, 2007), we included education as a covariate in our analyses. However, it is still possible that the highly educated nature of our study sample influenced our observed relationships between mnemonic discrimination, hippocampal subfield volumes, and cardiorespiratory fitness. In addition to a selection bias for highly educated participants, the descriptive statistics for our study sample indicated the presence of a selection bias for highly fit participants. Perceptibly, our participants demonstrated higher than average (50th percentile) fitness compared to age- and sex-specific national normative values. It is possible that the skewed nature of the data may have affected the observed relationships between cardiorespiratory fitness, mnemonic discrimination, and hippocampal subfield volumes, especially given that lower-fit individuals may receive greater cognitive benefit with increased cardiorespiratory fitness (Nauer, Dunne, et al., 2020). Future studies should endeavor to include participants with a wider range of education and fitness levels to examine the relationships between cardiorespiratory fitness and brain structure and function within a more representative sample of older adults.

Other points to acknowledge are the limited number of participants included in the study, the potential error in predicting $\dot{V}O_{2max}$ from heart rate and work rate, and the cross-sectional nature of the current study's experimental design. Regarding sample size, recruiting

a greater number of participants may have increased our statistical power, and therefore, may have provided a greater likelihood of identifying statistically significant relationships between our variables of interest. Regarding predicting $\dot{V}O_{2max}$, although the submaximal treadmill test protocol was selected specifically to assure safety and accuracy in older adult participants (Cooper & Storer, 2004; Pescatello & American College of Sports Medicine, 2014), there is potential error associated with predicting $\dot{V}O_{2max}$ from heart rate and work rate rather than measuring it directly from gas exchange. Using a maximal treadmill test protocol to directly measure $\dot{V}O_{2max}$ would lower estimation error but has to be weighed against participant safety. Finally, the cross-sectional design of the current study renders us unable to make causational claims regarding how the manipulation of cardiorespiratory fitness may affect brain function and structure in older adults. Altogether, future studies should consider these factors while developing their experimental designs, namely, the use of adequate sample sizes, direct measurement of $\dot{V}O_{2max}$ when it is safe for the study sample, and manipulation of cardiorespiratory fitness through aerobic exercise training.

5 | CONCLUSION

In the current study, we demonstrated that older adults exhibited the lowest MDT accuracy in the condition with the highest stimuli similarity and that left DG/CA3 body volume significantly predicted performance in this high-similarity condition. Furthermore, cardiorespiratory fitness was not associated with DG/CA3 volume as predicted, but instead, was associated with bilateral subiculum volume in women. Through these results, we build upon the extant literature that indicates that aging is associated with decreased mnemonic discrimination ability, and that the DG/CA3 circuit is a constituent of the neural substrate responsible for pattern separation-like task performance in humans. Furthermore, through these results, we offer a unique contribution to the literature by demonstrating that greater cardiorespiratory fitness is associated with greater bilateral subiculum volume in older adult women. This finding is of importance given that it provides additional evidence that hippocampal subfields other than the DG/CA3 may be modulated by cardiorespiratory fitness in older adults. Given the proposed role of the subiculum in both spatial navigation and episodic memory retrieval, future studies should examine subicular integrity as a modulator of the relationships between cardiorespiratory fitness or aerobic exercise training and performance on cognitive tasks involving spatial navigation or episodic memory. Finally, this finding indicates that future studies should examine how physiological and neurobiological sex differences, such as sex steroid hormones and sex-dependent differences in neurotrophin levels, may modulate associations between cardiorespiratory fitness and brain structure and function. Altogether, these findings further substantiate evidence for a role of the DG/CA3 circuit in behavioral pattern separation in humans and suggest that cardiorespiratory fitness may have differential effects on hippocampal subfield integrity in older adult men and women.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any interest or relationship, financial or otherwise, that may be perceived as influencing an author's objectivity.

ETHICS STATEMENT

All study procedures followed the guidelines set by the Code of Ethics of the World Medical Association and were approved by the Boston University Medical Campus Institutional Review Board.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

PATIENT CONSENT STATEMENT

All participants provided signed, informed consent prior to participation in the study.

ORCID

Kathryn L. Kern https://orcid.org/0000-0001-7382-8472

REFERENCES

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science (New York, N.Y.)*, *319*(5870), 1640–1642. https://doi.org/10.1126/science.1152882

Bakos, J., Hlavacova, N., Rajman, M., Ondicova, K., Koros, C., Kitraki, E., … Jezova, D. (2009). Enriched environment influences hormonal status and hippocampal brain derived neurotrophic factor in a sex dependent manner. *Neuroscience*, *164*(2), 788–797. http://dx.doi.org/10.1016/j.neuroscience.2009.08.054

Bannerman, D. M., Rawlins, J. N. P., McHugh, S. B., Deacon, R. M. J., Yee, B. K., Bast, T., … Feldon, J. (2004). Regional dissociations within the hippocampus—Memory and anxiety. *Neuroscience & Biobehavioral Reviews*, *28*(3), 273–283. https://doi.org/10.1016/j.neubiorev.2004.03.004

Barbato, J. C., Koch, L. G., Darvish, A., Cicila, G. T., Metting, P. J., & Britton, S. L. (1998). Spectrum of aerobic endurance running performance in eleven inbred strains of rats. *Journal of Applied Physiology*, *85*(2), 530–536. https://doi.org/10.1152/jappl.1998.85.2.530

Beck, A. T., & Steer, R. A. (1993). *Beck anxiety inventory manual*. New York, NY: Psychological Corporation.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*(1), 289–300 JSTOR.

Berron, D., Schutze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H. J., Kumaran, D., & Duzel, E. (2016). Strong evidence for pattern separation in human dentate gyrus. *Journal of Neuroscience*, *36*(29), 7569–7579. https://doi.org/10.1523/JNEUROSCI.0518-16.2016

Boldrini, M., Fulmore, C. A., Tartt, A. N., Simeon, L. R., Pavlova, I., Poposka, V., … Mann, J. J. (2018). Human hippocampal neurogenesis persists throughout aging. *Cell Stem Cell*, *22*(4), 589–599.e5. https://doi.org/10.1016/j.stem.2018.03.015

Brotons-Mas, J. R., Schaffelhofer, S., Guger, C., O'Mara, S. M., & Sanchez-Vives, M. V. (2017). Heterogeneous spatial representation by different subpopulations of neurons in the subiculum. *Neuroscience*, *343*, 174–189. https://doi.org/10.1016/j.neuroscience.2016.11.042

Brown, B. M., Peiffer, J. J., & Martins, R. N. (2013). Multiple effects of physical activity on molecular and cognitive signs of brain aging: Can exercise slow neurodegeneration and delay Alzheimer's disease? *Molecular Psychiatry*, *18*(8), 864–874. https://doi.org/10.1038/mp.2012.162

Chan, C. B., & Ye, K. (2017). Sex differences in brain-derived neurotrophic factor signaling and functions. *Journal of Neuroscience Research*, *95*(1–2), 328–335. http://dx.doi.org/10.1002/jnr.23863

Clark, P. J., Brzezinska, W. J., Puchalski, E. K., Krone, D. A., & Rhodes, J. S. (2009). Functional analysis of neurovascular adaptations to exercise in the dentate gyrus of young adult mice associated with cognitive gain. *Hippocampus*, *19*(10), 937–950. https://doi.org/10.1002/hipo.20543

Clelland, C. D., Choi, M., Romberg, C., Clemenson, G. D., Fragniere, A., Tyers, P., … Bussey, T. J. (2009). A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science (New York, NY)*, *325*(5937), 210–213. https://doi.org/10.1126/science.1173215

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. https://doi.org/10.2307/2136404

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science*, *14*(2), 125–130. https://doi.org/10.1111/1467-9280.t01-1-01430

Cooper, C. B., & Storer, T. W. (2004). *Exercise testing and interpretation. A practical approach*. Cambridge, MA: Cambridge University Press.

Creer, D. J., Romberg, C., Saksida, L. M., van Praag, H., & Bussey, T. J. (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. https://doi.org/10.1073/pnas.0911725107

Déry, N., Pilgrim, M., Gibala, M., Gillen, J., Wojtowicz, J. M., Macqueen, G., & Becker, S. (2013). Adult hippocampal neurogenesis reduces memory interference in humans: Opposing effects of aerobic exercise and depression. *Frontiers in Neuroscience*, *7*, 66. https://doi.org/10.3389/fnins.2013.00066

Doxey, C. R., & Kirwan, C. B. (2015). Structural and functional correlates of behavioral pattern separation in the hippocampus and medial temporal lobe. *Hippocampus*, *25*(4), 524–533. https://doi.org/10.1002/hipo.22389

Eadie, B. D., Redila, V. A., & Christie, B. R. (2005). Voluntary exercise alters the cytoarchitecture of the adult dentate gyrus by increasing cellular proliferation, dendritic complexity, and spine density. *The Journal of Comparative Neurology*, 486(1), 39–47. https://doi.org/10.1002/cne.20493

Eldridge, L. L. (2005). A dissociation of encoding and retrieval processes in the human hippocampus. *Journal of Neuroscience*, 25(13), 3280–3286. https://doi.org/10.1523/JNEUROSCI.3420-04.2005

Erickson, C. (2003). The neurobiology of memory changes in normal aging. *Experimental Gerontology*, 38(1–2), 61–69. https://doi.org/10.1016/S0531-5565(02)00160-2

Erickson, K. I., Prakash, R. S., Voss, M. W., Chaddock, L., Hu, L., Morris, K. S., ... Kramer, A. F. (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus*, 19(10), 1030–1039. https://doi.org/10.1002/hipo.20547

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

Eriksson, P. S., Perfilieva, E., Björk-Eriksson, T., Alborn, A.-M., Nordborg, C., Peterson, D. A., & Gage, F. H. (1998). Neurogenesis in the adult human hippocampus. *Nature Medicine*, 4(11), 1313–1317. https://doi.org/10.1038/3305

Fanselow, M. S., & Dong, H.-W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron*, 65(1), 7–19. https://doi.org/10.1016/j.neuron.2009.11.031

Farah, M. J. (2017). The neuroscience of socioeconomic status: Correlates, causes, and consequences. *Neuron*, 96(1), 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

Firth, J., Stubbs, B., Vancampfort, D., Schuch, F., Lagopoulos, J., Rosenbaum, S., & Ward, P. B. (2018). Effect of aerobic exercise on hippocampal volume in humans: A systematic review and meta-analysis. *NeuroImage*, 166, 230–238. https://doi.org/10.1016/j.neuroimage.2017.11.007

Fratiglioni, L., & Wang, H.-X. (2007). Brain reserve hypothesis in dementia. *Journal of Alzheimer's Disease*, 12(1), 11–22. https://doi.org/10.3233/JAD-2007-12103

Gabrieli, J. D. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science*, 276(5310), 264–266. https://doi.org/10.1126/science.276.5310.264

Gazzaley, A. H., Siegel, S. J., Kordower, J. H., Mufson, E. J., & Morrison, J. H. (1996). Circuit-specific alterations of N-methyl-D-aspartate receptor subunit 1 in the dentate gyrus of aged monkeys. *Proceedings of the National Academy of Sciences*, 93(7), 3121–3125. https://doi.org/10.1073/pnas.93.7.3121

Gould, E., Reeves, A. J., Fallah, M., Tanapat, P., Gross, C. G., & Fuchs, E. (1999). Hippocampal neurogenesis in adult Old World primates. *Proceedings of the National Academy of Sciences*, 96(9), 5263–5267. https://doi.org/10.1073/pnas.96.9.5263

Hagberg, J. M., Graves, J. E., Limacher, M., Woods, D. R., Leggett, S. H., Cononie, C., ... Pollock, M. L. (1989). Cardiovascular responses of 70- to 79-yr-old men and women to exercise training. *Journal of Applied Physiology (Bethesda, MD: 1985)*, 66(6), 2589–2594. https://doi.org/10.1152/jappl.1989.66.6.2589

Head, D., Rodrigue, K. M., Kennedy, K. M., & Raz, N. (2008). Neuroanatomical and cognitive mediators of age-related differences in episodic memory. *Neuropsychology*, 22(4), 491–507. https://doi.org/10.1037/0894-4105.22.4.491

Herman, J., & Mueller, N. (2006). Role of the ventral subiculum in stress integration. *Behavioural Brain Research*, 174(2), 215–224. https://doi.org/10.1016/j.bbr.2006.05.035

Holden, H. M., Toner, C., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2013). Visual object pattern separation varies in older adults. *Learning & Memory*, 20(7), 358–362. https://doi.org/10.1101/lm.030171.112

Howley, E. T. (2001). Type of activity: Resistance, aerobic and leisure versus occupational physical activity. *Medicine and Science in Sports and Exercise*, 33(6), S364–S369.

Hunsaker, M. R., Chen, V., Tran, G. T., & Kesner, R. P. (2013). The medial and lateral entorhinal cortex both contribute to contextual and item recognition memory: A test of the binding ofitems and context model. *Hippocampus*, 23(5), 380–391. https://doi.org/10.1002/hipo.22097

Jack, C. R., Petersen, R. C., Xu, Y. C., Waring, S. C., O'Brien, P. C., Tangalos, E. G., ... Kokmen, E. (1997). Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology*, 49(3), 786–794. https://doi.org/10.1212/WNL.49.3.786

Jonasson, L. S., Nyberg, L., Kramer, A. F., Lundquist, A., Riklund, K., & Boraxbekk, C.-J. (2017). Aerobic exercise intervention, cognitive performance, and brain structure: Results from the physical influences on brain in aging (PHIBRA) study. *Frontiers in Aging Neuroscience*, 8, 336. https://doi.org/10.3389/fnagi.2016.00336

Kaminsky, L. A., Arena, R., & Myers, J. (2015). Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing. *Mayo Clinic Proceedings*, 90(11), 1515–1523. https://doi.org/10.1016/j.mayocp.2015.07.026

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., ... Petersen, S. E. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron*, 20(5), 927–936. https://doi.org/10.1016/S0896-6273(00)80474-2

Kesner, R. P., Lee, I., & Gilbert, P. (2004). A behavioral assessment of hippocampal function based on a subregional analysis. *Reviews in the Neurosciences*, 15(5), 333–351. https://doi.org/10.1515/REVNEURO.2004.15.5.333

Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory*, 14(9), 625–633. https://doi.org/10.1101/lm.663507

Kohrt, W. M., Malley, M. T., Coggan, A. R., Spina, R. J., Ogawa, T., Ehsani, A. A., ... Holloszy, J. O. (1991). Effects of gender, age, and fitness level on response of VO_{2max} to training in 60-71 yr olds. *Journal of Applied Physiology (Bethesda, MD: 1985)*, 71(5), 2004–2011. https://doi.org/10.1152/jappl.1991.71.5.2004

Kronman, C. A., Kern, K. L., Nauer, R. K., Dunne, M. F., Storer, T. W., & Schon, K. (2020). Cardiorespiratory fitness predicts effective connectivity between the hippocampus and default mode network nodes in young adults. *Hippocampus*, 30(5), 526–541. https://doi.org/10.1002/hipo.23169

Kuhn, H. G., Dickinson-Anson, H., & Gage, F. H. (1996). Neurogenesis in the dentate gyrus of the adult rat: Age-related decrease of neuronal progenitor proliferation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 16(6), 2027–2033.

Kyle, C. T., Stokes, J. D., Lieberman, J. S., Hassan, A. S., & Ekstrom, A. D. (2015). Successful retrieval of competing spatial environments in humans involves hippocampal pattern separation mechanisms. *eLife*, 4, e10499. https://doi.org/10.7554/eLife.10499

Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. L. (2010). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory*, 18(1), 15–18. https://doi.org/10.1101/lm.1971111

Leal, S. L., & Yassa, M. A. (2018). Integrating new findings and examining clinical applications of pattern separation. *Nature Neuroscience*, 21(2), 163–173. https://doi.org/10.1038/s41593-017-0065-1

Ledergerber, D., & Moser, E. I. (2017). Memory retrieval: Taking the route via subiculum. *Current Biology*, 27(22), R1225–R1227. https://doi.org/10.1016/j.cub.2017.09.042

Lupien, S. J., de Leon, M., de Santi, S., Convit, A., Tarshish, C., Nair, N. P. V., ... Meaney, M. J. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience*, 1(1), 69–73. https://doi.org/10.1038/271

Lupien, S. J., Wilkinson, C. W., Brière, S., Ménard, C., Ng Ying Kin, N. M. K., & Nair, N. P. V. (2002). The modulatory effects of corticosteroids on cognition: Studies in young human populations. *Psychoneuroendocrinology*, 27(3), 401–416. https://doi.org/10.1016/S0306-4530(01)00061-0

Marlatt, M. W., Potter, M. C., Lucassen, P. J., & van Praag, H. (2012). Running throughout middle-age improves memory function, hippocampal neurogenesis, and BDNF levels in female C57BL/6J mice. *Developmental Neurobiology*, 72(6), 943–952. https://doi.org/10.1002/dneu.22009

Marr, D. (1971). Simple memory: A theory for Archicortex. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 262(841), 23–81. https://doi.org/10.1098/rstb.1971.0078

Mattis, S. (1976). Mental status examination for organic mental syndrome in the elderly patient. In L. Bellak & T. B. Karasu (Eds.), *Geriatric psychiatry* (pp. 77–121). Orlando, FL: Grune & Stratton.

Mitchell, J. H., Sproule, B. J., & Chapman, C. B. (1958). The physiological meaning of the maximal oxygen intake test. *Journal of Clinical Investigation*, 37(4), 538–547. https://doi.org/10.1172/JCI103636

Morris, R. G. M., Schenk, F., Tweedie, F., & Jarrard, L. E. (1990). Ibotenate lesions of hippocampus and/or subiculum: Dissociating components of Allocentric spatial learning. *European Journal of Neuroscience*, 2(12), 1016–1028. https://doi.org/10.1111/j.1460-9568.1990.tb00014.x

Motley, S. E., & Kirwan, C. B. (2012). A parametric investigation of pattern separation processes in the medial temporal lobe. *Journal of Neuroscience*, 32(38), 13076–13084. https://doi.org/10.1523/JNEUROSCI.5920-11.2012

Nauer, R. K., Dunne, M. F., Stern, C. E., Storer, T. W., & Schon, K. (2020). Improving fitness increases dentate gyrus/CA3 volume in the hippocampal head and enhances memory in young adults. *Hippocampus*, 30(5), 488–504. https://doi.org/10.1002/hipo.23166

Nauer, R. K., Schon, K., & Stern, C. E. (2020). Cardiorespiratory fitness and mnemonic discrimination across the adult lifespan. *Learning & Memory*, 27(3), 91–103. https://doi.org/10.1101/lm.049197.118

Ngwenya, L. B., Heyworth, N. C., Shwe, Y., Moore, T. L., & Rosene, D. L. (2015). Age-related changes in dentate gyrus cell numbers, neurogenesis, and associations with cognitive impairments in the rhesus monkey. *Frontiers in Systems Neuroscience*, 9, 102. https://doi.org/10.3389/fnsys.2015.00102

Nokia, M. S., Lensu, S., Ahtiainen, J. P., Johansson, P. P., Koch, L. G., Britton, S. L., & Kainulainen, H. (2016). Physical exercise increases adult hippocampal neurogenesis in male rats provided it is aerobic and sustained. *The Journal of Physiology*, 594(7), 1855–1873. https://doi.org/10.1113/JP271552

O'Mara, S. M., Sanchez-Vives, M. V., Brotons-Mas, J. R., & O'Hare, E. (2009). Roles for the subiculum in spatial information processing, memory, motivation and the temporal control of behaviour. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 33(5), 782–790. https://doi.org/10.1016/j.pnpbp.2009.03.040

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, 4(6), 661–682. https://doi.org/10.1002/hipo.450040605

Penner, M. R., Roth, T. L., Chawla, M. K., Hoang, L. T., Roth, E. D., Lubin, F. D., ... Barnes, C. A. (2011). Age-related changes in arc transcription and DNA methylation within the hippocampus. *Neurobiology of Aging*, 32(12), 2198–2210. https://doi.org/10.1016/j.neurobiolaging.2010.01.009

Pereira, A. C., Huddleston, D. E., Brickman, A. M., Sosunov, A. A., Hen, R., McKhann, G. M., ... Small, S. A. (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. https://doi.org/10.1073/pnas.0611721104

Pescatello, L. S., & American College of Sports Medicine (Eds.). (2014). *ACSM's guidelines for exercise testing and prescription* (9th ed.). Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins Health.

Pike, C. J., Carroll, J. C., Rosario, E. R., & Barron, A. M. (2009). Protective actions of sex steroid hormones in Alzheimer's disease. *Frontiers in Neuroendocrinology*, 30(2), 239–258. https://doi.org/10.1016/j.yfrne.2009.04.015

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, 17(5), 230–240. https://doi.org/10.1016/j.tics.2013.03.005

Powell, H. W. R., Koepp, M. J., Symms, M. R., Boulby, P. A., Salek-Haddadi, A., Thompson, P. J., ... Richardson, M. P. (2005). Material-specific lateralization of memory encoding in the medial temporal lobe: Blocked versus event-related design. *NeuroImage*, 27(1), 231–239. https://doi.org/10.1016/j.neuroimage.2005.04.033

Pruessner, J. C. (2000). Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: Minimizing the discrepancies between laboratories. *Cerebral Cortex*, 10(4), 433–442. https://doi.org/10.1093/cercor/10.4.433

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13(10), 713–726. https://doi.org/10.1038/nrn3338

Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M., & Acker, J. D. (2004). Differential aging of the medial temporal lobe: A study of a five-year change. *Neurology*, 62(3), 433–438. https://doi.org/10.1212/01.WNL.0000106466.09835.46

Raz, N., Gunning-Dixon, F., Head, D., Rodrigue, K. M., Williamson, A., & Acker, J. D. (2004). Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: Replicability of regional differences in volume. *Neurobiology of Aging*, 25(3), 377–396. https://doi.org/10.1016/S0197-4580(03)00118-0

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., ... Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex*, 15(11), 1676–1689. https://doi.org/10.1093/cercor/bhi044

Redila, V. A., & Christie, B. R. (2006). Exercise-induced changes in dendritic structure and complexity in the adult hippocampal dentate gyrus. *Neuroscience*, 137(4), 1299–1307. https://doi.org/10.1016/j.neuroscience.2005.10.050

Rentz, D. M., Locascio, J. J., Becker, J. A., Moran, E. K., Eng, E., Buckner, R. L., ... Johnson, K. A. (2009). Cognition, reserve and amyloid deposition in normal aging. *Annals of Neurology*, 67(3), 353–364. https://doi.org/10.1002/ana.21904

Ritchey, M., Libby, L. A., & Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition. In *Progress in brain research* (Vol. 219, pp. 45–64). Amsterdam: Elsevier. https://doi.org/10.1016/bs.pbr.2015.04.001

Rosano, C., Guralnik, J., Pahor, M., Glynn, N. W., Newman, A. B., Ibrahim, T. S., ... Aizenstein, H. J. (2017). Hippocampal response to a 24-month physical activity intervention in sedentary older adults. *The American Journal of Geriatric Psychiatry*, 25(3), 209–217. https://doi.org/10.1016/j.jagp.2016.11.007

Roy, D. S., Kitamura, T., Okuyama, T., Ogawa, S. K., Sun, C., Obata, Y., ... Tonegawa, S. (2017). Distinct neural circuits for the formation and retrieval of episodic memories. *Cell*, 170(5), 1000–1012.e19. https://doi.org/10.1016/j.cell.2017.07.013

Schaie, K. W. (2013). *Developmental influences on adult intelligence: The Seattle longitudinal study* (2nd ed.). Oxford: Oxford University Press.

Seo, S. (2006). *A review and comparison of methods for detecting outliers in univariate data sets*. Pittsburg, PA: University of Pittsburgh.

Siddiqui, A. N., Siddiqui, N., Khan, R. A., Kalam, A., Jabir, N. R., Kamal, M. A., ... Tabrez, S. (2016). Neuroprotective role of steroidal sex hormones: An overview. *CNS Neuroscience & Therapeutics*, 22(5), 342–350. https://doi.org/10.1111/cns.12538

Small, S. A., Chawla, M. K., Buonocore, M., Rapp, P. R., & Barnes, C. A. (2004). From the cover: Imaging correlates of brain function in monkeys and rats isolates a hippocampal subregion differentially

vulnerable to aging. *Proceedings of the National Academy of Sciences of the United States of America*, 101(18), 7181–7186. https://doi.org/10.1073/pnas.0400285101

Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P., & Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience*, 12(10), 585–601. https://doi.org/10.1038/nrn3085

Small, S. A., Stern, Y., Tang, M., & Mayeux, R. (1999). Selective decline in memory function among healthy elderly. *Neurology*, 52(7), 1392–1392, 1396. https://doi.org/10.1212/WNL.52.7.1392

Small, S. A., Tsai, W. Y., DeLaPaz, R., Mayeux, R., & Stern, Y. (2002). Imaging hippocampal function across the human life span: Is memory decline normal or not? *Annals of Neurology*, 51(3), 290–295. https://doi.org/10.1002/ana.10105

Sorrells, S. F., Paredes, M. F., Cebrian-Silla, A., Sandoval, K., Qi, D., Kelley, K. W., ... Alvarez-Buylla, A. (2018). Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature*, 555(7696), 377–381. https://doi.org/10.1038/nature25975

Spalding, K. L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H. B., ... Frisén, J. (2013). Dynamics of hippocampal neurogenesis in adult humans. *Cell*, 153(6), 1219–1227. https://doi.org/10.1016/j.cell.2013.05.002

Stark, S. M., & Stark, C. E. L. (2017). Age-related deficits in the mnemonic similarity task for objects and scenes. *Behavioural Brain Research*, 333, 109–117. https://doi.org/10.1016/j.bbr.2017.06.049

Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia*, 51(12), 2442–2449. https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, 15(10), 655–669. https://doi.org/10.1038/nrn3785

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford: Oxford University Press.

Suwabe, K., Byun, K., Hyodo, K., Reagh, Z. M., Roberts, J. M., 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*, 115(41), 10487–10492. https://doi.org/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. https://doi.org/10.1038/s41598-017-04850-y

Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, 37(1), 153–156. https://doi.org/10.1016/S0735-1097(00)01054-8

Teran-Garcia, M., Rankinen, T., & Bouchard, C. (2008). Genes, exercise, growth, and the sedentary, obese child. *Journal of Applied Physiology*, 105(3), 988–1001. https://doi.org/10.1152/japplphysiol.00070.2008

Thomas, A. G., Dennis, A., Rawlings, N. B., Stagg, C. J., 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. https://doi.org/10.1016/j.neuroimage.2015.10.090

Tomporowski, P. D. (2003). Effects of acute bouts of exercise on cognition. *Acta Psychologica*, 112(3), 297–324.

Toner, C. K., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2009). Visual object pattern separation deficits in nondemented older adults. *Learning & Memory*, 16(5), 338–342. https://doi.org/10.1101/lm.1315109

Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4(3), 374–391. https://doi.org/10.1002/hipo.450040319

van der Kouwe, A. J. W., Benner, T., Salat, D. H., & Fischl, B. (2008). Brain morphometry with multiecho MPRAGE. *NeuroImage*, 40(2), 559–569. https://doi.org/10.1016/j.neuroimage.2007.12.025

van Praag, H., Christie, B. R., Sejnowski, T. J., & Gage, F. H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proceedings of the National Academy of Sciences*, 96(23), 13427–13431. https://doi.org/10.1073/pnas.96.23.13427

van Praag, H., Kempermann, G., & Gage, F. H. (1999). Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. *Nature Neuroscience*, 2(3), 266–270. https://doi.org/10.1038/6368

van Praag, H., Shubert, T., Zhao, C., & Gage, F. H. (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. https://doi.org/10.1523/JNEUROSCI.1731-05.2005

Varma, V. R., Chuang, Y.-F., Harris, G. C., Tan, E. J., & Carlson, M. C. (2015). Low-intensity daily walking activity is associated with hippocampal volume in older adults: Daily walking activity and hippocampal volume. *Hippocampus*, 25(5), 605–615. https://doi.org/10.1002/hipo.22397

Varma, V. R., Tang, X., & Carlson, M. C. (2016). Hippocampal sub-regional shape and physical activity in older adults: Hippocampal shape and physical activity. *Hippocampus*, 26(8), 1051–1060. https://doi.org/10.1002/hipo.22586

Wagner, P. D. (1996). Determinants of maximal oxygen transport and utilization. *Annual Review of Physiology*, 58(1), 21–50. https://doi.org/10.1146/annurev.ph.58.030196.000321

Wasserman, K. (Ed.). (2012). *Principles of exercise testing and interpretation: Including pathophysiology and clinical applications* (5th ed.). Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2011). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*, 21(9), 968–979. https://doi.org/10.1002/hipo.20808

Yassa, M. A., Mattfeld, A. T., Stark, S. M., & Stark, C. E. L. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences*, 108(21), 8873–8878. https://doi.org/10.1073/pnas.1101567108

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, 34(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49. https://doi.org/10.1016/0022-3956(82)90033-4

Young, A. H., Sahakian, B. J., Robbins, T. W., & Cowen, P. J. (1999). The effects of chronic administration of hydrocortisone on cognitive function in normal male volunteers. *Psychopharmacology*, 145(3), 260–266. https://doi.org/10.1007/s002130051057

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3), 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015

Yushkevich, P. A., Pluta, J. B., Wang, H., Xie, L., Ding, S.-L., Gertje, E. C., ... Wolk, D. A. (2015). Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment: Automatic morphometry of MTL subfields in MCI. *Human Brain Mapping*, 36(1), 258–287. https://doi.org/10.1002/hbm.22627

Zhu, S.-W., Yee, B. K., Nyffeler, M., Winblad, B., Feldon, J., & Mohammed, A. H. (2006). Influence of differential housing on emotional behaviour

and neurotrophin levels in mice. *Behavioural Brain Research, 169*(1), 10–20. http://dx.doi.org/10.1016/j.bbr.2005.11.024

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.