Reproducibility of the correlative triad among aging, dopamine receptor availability, and cognition

Eric J. Juarez[1, 2], Jaime J. Castrellon[1, 2], Mikella A. Green[1, 2], Jennifer L. Crawford[3], Kendra L. Seaman[2, 4], Christopher T. Smith[5], Linh C. Dang[5], David Matuskey[6,7], Evan D. Morris[6,7,8], Ronald L. Cowan[5,9,10], David H. Zald[5,9], Gregory R. Samanez-Larkin[1,2]

[1]Department of Psychology and Neuroscience, Duke University

[2]Center for Cognitive Neuroscience, Duke University

[3]Department of Psychology, Yale University

[4]Center for the Study of Aging and Human Development, Duke University

[5]Department of Psychology, Vanderbilt University

[6]Department of Radiology and Biomedical Imaging, Yale University

[7]Department of Psychiatry, Yale University

[8]Department of Biomedical Engineering, Yale University

[9]Department of Psychiatry and Behavioral Sciences, Vanderbilt University School of Medicine

[10]Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center

Author Note

Jennifer Crawford is now at Washington University in St Louis. Kendra Seaman is now at University of Texas Dallas. Christopher T. Smith is now at North Carolina State University. Linh Dang is now at Google. Some of the results reported in this manuscript were presented in a poster at the Cognitive Neuroscience Society annual meeting (2018) and the Cognitive Aging Conference (2018). Correspondence concerning this article should be addressed to Eric Juarez, Department of Psychology and Neuroscience, Duke University, Box 90999, Durham, NC 27708. Email: eric.juarez@duke.edu

Acknowledgements: This research was supported by National Institute on Aging Pathway to Independence Award R00-AG042596, National Institute on Aging grant R01-AG044838, and National Center for Advancing Translational Science grant UL1TR000445. K.L.S. was supported by T32-AG000029.

Data: Data used in the manuscript can be viewed and downloaded from: https://osf.io/xjqe9/

Abstract

The evidence that dopamine function mediates the association between aging and cognition is one of the most cited findings in the cognitive neuroscience of aging. However, few and relatively small studies have directly examined these associations. Here we examined correlations among adult age, dopamine D2-like receptor (D2R) availability, and cognition in two cross-sectional studies of healthy human adults. Participants completed a short cognitive test battery and, on a separate day, a PET scan with either the high-affinity D2R tracer [^{18}F]Fallypride (Study 1) or [^{11}C]FLB457 (Study 2). Digit span, a measure of short-term memory maintenance and working memory, was the only cognitive test for which dopamine D2R availability partially mediated the age effect on cognition. In Study 1, age was negatively correlated with digit span. Striatal D2R availability was positively correlated with digit span controlling for age. The age effect on digit span was smaller when controlling for striatal D2R availability. Although other cognitive measures used here have individually been associated with age and D2R availability in prior studies, we found no consistent evidence for significant associations between low D2R availability and low cognitive performance on these measures. These results at best only partially supported the correlative triad of age, dopamine D2R availability, and cognition. While a wealth of other research in human and non-human animals demonstrates that dopamine makes critical contributions to cognition, the present studies suggest caution in interpreting PET findings as evidence that dopamine D2R loss is a primary cause of broad age-related declines in fluid cognition.

Keywords: cognition, working memory, aging, dopamine, PET

Introduction

There is considerable evidence from hundreds of studies over the past half century that fluid cognitive abilities decline over the course of the adult life span (Horn & Cattell, 1967; Park & Schwarz, 2000). These fluid cognitive abilities include processing speed, working memory, executive control, and episodic memory. Over the past few decades, functional and structural brain imaging have been used to identify the aspects of neural decline that account for these behavioral losses (Cabeza, Nyberg, & Park, 2016; Grady, 2012; Jagust, & D'Esposito, 2009).

Perhaps the earliest and most consistently reported adult age differences identified in human neuroimaging studies were reductions in aspects of dopamine function. Nearly 100 studies using PET or SPECT imaging have been conducted over the past thirty years documenting lower dopamine receptor and transporter availability in older compared to younger adults. According to a recent meta-analysis (Karrer et al., 2017), adult age is strongly and negatively correlated with D1-like receptor availability ($r = -.77$), D2-like receptor (D2R) availability ($r = -.56$), and dopamine transporter availability ($r = -.68$). All of these measures of dopamine targets (receptors, transporters, or relevant enzymes) have been associated with fluid cognitive abilities (e.g., episodic memory, working memory, cognitive control, psychomotor speed, motor speed) in aging adults (Bäckman, Lindenberger, Li, & Nyberg, 2010).

By the late 1990s and early 2000s, there was evidence for a "correlative triad" suggesting that individual differences in dopaminergic status (assessed with PET/SPECT imaging of dopamine targets), but not chronological age, explain age-related deficits in fluid cognition (Volkow et al., 1998, Bäckman et al., 2000). The implication was that dopaminergic decline makes a critical contribution to fluid cognitive decline with age. The theory is one of the most highly cited in the neuroscience and psychology of aging. However, there are a number of limitations of the existing empirical studies. For instance, although some of the foundational studies included participants across much of the adult life span (Bäckman et al., 2000: ages 21-68, Volkow et al., 1998: ages 24-86) and used the same tracer, [^{11}C]Raclopride, each study relied on measures tapping different cognitive functions to provide evidence for the correlative triad. Bäckman and colleagues (2000) reported associations between D2R and perceptual speed and episodic memory whereas Volkow and colleagues (1998) reported associations between D2R and motor and executive function. As can be seen in Table 1, mixed findings, inconsistent reporting of results, and small sample sizes across these studies make evaluation of the

reproducibility of these effects difficult. Potential meta-analysis of the previous empirical work is complicated by the inconsistent use of cognitive tests and/or selective reporting of only statistically significant dopamine-cognition associations (Karrer, 2017).

Only 10 out of 95 previously published datasets on adult age effects on dopamine targets (receptors, transporters, or relevant enzymes) using PET or SPECT reported cognitive associations (Karrer et al., 2017). Ten papers would be sufficient for meta-analysis, but almost none of the papers report associations with the same cognitive measures. The reported associations between different measures of dopamine function and cognitive function spanned several domains such as cognitive control (Lappin et al., 2009), episodic memory (Bäckman et al., 2000), or working memory (Bäckman et al., 2011), but the same cognitive measures were not reported across studies. Recent studies have questioned and begun to revise the broad claims of the original correlative triad hypothesis (Nyberg et al., 2016, Lövdén et al., 2017, Rieckmann et al., 2018, Karalija et al, 2019, Salami, et al., 2019). For example, third variables such as education, lifestyle, or genetics have been implicated as potential critical moderators of dopamine-cognition relationships with aging (Lövdén et al., 2017, Karalija et al., 2019). Thus, given the assumptions that dopaminergic loss underlies fluid cognitive deficits, or may even be a biomarker for cognitive decline, it is important to establish the replicability of the effects.

Here, in two adult life-span studies at different sites using different tracers, with complementary coverage of cortical and subcortical regions, we investigated whether adult age differences in standard neuropsychological measures of fluid cognition were due to individual differences in D2R availability.

Method

Two studies at two different universities examined the replicability of associations between age, D2R availability, and cognitive test performance.

Participants

Study 1. Healthy participants were recruited from the greater Nashville community for PET scans performed at Vanderbilt University. Study 1 consisted of 83 adults (ages 22 to 83 years, mean=49.80, SD=17.68, 35 males). See Supplementary Figure 1 for age histogram. Participants were medically and mentally healthy and were assessed by a physical examination, comprehensive metabolic panel, complete blood count, EKG, and interviews of medical and psychiatric history. Inclusion criteria were the following: no illicit drug use in last 2 months, no

use of any psychotropic medication in last 6 months, no current uncontrolled medical condition such as neurological, cardiovascular, endocrine, renal, liver, or thyroid pathology, no history of neurological or psychiatric disorders, no current tobacco or nicotine use, alcohol consumption no greater than 8 ounces of whiskey (~5 standard alcoholic drinks)/week, and no claustrophobia or other MRI contraindications. Females had negative pregnancy tests at intake and on the day of the scan. Approval for the Study 1 protocol was obtained from the Vanderbilt University Human Research Protection Program and the Radioactive Drug Research Committee and all participants completed written informed consent.

Study 2. Healthy participants were recruited from the greater New Haven community for PET scans performed at Yale University. Study 2 consisted of 37 adults (ages 26 to 79 years, mean=47.81, SD=16.93, 17 males). See Supplementary Figure 2 for age histogram. Participants were medically and mentally healthy and were assessed by a physical examination, comprehensive metabolic panel, complete blood count, EKG, and interviews of medical and psychiatric histories. Inclusion criteria were the following: no prescription or illicit drug use, no history of tobacco or nicotine use, no current uncontrolled medical conditions such as cardiovascular, endocrine, renal, liver, or thyroid pathologies, no history of neurological or psychiatric disorders, alcohol consumption no greater than 8 ounces of whiskey (~5 standard alcoholic drinks)/week, and no claustrophobia or other MRI contraindications. Females had negative pregnancy tests at intake and on the day of the scan. Approval for the Study 2 protocol was obtained from the Yale University Human Investigation Committee and the Yale University Hospital Radiation Safety Committee and all participants completed written informed consent.

Cognitive measures

Four cognitive tests were collected in both studies. A relatively brief neuropsychological battery was selected to characterize age differences in common measures of aspects of fluid cognition that had been previously correlated with D2R. All measures were included based on literature indicating that these functions decline with age (Hicks & Birren, 1970; Park & Schwarz, 2000; Salthouse, 2004; Buckner, 2004) and have been linked to prefrontal function or DA modulation (West, 1996; Luciana, Depue, Arbisi, & Leon, 1992; Rubin, 1999). The Trail-Making Test (TMT) measures speed of processing, visual search, and cognitive flexibility (Crowe, 1998). The test includes two parts. In Part A, participants draw lines connecting numbers in numerical order (1-2-3-4) while an experimenter records time to completion.

Completion time on Part A is used as a measure of psychomotor speed and visual search. Part B includes similar psychomotor and search demands but also includes alternation between numbers and letters (e.g., 1-A-2-B-3-C). The difference between these scores (B – A) is used as a measure of cognitive flexibility. Three measures of memory from the Wechsler Memory Scale WMS-III, two that assessed short-term maintenance (Digit Span) and working memory (Digit Span, Letter-Number Sequencing) and one that assessed longer-term memory (Verbal Paired Associates Delayed Recall) were also collected in both studies. The total score of forward plus backward digit span was used for analyses. In the forward digit span task, participants repeat a list of numbers as they were presented by the experimenter. In the backward digit span task, participants repeat a list of numbers in the reverse order. In the letter-number sequencing task, participants must reorder increasingly longer strings of letters and numbers in alphanumeric order. In the verbal paired associates task, participants learn a series of two-word associations which they recall four times immediately after presentation by the experimenter. In the delayed recall portion of the task, participants repeat the paired associates after an extended delay. Only the delayed recall portion of the verbal paired associates task was used for analysis.

PET Acquisition

Study 1. [^{18}F]Fallypride, (S)-N-[(1-allyl-2-pyrrolidinyl)methyl]-5-(3[^{18}F]fluoropropyl)-2,3-dimethoxybenzamide (hereafter referred to as Fallypride), was produced in the radiochemistry laboratory attached to the PET unit at Vanderbilt University Medical Center, following synthesis and quality control procedures described in US Food and Drug Administration IND 47,245. Prior to the PET scan, T1-weighted magnetic resonance (MR) images (TFE SENSE protocol; Act. TR=8.9 ms, TE=4.6 ms, 192 TFE shots, TFE duration=1201.9 s, FOV=256x256 mm, voxel size=1x1x1 mm) were acquired on a 3T Philips Intera Achieva whole-body scanner (Philips Healthcare, Best, The Netherlands). PET data were collected on a GE Discovery STE (DSTE) PET scanner (General Electric Healthcare, Chicago, IL, USA). Serial scan acquisition was started simultaneously with a 5.0 mCi (185 MBq; average = 5.06 mCi, SD = 0.23) slow bolus injection of the dopamine D2/3 tracer Fallypride (median specific activity: 9.24 mCi/nmol). CT scans were collected for attenuation correction prior to each of the three emission scans, which together lasted approximately 3.5 hours with two breaks for participant comfort. Acquisition times for the dynamic PET scans have been reported previously (Smith et al., 2016). After decay correction and attenuation correction, PET scan

frames were corrected for motion using SPM8 (Friston et al., 1995) with the 20th dynamic image frame of the first series serving as the reference image. The realigned PET frames were then merged and re-associated with their acquisition timing information in PMOD (PMOD Technologies, Zurich, Switzerland) 's PVIEW module to create a single 4D file for use in PMOD's PNEURO tool for further analysis (see below).

Study 2. [^{11}C]FLB 457, 5-bromo-N-[[(2S)-1- ethyl-2-pyrrolidinyl]methyl]-3-methoxy-2-(methoxy-^{11}C) benzamide (hereafter referred to as FLB), was synthesized as previously described by Sandiego et al. (2015). PET scans were acquired on the High Resolution Research Tomograph (HRRT; Siemens Medical Solutions, Knoxville, TN, USA). FLB (median specific activity: 7.80 mCi/nmol) was injected intravenously as a bolus (315 MBq; average = 8.62 mCi, SD = 2.03) over one minute by an automated infusion pump (Harvard Apparatus, Holliston, MA, USA). Prior to each scan a six-minute transmission scan was performed for attenuation correction. Dynamic scan data were acquired in list mode for 90 min following the administration of FLB and reconstructed into 27 frames (6 × 0.5 mins, 3 × 1 min, 2 × 2 mins, 16 × 5 mins) with corrections for attenuation, normalization, scatter, randoms, and dead time using the MOLAR (Motion-compensation OSEM List-mode Algorithm for Resolution-Recovery Reconstruction) algorithm (Carson, Barker, Liow, & Johnson, 2003). Event-by-event motion correction (Jin et al., 2013) was applied using a Polaris Vicra optical tracking system (NDI Systems, Waterloo, Canada) that detects motion using reflectors mounted on a cap worn by the participant throughout the duration of the scan. Prior to the PET scan, T1-weighted magnetic resonance (MR) images (MPRAGE protocol; TR=2.4 s, TE=1.9 ms, FOV=256x256 mm, voxel size=1x1x1 mm) were acquired on a 3T Trio whole-body scanner (Siemens Medical Systems, Erlangen, Germany).

PET data processing

Binding potential was estimated relative to non-displaceable tracer (BP_{ND}) within ROIs using PMOD's PNEURO module. The Hammers atlas (Hammers et al., 2003) available internally in PNEURO and PNEURO's deep nuclei option was used to parcellate each participant's grey matter as determined from a segmentation of their T1 MRI image into 30 bilateral cortical areas (including amygdala and hippocampus), 5 bilateral deep nuclei (caudate, putamen, ventral striatum, thalamus, and pallidum), plus bilateral cerebellum, and brainstem. After parcellation, the MRI and PET data were co-registered via a rigid matching procedure

based on the normalized mutual information criterion and PET data was resampled to the MRI space (1x1x1mm). Accurate co-registration is critical for the partial-volume correction methods (described later) (Hutton et al., 2013). To check registration quality, we calculated quality control metrics using the PFUS module in PMOD 4.0 for a subset of 42 participants including the oldest 10 participants in each study, since older participants tend to be the hardest to co-register due to grey matter loss with age. The average Dice coefficient between PET data warped to MRI space and each participant's own anatomical MRI data was 0.83±0.05. The Dice coefficient measures the ratio between the number of true positives (overlapping voxels) compared to the number of true positives (overlapping voxels) plus the number of false positives (non-overlapping voxels). This suggests that registration methods were successful and consistent across the sample. We also tested for age differences in the other registration quality control metrics provided by PFUS (sensitivity, specificity, Jaccard index) across the subsamples and found no relationships between age and any of these registration quality metrics.

The PNEURO parcellation ROIs were then used to extract time activity curves (TACs) from the PET data before modeling. These TACs were then fitted using a two-compartment simplified reference tissue model (Lammertsma & Hume, 1996) to obtain BP_{ND} values using PMOD's PKIN module with a merged, bilateral cerebellum ROI serving as the reference region. Volume of interest (VOI)-based partial volume-corrected (PVC) estimates of BP_{ND} were computed based on work from Rousset, Ma, & Evans (1998) following the methods of Smith and colleagues (2017) which includes details on the specification of VOIs.

Mean BP_{ND} in the bilateral midbrain was extracted from an ROI drawn in MNI standard space using previously described guidelines (Dang et al., 2012, Dang et al., 2013) and registered to PET images using the same transformations for cerebellum registration to PET images. The midbrain ROI was not PVC. The PVC procedure requires specifying boundaries around a region and we did not collect an anatomical image that would have allowed the contrast required to precisely define these boundaries in the midbrain (e.g., GRASE or FFE; Eapen et al., 2011). For consistency with the earlier papers reporting evidence for the correlative triad that also used a whole striatum ROI, here, a whole striatum measure of PVC BP_{ND} was computed as a volume-weighted average of the caudate, putamen, and ventral striatum ROIs.

Region-of-interest analyses

To limit the number of statistical tests, we created volume-weighted BP_{ND} summary measures for each ROI. We limited our analyses to regions with the highest averaged levels of uncorrected BP_{ND} exceeded group-averaged BP_{ND} that have been used in previous analyses of these datasets (e.g., Castrellon et al., 2019). For the following analyses, we used a whole striatum ROI consisting of caudate, putamen, and ventral striatum since much of the prior literature has investigated the correlative triad in the whole striatum. Data for striatal subregions is included on the OSF project page: https://osf.io/xjqe9/.

All statistical tests were conducted in R. Within each study, linear regression analyses were used to evaluate the correlative triad. First, regression analyses examined the effects of age on cognitive performance. Second, regression analyses examined the effects of D2R BP_{ND} on cognitive performance controlling for the linear effect of age. Separate regression analyses were run for each cognitive test and ROI. Both studies had ROI data for midbrain, anterior cingulate, thalamus, amygdala, hippocampus, and insula. Only Study 1 which used the tracer Fallypride included the striatum, because the FLB tracer used in Study 2 does not produce stable estimates of D2R BP_{ND} in striatum (Halldin et al., 1995). Given that previous studies reported significant associations between D2R BP_{ND} within many of these ROIs and some of these cognitive tests, we used a liberal statistical threshold of $p < .05$, two-tailed (Tables 1, 2) to indicate significant associations. We only conducted formal tests of mediation if two conditions were met: (1) age was negatively associated with cognitive performance and (2) D2R BP_{ND} was positively associated with cognitive performance after controlling for age. To assess mediation, we started by using an early approach that was standard in the field when the correlative triad was proposed: is the effect of age on cognition significant before but not after controlling for D2R BP_{ND}? We also computed bootstrapped estimates of the indirect effect (5000 replications) to evaluate whether age-related decline in D2R BP_{ND} significantly carried the influence of age to cognitive performance.

All data and code are publicly available on OSF: https://osf.io/xjqe9/

Whole-Brain Analyses

To explore potential associations outside of the ROIs, whole-brain voxelwise analyses used uncorrected estimates of D2R BP_{ND} (that were not PVC). Whole-brain analyses controlled for multiple comparisons using threshold-free cluster enhancement with 5000 permutations per contrast computed using the “randomise” function in FSL. Positive and negative contrasts were

computed for each of the cognitive tasks for each study while controlling for age. Significance was assessed using a 0.05 threshold on the corrected p maps. FLB BP_{ND} maps were smoothed at 8 mm prior to group-level analysis (Plaven-Sigray et al., 2017). If a region was identified that positively correlated with cognitive performance (i.e., higher BP_{ND} = better performance), we conducted formal tests of mediation in R using the same procedures as above.

All beta coefficients for linear effects are reported as standardized betas.

Results

Study 1

Effects of age on cognition. In Study 1, Trails A and Trails B-A were positively correlated with age, indicating longer time to completion as age increased (Trails A: $\beta_{81} = 0.461$, [0.268, 0.655], p = 0.0001, Trails B-A: $\beta_{81} = 0.222$, [0.009, 0.434], p = 0.044). The three memory-related tasks all negatively correlated with age (digit span total: $\beta_{81} = –.274$, [–0.483, –0.064], p = .012, letter number sequencing: $\beta_{81} = –.554$, [–0.735, –0.372], p < 0.0001, delayed recall: $\beta_{77} = –0.614$, [–0.790, –0.438], p < 0.0001).

Effects of age on dopamine D2R availability. For five out of the seven ROIs in Study 1, D2R BP_{ND} was significantly negatively correlated with age (amygdala: $\beta_{81} = –.292$, [–0.500,–0.084], p=0.007, thalamus: $\beta_{81}= –0.260$, [–0.470,–0.049], p=0.018, midbrain: $\beta_{81}=–0.504$, [–0.693,–0.316], p<0.0001, striatum: $\beta_{81}= –.450$, [–0.644,–0.255], p<0.0001, insula: $\beta_{81}= –0.359$, [–0.562,–0.156], p<0.0001). D2R BP_{ND} in the hippocampal and the anterior cingulate ROIs did not significantly correlate with age (hippocampus: $\beta_{81}= –0.201$, [–0.414, 0.012], p=0.068, anterior cingulate cortex: $\beta_{81}= –0.21$, [–0.425, 0.001], p=0.054.

Effects of dopamine D2R availability on cognition. When controlling for age in Study 1, only striatal D2R BP_{ND} positively correlated with task performance such that higher binding was associated with better performance (for all correlations between D2R BP_{ND} and cognitive tasks see Table 2). Striatal D2R BP_{ND} was positively correlated with digit span total score (β_{81}=0.24, 95% CI=[0.010, 0.471], p=0.044) suggesting that higher D2R BP_{ND} was associated with better task performance. Striatal D2R BP_{ND} and insula D2R BP_{ND} also positively correlated with length of time spent completing the Trails A task (striatum: β_{81}=0.270, [0.060,0.479], p=0.014, insula: β_{81}=0.239, [0.038, 0.440], p=0.023), a counterintuitive finding suggesting that higher D2R availability was associated with worse task performance. For associations between

D2R availability and both cognition and age using uncorrected (non-PVC) BP_{ND}, see Supplementary Tables 1–2. For associations between age and cognition controlling for D2R availability, see Supplementary Table 3.

The whole-brain analyses of the Fallypride data revealed several regions where higher D2R BP_{ND} was significantly associated with better memory on the verbal paired associates delayed recall test. There were positive correlations between delayed recall and D2R BP_{ND} along the medial frontal gyrus, within anterior portions of superior and middle frontal gyri, and in superior portions of the body of the caudate. We created spherical ROIs in each of the two frontal cortical regions to evaluate potential mediation. We used an existing caudate ROI that was estimated during the PVC procedure. Unthresholded statistical maps can be found on NeuroVault at: https://neurovault.org/collections/3707/.

Mediation Analyses. The correlation between striatal D2R BP_{ND} and digit span was the only effect within a priori ROIs to meet our criteria for mediation analyses. In Study 1, the effect of age on digit span was reduced from a significant effect of $\beta = -.27$ to a nonsignificant effect of $\beta = -.17$ after controlling for striatal D2R BP_{ND} (see Figure 1). Although 40% of the total effect was mediated by striatal BP_{ND}, bootstrapped estimates of the indirect effect were not significant, a*b= – .024, 95% CI = [–.053, –.001], z = –1.861, p = .063, despite the modest effect size.

We also evaluated mediation within three regions that were identified in whole brain analyses to be positively correlated with delayed recall for verbal paired associates (medial frontal gyrus, superior/middle frontal gyrus, superior caudate). Within all of these regions, the effect of age on delayed recall was significant before and after controlling for D2R BP_{ND} (medial frontal gyrus: $\beta = -.61$ to $\beta = -.46$; superior/middle frontal gyrus: $\beta = -.61$ to $\beta = -.52$; caudate: $\beta = -.61$ to $\beta = -.57$). Bootstrapped estimates of the indirect effect of D2R BP_{ND} were not significant in the superior/middle frontal gyrus, a*b=–0.012, 95% CI = [–0.033, 0.006], z=–1.217, p = 0.224, medial frontal gyrus, a*b = –.02, 95% CI=[–0.054, 0.011], z=–1.244,p=0.214, or caudate, a*b= –.006, 95% CI=[–.022, .005],z= –.925, p=.355. Thus, none of the associations in these additional regions identified in the whole-brain analyses were consistent with statistical mediation by D2R availability of the effect of age on delayed recall.

Study 2

Effects of age on cognition. In Study 2, Trails A and Trails B-A were significantly and positively correlated with age, indicating longer time to completion at older compared to younger

ages (Trails A: β_{35}=0.445, [0.149,0.742], p=0.006, Trails B-A: β_{35}=0.499, [0.211,0.786], p=0.002). Long-term memory and one working memory task were significantly and negatively correlated with age (delayed recall: β_{35}= –0.521, [–0.803, –0.238], p=0.001, letter number sequencing: β_{35}= –0.393, [–0.697,–0.088], p=0.016). Digit span did not correlate with age (β_{35}= –.070, [–0.401,0.260] p=0.678) in Study 2, although the confidence interval of the effect of age on digit span was highly overlapping with Study 1.

Effects of age on dopamine D2R availability. For four out of the six ROIs in Study 2, D2R BP_{ND} was significantly and negatively correlated with age (midbrain: β_{35}= –0.640, [–0.895,–0.386], p<0.0001, anterior cingulate cortex: β_{35}= –0.477, [–0.769,–0.186], p=0.003, amygdala: β_{35}= –0.340, [–0.652,–0.029], p =0.039, insula: β_{35}= –0.438, [–0.736, –0.140], p=0.007). The hippocampal and thalamic ROIs did not have significant negative correlations between age and D2R BP_{ND} (thalamus: β_{35}= –0.216, [–0.540,0.106], p=0.197, hippocampus: β_{35}= –0.161, [–0.488,0.166], p=0.341).

Effects of dopamine D2R availability on cognition. When controlling for age in Study 2, letter number sequencing total score negatively correlated with D2R BP_{ND} in the midbrain (β_{34}= –0.453, [–0.826,–0.081], p=0.023) and the hippocampus (β_{34}=–0.323, [–0.617,–0.030], p=0.038), (for all correlations between task and D2R BP_{ND} see Table 3) such that higher D2R BP_{ND} was associated with worse performance. Likewise, Trails B-A was positively correlated with D2R BP_{ND} in the midbrain (β_{34}= 0.390, [0.034,0.746], *p*=0.039) such that higher D2R availability was associated with worse performance. For associations between D2R availability and both cognition and age using uncorrected (non-PVC) BP_{ND}, see Supplementary Tables 4–5. For associations between age and cognition controlling for D2R availability, see Supplementary Table 6.

The whole-brain analyses of the FLB data did not reveal any regions where higher D2R BP_{ND} was significantly associated with better performance. There were negative correlations with letter-number sequencing and D2R BP_{ND} in lateral temporal cortex and lateral occipital cortex. There were positive correlations with Trails B-A and D2R BP_{ND} in superior and middle temporal gyrus. These effects were opposite of the predicted direction such that higher D2R availability corresponded to worse performance. Thus, we did not evaluate mediation for any of

these effects within the FLB data. Unthresholded statistical maps can be found on NeuroVault at: https://neurovault.org/collections/3707/.

Given the positive associations between delayed recall and Fallypride D2R BP_{ND} in Study 1, we also evaluated associations between memory tasks and FLB D2R BP_{ND} and potential mediation by FLB D2R BP_{ND} in Study 2 within the two regions identified in Study 1 that were available for analysis in Study 2 (medial frontal gyrus, superior/middle frontal gyrus). D2R BP_{ND} was not correlated with delayed recall in either of these regions (medial frontal gyrus: β=–.11 ; superior/middle frontal gyrus: β=.07). Given the lack of association between D2R BP_{ND} and delayed recall there was an expected lack of a mediation by D2R BP_{ND}. The effect of age on delayed recall was significant before and after controlling for D2R BP_{ND} (medial frontal gyrus: β=–.52[z] to β=–.60; superior/middle frontal gyrus: β=–.52 to β=–.48); bootstrapped estimates of the indirect effect were not significant in the medial frontal gyrus, a*b=–0.013, 95% CI=[–.024, .055], z=.658, p=.51, or superior/middle frontal gyrus, a*b=–0.007, 95% CI=[–.055, .022], z=–.388, p=.698. Overall the associations between cortical D2R availability and memory were not consistent across studies using different tracers (Study 1: Fallypride; Study 2: FLB) and neither set of results provided evidence for mediation.

Discussion

Here, in two cross-sectional studies of adulthood, we examined associations between aging, dopamine D2R availability using PET imaging, and cognition which form the bases for the correlative triad theory. We found that very few associations were significant and only one was in the predicted direction. While associations between aging and cognition and aging and D2R availability were mostly significant and highly reproducible across studies, the only region and measure that was consistent with the correlative triad was striatal D2R BP_{ND} and digit span in Study 1 that used Fallypride. Although inclusion of the dopamine D2R measure reduced the age effect on digit span from significant to non-significant, statistical evaluation of mediation was non-significant. Our findings provide partial support for the correlative triad in one data set for working memory, but qualify general claims of an association between age-related decline in the dopamine system and age-related decline in fluid cognition.

Several individual prior studies have reported associations between D2R availability across a variety of different cognitive measures. As a result, we did not have strong hypotheses

about which of the cognitive measures would show the strongest effects. Consistent with the positive associations reported here between D2R availability and short-term maintenance, several prior studies reported associations between dopamine target measures and short-term maintenance and/or working memory (positive: Erixon-Lindroth, et al., 2005 (dopamine transporter) ; Backman, et al., 2011 (D1) ; Salami, et al., 2018 (D2R); negative: Reeves, 2005 (D2R)). Two of the cognitive measures collected in the present studies assessed aspects of working memory: digit span total score and letter-number sequencing. Digit span, but not letter-number sequencing, was negatively associated with age and positively associated with striatal D2R availability in Study 1 using Fallypride. It is unclear why there was an association with digit span but not letter-number sequencing, given prior evidence that both tasks load onto the same latent factor (Parmenter, Shucard, Benedict, & Shucard, 2006), although it may be noted that letter number sequencing requires a more complex manipulation of information in working memory. One previous study indicated that digit span and logical memory were most strongly correlated with D2R availability when controlling for age (Chen et al., 2005). However, the majority of prior studies reporting associations between measures of dopamine targets (receptors, transporters, or relevant enzymes) and working memory used n-back tasks (see Salami et al., 2018). Interestingly, some research suggests that digit span does not strongly correlate with n-back, although these studies did not specifically use the combined forward and backward score used in the present studies (Miller, Price, Okkun, Montijo, & Bowers, 2009; Parmenter, Shucard, Benedict, & Shucard, 2006). Nevertheless, it is possible that we would have detected stronger associations between D2R availability and working memory if we used an n-back task. That said, Nyberg et al. (2016) did not find a relationship between working memory (and perceptual speed) and D2R availability measured using the tracer [^{11}C]Raclopride, prompting a hypothesis that D2R might be more strongly related with episodic memory and the hippocampus.

We observed positive associations between digit span and D2R availability in the striatum but not in any other regions of interest. Although task-based neuroimaging studies most often highlight associations between prefrontal cortical function and working memory (Braver et al., 1997; Wager & Smith, 2003; Constantinidis & Klingberg, 2016), striatal dopamine synthesis capacity has also been associated with working memory (Landau, 2009). Additionally, lesions to the striatum produce deficits in working memory and many other fluid cognitive tasks assumed to be primarily dependent on the prefrontal cortex (Rubin, 1999). Furthermore, a

pharmacological fMRI study showed that amphetamine administration increased striatal as well as frontal cortical BOLD signal during an n-back working memory task in older and younger adults (Garrett et al. 2015). The lack of associations between frontal D2R availability reported here might reflect the lower signal in lateral cortical regions for the radiotracers used in this study, particularly for Fallypride (Mukherjee et al., 2002; Vandehey et al., 2010; Zald et al., 2010). Estimates of frontal D2R availability using FLB are higher compared to Fallypride, but many lateral regions still have quite low signal, which may constrain the variance available to explain variation in cognition. Thus, it is possible that there are positive associations between D2R availability and cognition in regions of the frontal cortex that we were not able to detect. While we did observe an association between striatal D2R availability and digit span, this could only be tested within Study 1 using Fallypride (since FLB used in Study 2 does not produce stable estimates of striatal BP_{ND}). Thus, we could not evaluate the replicability of that effect across data sets. Differences in tracer pharmacokinetics might also contribute to differences in the size of observed associations. The present studies utilized different tracers (i.e., the high affinity tracers Fallypride and FLB457) compared to the foundational studies in this area (which mostly used Raclopride), but we would not expect this to lead to a fundamentally different pattern of associations.

Exploratory whole-brain analyses identified additional frontal cortical regions that were positively associated with delayed recall in Study 1 using Fallypride. However, additional analyses provided no evidence for statistical mediation. In fact, the effects of age on delayed recall were nearly the same before or after controlling for D2R availability in these frontal regions in both studies. Finally, the associations between frontal D2R availability and delayed recall were not significant in Study 2 that used FLB which produces higher estimates of cortical D2R availability. Overall, there was no evidence for frontal cortical mediation within Study 1 using Fallypride or Study 2 using FLB and the associations between dopamine D2R availability and delayed recall were inconsistent across Studies 1 and 2.

Contrary to our hypotheses and the correlative triad theory, we also observed some *negative* associations between dopamine D2R availability and cognition such that higher D2R availability was associated with worse task performance. Other studies have also reported such negative associations. For example, in a recent study of a cohort of older adults, one subgroup had lower D2R availability in the striatum but also had higher working memory scores (Lövdén

et al. 2017). Levels of D2R availability could reflect compensatory changes in response to lower endogenous DA levels. Thus, it is possible that the observed negative associations are indicators that relatively higher endogenous dopamine levels are associated with better cognitive performance. However, the consistent and highly replicable evidence for lower D2R availability in older age and lower fluid cognitive performance in older age runs counter to this interpretation especially in cross-sectional life-span studies. Note, also, that these effects were not consistent across samples within regions that were reported across both studies.

Though studies that laid the foundation for the correlative triad theory also used D2R tracers, it is possible that other components of dopamine function, such as other receptor subtypes or dopamine release, are better correlated with fluid cognitive ability. For example, D1-like receptor availability has been associated with motor performance (Wang et al., 1998) as well as several memory measures (Sawaguchi & Goldman-Rakic, 1991; Sawaguchi, 2001; Karlsson, 2009, Backman, 2011). However, even studies with other measures of dopamine function have failed to replicate some of these associations. For example, Rieckmann et al., 2018 found no significant associations between dopamine transporter density in the striatum and processing speed or executive function (factors composing some of the tests used in the present study). However, as noted earlier it is difficult to assess the reproducibility of these effects given the limited number of studies and inconsistent use of tasks.

Overall, many of the individual previously reported associations between dopamine D2R availability and cognition in smaller samples were not replicated in the present studies (e.g., Bäckman, 2000; Erixon-Lindroth et al., 2005). It is possible that the neuropsychological test measures used in the present studies lacked sensitivity relative to the tasks used in prior studies. Some prior studies reported associations between measures of dopamine function and measures of memory or cognitive flexibility based on scores from experimental behavioral tasks rather than neuropsychological measures. It is possible that those tasks better assess aspects of cognitive function that are most related to dopamine function. One possibility is that the neuropsychological tests in these two studies were not difficult enough. Salami and colleagues (2019) found that D2R-BOLD associations were greatest in a 3-back task compared to 1 or 2-back in older adults who performed normally on the working memory task, suggesting a relationship between D2R availability and task efficiency. Additionally, recent work suggests that individual differences in gene expression moderates the relationship between D2R

availability and cognitive test scores (Karalija et al., 2019). Thus, associations between D2R availability and cognitive measures in the present studies may be masked by genetic variation, which has contributed to a more recent refinement of the correlative triad hypothesis (Karalija et al., 2019). Relatedly, other individual difference factors (e.g., education and body-mass index) modify associations between dopamine and cognition in initial analyses of the COBRA longitudinal dataset (Lovden et al., 2017). Given that the emerging moderating effects are based on statistical interactions and multivariate associations which require much more statistical power than simple main effects, future studies should evaluate the replicability of these more complicated effects. While both of the studies reported here, particularly Study 1, were larger than the majority of studies that originally explored the correlative triad, they were underpowered to investigate individual differences independent of age or potential interactions between age or D2R availability and other variables of interest. If others would like to conduct exploratory analyses of potential interactions with some of these variables, we also include sex, education, income, race/ethnicity, and BMI in our publicly available data files (https://osf.io/xjqe9/). While we have not ourselves tested for such moderating effects, the present studies could be consistent with the recent evidence for moderating factors given the lack of support for a broad traditional linear mediation model of the correlative triad.

Recent research has identified broad variation in the effects of age on D2R availability across subregions of the brain (Seaman et al., 2019). It is possible that there are subregional effects consistent with the correlative triad that we missed in the present analyses. The studies here focused on a limited set of somewhat larger ROIs. Many of these ROIs were the same size (e.g., whole striatum) or even smaller than what has been used in previous studies of the correlative triad. The primary goal of the present analyses was to evaluate reproducibility so we used ROIs that would be most consistent with previously published research. The whole-brain analyses were included to identify potential subregional effects, although no clear subregional effects emerged from these analyses.

A final and critical limitation of our findings, and the literature in general, is that despite our relatively large samples, conclusions about aging are drawn from cross-sectional data. It is challenging to accurately evaluate statistical mediation of aging from cross-sectional data. Mediation in cross-sectional data may not generalize in longitudinal samples, and a mediation effect in a longitudinal sample may exist despite a lack of significant mediation in cross-sectional

data (Raz & Lindenberger, 2011). As in any cross-sectional study, the results reported here may be related to cohort rather than aging effects. Ongoing longitudinal work will be able to better address potential longitudinal associations between the dopamine system and cognition (Nevalainen et al., 2015).

Overall, the present findings suggest that associations between age differences in dopamine D2R availability and fluid cognition may not be as replicable or broad as initially assumed. There was some suggestive evidence for an association between D2R availability and short-term memory maintenance, but the majority of previously reported findings were not replicated here. A wealth of other research in human and non-human animals demonstrates that dopamine function makes critical contributions to cognition, but the present studies suggest caution in using PET data to make generalizations that dopamine receptor loss is a critical cause of broad age-related declines in fluid cognition.

Tables

Table 1. Summary of previous studies investigating aging, dopamine targets (receptors, transporters, or relevant enzymes), and cognition. D1R = dopamine D1-like receptors; D2R = dopamine D2-like receptors; DAT = dopamine transporter.

First author	Year	N	Cognitive measure	Other cognitive measures	Target/Tracer
Bäckman	2000	11	Dots, Trails A, FaceRecognition, Word recognition	NA	D2R [^{11}C]Raclopride
Bäckman	2011	40	spatial working memory task	Verbal free recall, digit symbol, vocabulary, information,	D1R [^{11}C]SCH23390
Berry	2016	36	Two condition task switching	California Verbal learning test, WMSIII logical memory story and visual reproduction, listening span, category fluency, WMS II digit span (forward/backward), arthmetic tasks, digit symbol test,	Synthesis capacity 6-[^{18}F]fluoro-L-*m*-tyrosine

				stroop, WMSIII mental control test, trails a, finger tapping	
Braskie	2008	37	Listening span, controlled oral word association (FAS), category fuency, stroop, trail making b-a	wisconsin card sort test, digit span	Synthesis capacity 6-[^{18}F]fluoro-L-*m*-tyrosine
Chen	2005	62	WMS-R	NA	D2R [^{123}I]IBZM
Dreher	2008	21	Slot machine task	NA	Synthesis capacity 6-[^{18}F]FluoroDOPA
Erixon-Lindroth	2005	12	word recall, figure recall, face recognition, visuospatial working memory, controlled oral association test (verbal fluency), information test from WAIS_R	NA	DAT [^{11}C]β-CIT-FE
Lappin	2009	6/10	Sequential motor learning /adapted version of tower of london	NA	D2R ligand displacement (DA release) [^{11}C]Raclopride

Mozley	2001	66	Pennsylvania verbal learning test	Auditory consonant trigrams task, stroop task	DAT TRODAT-1
Nevalainen	2015	181	Word recall, number-word recall, object-position recall, letter updating, numerical 3-back, spatial updating, letter comparison, number comparison, figure comparison	NA	D2R [^{11}C]Raclopride
Rieckmann	2018	54	Phonemic fluency, letter-number sequencing, Trail Making Test, digit-symbol test, logical memory test, free and cued selective reminding test, six-trial selective reminding test	NA	DAT [^{11}C]Altropane
Reeves	2005	30	Spatial span, spatial working memory task, tower of london	NA	D2R [^{11}C]Raclopride

Volkow	1998	30	Stroop Color-Word Test, Wisconsin Card Sorting Test, Symbol Digit Modalities test	Raven Standard Progressive Matrices	D2R [^{11}C]Raclopride

Table 2. Region of interest analyses for Study 1. Correlations between cognitive test scores and D2R BP_{ND} controlling for age (N=83). BP_{ND} estimates are PVC except midbrain which is uncorrected. * p < .05 uncorrected

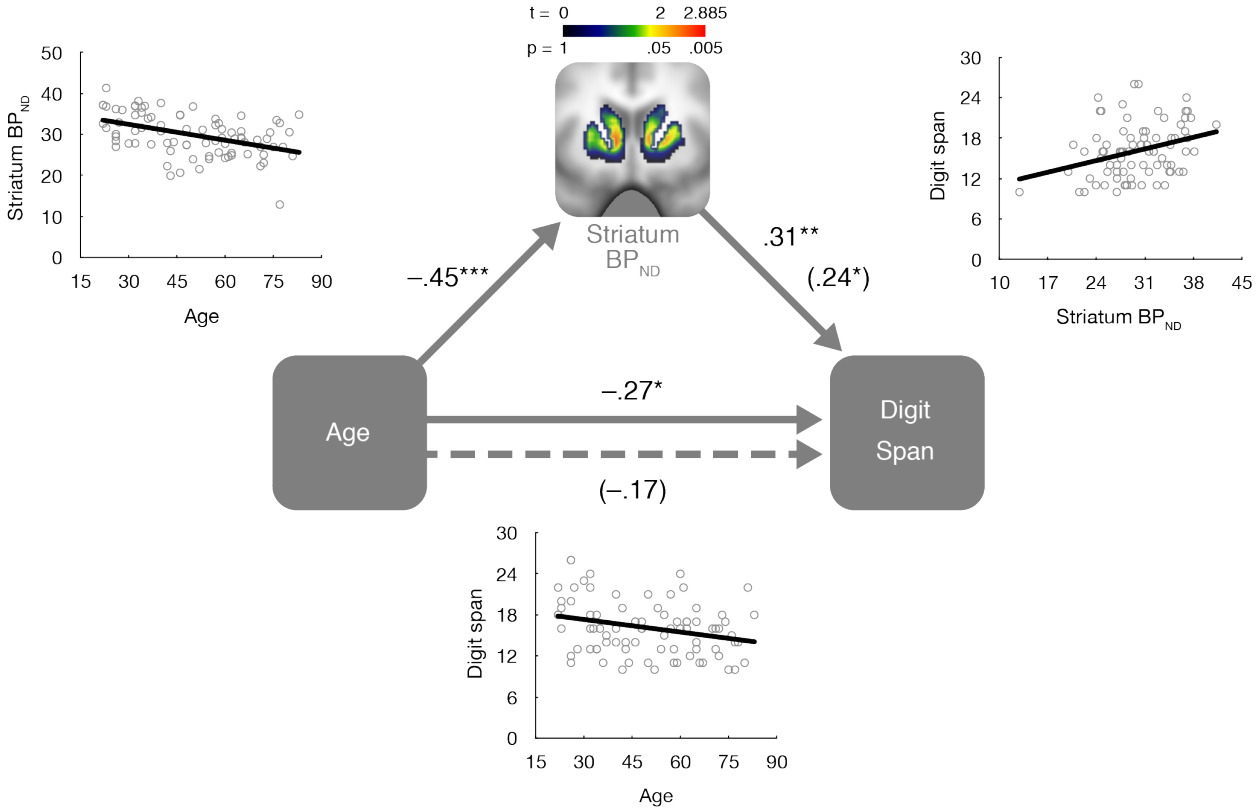
	Trails A	Trails B-A	Digit Span	Letter-Number Sequencing	Verbal Paired Associates Delayed Recall
Midbrain	.197 [–0.0237,0.418]	.032 [–0.216,0.279]	.115 [–0.127,0.359]	–.104 [–0.314,0.106]	–.007 [–0.213,0.199]
Striatum	**.270* [0.060,0.479]**	–.123 [–0.360,0.115]	**.241* [0.010,0.471]**	.039 [–0.165,0.243]	.141 [–0.053,0.336]
Anterior cingulate	.140 [–0.0564,0.337]	–.017 [–0.202,0.235]	–.007 [–0.223,0.208]	–.089 [–0.275,0.097]	.112 [–0.067,0.292]
Thalamus	.112 [–0.0876,0.312]	–.140 [–0.359,0.079]	.042 [–0.176,0.260]	–.063 [–0.251,0.126]	–.020 [–0.204,0.163]
Amygdala	.084 [–0.118,0.287]	.098 [–0.124,0.320]	.025 [–0.195,0.245]	–.149 [–0.337,0.039]	.039 [–0.146,0.224]
Hippocampus	.0294 [–0.169,0.228]	.028 [–0.191,0.246]	–.010 [–0.225,0.205]	–.090 [–0.275,0.095]	–.052 [–0.232,0.128]
Insula	**.239* [0.038,0.441]**	.028 [–0.201,0.256]	.055 [–0.171, 0.285]	–.094 [–0.288, 0.101]	–.076 [–0.265,0.113]

Table 3. Region of interest analyses for Study 2. Correlations between cognitive test scores and D2R BP_{ND} controlling for age (N=37). BP_{ND} estimates are PVC except midbrain which is uncorrected. * p < .05 uncorrected

	Trails A	Trails B-A	Digit Span	Letter-Number Sequencing	Verbal Paired Associates Delayed Recall
Midbrain	–.289 [–0.669,0.090]	**.390* [0.0341, 0.746]**	–.376 [–0.794,0.0416]	**–.453* [–0.826,–0.081]**	–.169 [–0.538,0.200]
Anterior cingulate	–.008 [–0.350,0.335]	.189 [–0.136,0.515]	–.076 [–0.457,0.305]	–.184 [–0.530,0.163]	–.048 [–0.374,0.278]
Thalamus	–.092 [–0.399,0.215]	.213 [–0.076,0.503]	–.146 [–0.486,0.194]	–.226 [–0.534,0.081]	.136 [–0.155,0.426]
Amygdala	–.191 [–0.505,0.122]	.199 [–0.103,0.502]	–.080 [–0.436,0.275]	–.313 [–0.624,–0.001]	–.155 [–0.456,0.146]
Hippocampus	.012 [–0.292,0.317]	.140 [–0.151,0.431]	–.143 [–0.479,0.193]	**–.323* [–0.617,–0.030]**	.082 [–0.208,0.371]
Insula	.016 [–0.312,0.351]	.309 [0.002, 0.616]	–.183 [–0.551,0.184]	–.187 [–.525,0.151]	–.081 [–0.399,0.237]

Figures

Figure 1. Associations between age, striatal dopamine D2R availability (BP_{ND}), and digit span. Age significantly and negatively correlated with striatal dopamine D2R availability. Striatal dopamine D2R availability was significantly and positively correlated with digit span before controlling for age and after controlling for age (effect within parentheses). The brain image displays a voxelwise t-statistic map within the striatal ROI for visualization of regional variation in the association between individual differences in D2R availability and digit span. The effect of age on digit span was reduced from significant to nonsignificant (effect reported within parentheses) after controlling for striatal D2R availability. All scatterplots display simple pairwise associations between variables without covariates. Effect sizes displayed are standardized linear effects. * p < .05, ** p < .01, *** p < .001



References

Bäckman, L., Ginovart, N., Dixon, R. A., Wahlin, T.-B. R., Wahlin, Å., Halldin, C., & Farde, L. (2000). Age-Related Cognitive Deficits Mediated by Changes in the Striatal Dopamine System. *American Journal of Psychiatry*, *157*(4), 635–637. https://doi.org/10.1176/ajp.157.4.635

Bäckman, L., Karlsson, S., Fischer, H., Karlsson, P., Brehmer, Y., Rieckmann, A., … Nyberg, L. (2011). Dopamine D1 receptors and age differences in brain activation during working memory. *Neurobiology of Aging*, *32*(10), 1849–1856. https://doi.org/10.1016/j.neurobiolaging.2009.10.018

Bäckman, L., Lindenberger, U., Li, S.-C., & Nyberg, L. (2010). Linking cognitive aging to alterations in dopamine neurotransmitter functioning: Recent data and future avenues. *Neuroscience & Biobehavioral Reviews*, *34*(5), 670–677. https://doi.org/10.1016/j.neubiorev.2009.12.008

Berry, A. S., Shah, V. D., Baker, S. L., Vogel, J. W., O'Neil, J. P., Janabi, M., … Jagust, W. J. (2016). Aging Affects Dopaminergic Neural Mechanisms of Cognitive Flexibility. *The Journal of Neuroscience*, *36*(50), 12559–12569. https://doi.org/10.1523/JNEUROSCI.0626-16.2016

Braskie, M. N., Wilcox, C. E., Landau, S. M., O'Neil, J. P., Baker, S. L., Madison, C. M., … Jagust, W. J. (2008). Relationship of Striatal Dopamine Synthesis Capacity to Age and Cognition. *Journal of Neuroscience*, *28*(52), 14320–14328. https://doi.org/10.1523/JNEUROSCI.3729-08.2008

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A Parametric Study of Prefrontal Cortex Involvement in Human Working Memory. *NeuroImage*, *5*(1), 49–62. https://doi.org/10.1006/nimg.1996.0247

Buckner, R. L. (2004). Memory and Executive Function in Aging and AD: Multiple Factors that Cause Decline and Reserve Factors that Compensate. *Neuron*, *44*(1), 195–208.

Cabeza, R., Nyberg, L., & Park, D. C. (Eds.). (2016). *Cognitive neuroscience of aging: Linking cognitive and cerebral aging*. Oxford University Press.

Carson, R. E., Barker, W. C., Liow, J. S., & Johnson, C. A. (2003, October). Design of a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT. In *Nuclear Science Symposium Conference Record*, 2003 IEEE (Vol. 5, pp. 3281-3285). IEEE.

Castrellon, J.J., Seaman, K.L., Crawford, J.L., Young, J.S., Smith, C.T., Dang, L.C., Hsu, M., Cowan, R.L., Zald, D.H., & Samanez-Larkin, G.R. (2019). Individual differences in dopamine are associated with reward discounting in clinical groups but not in healthy adults. *Journal of Neuroscience, 39*(2), 321-332. https://doi.org/10.1523/JNEUROSCI.1984-18.2018

Chen, P. S., Kuang Yang, Y., Lee, Y.-S., Yeh, T. L., Lee, I. H., Chiu, N. T., & Chu, C. L. (2005). Correlation between different memory systems and striatal dopamine D2/D3 receptor density: a single photon emission computed tomography study. *Psychological Medicine*, *35*(2), 197–204. https://doi.org/10.1017/S0033291704003101

Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, *17*(7), 438–449. https://doi.org/10.1038/nrn.2016.43

Crowe, S. F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. Journal of clinical psychology, 54(5), 585-591.

Dang, L. C., O'Neil, J. P., & Jagust, W. J. (2012). Dopamine Supports Coupling of Attention-Related Networks. *Journal of Neuroscience*, *32*(28), 9582–9587. https://doi.org/10.1523/JNEUROSCI.0909-12.2012

Dang, Linh C., O'Neil, J. P., & Jagust, W. J. (2013). Genetic effects on behavior are mediated by neurotransmitters and large-scale neural networks. *NeuroImage*, *66*, 203–214. https://doi.org/10.1016/j.neuroimage.2012.10.090

Dreher, J.-C., Meyer-Lindenberg, A., Kohn, P., & Berman, K. F. (2008). Age-related changes in midbrain dopaminergic regulation of the human reward system. *Proceedings of the National Academy of Sciences*, *105*(39), 15106–15111. https://doi.org/10.1073/pnas.0802127105

Eapen, M., Zald, D. H., Gatenby, J. C., Ding, Z., & Gore, J. C. (2011). Using high-resolution MR imaging at 7T to evaluate the anatomy of the midbrain dopaminergic system. *American Journal of Neuroradiology*, *32*(4), 688-694.

Erixon-Lindroth, N., Farde, L., Robins Wahlin, T.-B., Sovago, J., Halldin, C., & Bäckman, L. (2005). The role of the striatal dopamine transporter in cognitive aging. *Psychiatry Research: Neuroimaging*, *138*(1), 1–12. https://doi.org/10.1016/j.pscychresns.2004.09.005

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. (1995). Spatial registration and normalization of images. *Human brain mapping, 3*(3), 165-189.

Garrett, D. D., Nagel, I. E., Preuschhof, C., Burzynska, A. Z., Marchner, J., Wiegert, S., ... & Heekeren, H. R. (2015). Amphetamine modulates brain signal variability and working memory in younger and older adults. *Proceedings of the National Academy of Sciences*, *112*(24), 7593-7598. https://doi.org/10.1073/pnas.1504090112

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, *13*(7), 491.

Halldin, C., Farde, L., Högberg, T., Mohell, N., Hall, H., Suhara, T., ... & Swahn, C. G. (1995). Carbon-11-FLB 457: a radioligand for extrastriatal D2 dopamine receptors. *Journal of nuclear medicine*, *36*(7), 1275-1281.

Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., … Duncan, J. S. (2003). Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping*, *19*(4), 224–247. https://doi.org/10.1002/hbm.10123

Hicks, L. H., & Birren, J. E. (1970). Aging, brain damage, and psychomotor slowing. *Psychological Bulletin*, *74*(6), 377–396. https://doi.org/10.1037/h0033064

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107–129.

Hutton, B. F., Thomas, B. A., Erlandsson, K., Bousse, A., Reilhac-Laborde, A., Kazantsev, D., ... & Ourselin, S. (2013). What approach to brain partial volume correction is best for PET/MRI?. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *702*, 29-33.

Jagust, W., & D'Esposito, M. (Eds.). (2009). *Imaging the aging brain*. New York.: Oxford University Press.

Jin, X., Chan, C., Mulnix, T., Panin, V., Casey, M. E., Liu, C., & Carson, R. E. (2013). List-mode reconstruction for the Biograph mCT with physics modeling and event-by-event motion correction. *Physics in Medicine and Biology*, *58*(16), 5567–5591. https://doi.org/10.1088/0031-9155/58/16/5567

Karalija, N., Papenberg, G., Wåhlin, A., Johansson, J., Andersson, M., Axelsson, J., ... & Nyberg, L. (2019). C957T-mediated Variation in Ligand Affinity Affects the Association between 11C-raclopride Binding Potential and Cognition. *Journal of cognitive neuroscience*, 1-12.

Karlsson, S., Nyberg, L., Karlsson, P., Fischer, H., Thilers, P., MacDonald, S., … Bäckman, L. (2009). Modulation of striatal dopamine D1 binding by cognitive processing. *NeuroImage*, *48*(2), 398–404. https://doi.org/10.1016/j.neuroimage.2009.06.030

Karrer, T.M., Josef, A.K., Mata, R., Morris, E.D., Samanez-Larkin, G.R. (2017) Reduced dopamine receptors and transporters but not synthesis capacity in normal aging adults: a meta-analysis. *Neurobiology of Aging*, 57, 36–46.

Lammertsma, A. A., & Hume, S. P. (1996). Simplified Reference Tissue Model for PET Receptor Studies. *NeuroImage*, *4*(3), 153–158. https://doi.org/10.1006/nimg.1996.0066

Landau, S. M., Lal, R., O'Neil, J. P., Baker, S., & Jagust, W. J. (2009). Striatal Dopamine and Working Memory. *Cerebral Cortex*, *19*(2), 445–454. https://doi.org/10.1093/cercor/bhn095

Lappin, J. M., Reeves, S. J., Mehta, M. A., Egerton, A., Coulson, M., & Grasby, P. M. (2009). Dopamine Release in the Human Striatum: Motor and Cognitive Tasks Revisited. *Journal of Cerebral Blood Flow & Metabolism*, *29*(3), 554–564. https://doi.org/10.1038/jcbfm.2008.146

Lövdén, M., Karalija, N., Andersson, M., Wåhlin, A., Jan Axelsson, Köhncke, Y., … Lindenberger, U. (2017). Latent-Profile Analysis Reveals Behavioral and Brain Correlates of Dopamine-Cognition Associations. *Cerebral Cortex*, 1–14. https://doi.org/10.1093/cercor/bhx253

Luciana, M., Depue, R. A., Arbisi, P., & Leon, A. (1992). Facilitation of Working Memory in Humans by a D_2 Dopamine Receptor Agonist. *Journal of Cognitive Neuroscience*, *4*(1), 58–68. https://doi.org/10.1162/jocn.1992.4.1.58

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the N-Back Task a Valid Neuropsychological Measure for Assessing Working Memory? *Archives of Clinical Neuropsychology*, *24*(7), 711–717. https://doi.org/10.1093/arclin/acp063

Mozley, L. H., Gur, R. C., Mozley, P. D., & Gur, R. E. (2001). Striatal Dopamine Transporters and Cognitive Functioning in Healthy Men and Women. *American Journal of Psychiatry*, *158*(9), 1492–1499. https://doi.org/10.1176/appi.ajp.158.9.1492

Mukherjee, J., Christian, B. T., Dunigan, K. A., Shi, B., Narayanan, T. K., Satter, M., & Mantil, J. (2002). Brain imaging of 18F-fallypride in normal volunteers: Blood analysis, distribution, test-retest studies, and preliminary assessment of sensitivity to aging effects on dopamine D-2/D-3 receptors. *Synapse*, *46*(3), 170–188. https://doi.org/10.1002/syn.10128

Nevalainen, N., Riklund, K., Andersson, M., Axelsson, J., Ögren, M., Lövdén, M., ... & Nyberg, L. (2015). COBRA: A prospective multimodal imaging study of dopamine, brain structure and function, and cognition. *Brain research, 1612*, 83-103.

Nyberg, L., Karalija, N., Salami, A., Andersson, M., Wåhlin, A., Kaboovand, N., ... & Garrett, D. D. (2016). Dopamine D2 receptor availability is linked to hippocampal–caudate functional connectivity and episodic memory. *Proceedings of the National Academy of Sciences, 113*(28), 7918-7923.

Park, D.C., & Schwarz, N. (Eds.). (2000). *Cognitive Aging: A Primer*. Psychology Press.

Parmenter, B. A., Shucard, J. L., Benedict, R. H., & Shucard, D. W. (2006). Working memory deficits in multiple sclerosis: Comparison between the n-back task and the Paced Auditory Serial Addition Test. *Journal of the International Neuropsychological Society, 12*(5), 677-687.

Plavén-Sigray, P., Hedman, E., Victorsson, P., Matheson, G. J., Forsberg, A., Djurfeldt, D. R., ... & Cervenka, S. (2017). Extrastriatal dopamine D2-receptor availability in social anxiety disorder. *European Neuropsychopharmacology*, *27*(5), 462-469.

Raz, N., & Lindenberger, U. (2011). Only Time will Tell: Cross-sectional Studies Offer no Solution to the Age-Brain-Cognition Triangle—Comment on. *Psychological Bulletin*, *137*(5), 790–795. https://doi.org/10.1037/a0024503

Reeves, S. J., Grasby, P. M., Howard, R. J., Bantick, R. A., Asselin, M.-C., & Mehta, M. A. (2005). A positron emission tomography (PET) investigation of the role of striatal dopamine (D2) receptor availability in spatial cognition. *NeuroImage*, *28*(1), 216–226. https://doi.org/10.1016/j.neuroimage.2005.05.034

Rieckmann, A., Johnson, K. A., Sperling, R. A., Buckner, R. L., & Hedden, T. (2018). Dedifferentiation of caudate functional connectivity and striatal dopamine transporter density predict memory change in normal aging. *Proceedings of the National Academy of Sciences*, *115*(40), 10160–10165. https://doi.org/10.1073/pnas.1804641115

Rousset, O. G., Ma, Y., & Evans, A. C. (1998). Correction for partial volume effects in PET: principle and validation. *Journal of Nuclear Medicine*, *39*(5), 904-911.

Rubin, D. C. (1999). Frontal-Striatal Circuits in Cognitive Aging: Evidence for Caudate Involvement. *Aging, Neuropsychology, and Cognition*, *6*(4), 241–259. https://doi.org/10.1076/1382-5585(199912)06:04;1-B;FT241

Salami, A., Rieckmann, A., Karalija, N., Avelar-Pereira, B., Andersson, M., Wåhlin, A., … Nyberg, L. (2018). Neurocognitive Profiles of Older Adults with Working-Memory Dysfunction. *Cerebral Cortex*, *28*(7), 2525–2539. https://doi.org/10.1093/cercor/bhy062

Salami, A., Garrett, D. D., Wåhlin, A., Rieckmann, A., Papenberg, G., Karalija, N., … Nyberg, L. (2019). Dopamine $D_{2/3}$ Binding Potential Modulates Neural Signatures of Working Memory in a Load-Dependent Fashion. *The Journal of Neuroscience*, *39*(3), 537–547. https://doi.org/10.1523/JNEUROSCI.1493-18.2018

Salthouse, T. A. (2004). What and When of Cognitive Aging. *Current Directions in Psychological Science*, *13*(4), 140–144. https://doi.org/10.1111/j.0963-7214.2004.00293.x

Sandiego, C. M., Gallezot, J.-D., Lim, K., Ropchan, J., Lin, S., Gao, H., … Cosgrove, K. P. (2015). Reference region modeling approaches for amphetamine challenge studies with [11C]FLB 457 and PET. *Journal of Cerebral Blood Flow & Metabolism*, *35*(4), 623–629. https://doi.org/10.1038/jcbfm.2014.237

Sawaguchi, T. (2001). The effects of dopamine and its antagonists on directional delay-period activity of prefrontal neurons in monkeys during an oculomotor delayed-response task. *Neuroscience Research*, *41*(2), 115–128. https://doi.org/10.1016/S0168-0102(01)00270-X

Sawaguchi, T., & Goldman-Rakic, P. S. (1991). D1 Dopamine Receptors in Prefrontal Cortex: Involvement in Working Memory. *Science; Washington*, *251*(4996), 947.

Seaman, K. L., Smith, C. T., Juarez, E. J., Dang, L. C., Castrellon, J. J., Burgess, L. L., ... & Samanez-Larkin, G. R. (2019). Differential regional decline in dopamine receptor availability across adulthood: Linear and nonlinear effects of age. *Human brain mapping*, *40*, 3125– 3138. https://doi.org/10.1002/hbm.24585

Smith, C. T., Crawford, J. L., Dang, L. C., Seaman, K. L., San Juan, M. D., Vijay, A., … Samanez-Larkin, G. R. (2017). Partial-volume correction increases estimated dopamine D2-like receptor binding potential and reduces adult age differences. *Journal of Cerebral Blood Flow & Metabolism*, 0271678X1773769. https://doi.org/10.1177/0271678X17737693

Smith, C. T., Dang, L. C., Cowan, R. L., Kessler, R. M., & Zald, D. H. (2016). Variability in paralimbic dopamine signaling correlates with subjective responses to d-amphetamine. *Neuropharmacology*, *108*, 394–402. https://doi.org/10.1016/j.neuropharm.2016.05.004

Vandehey, N. T., Moirano, J. M., Converse, A. K., Holden, J. E., Mukherjee, J., Murali, D., … Christian, B. T. (2010). High-Affinity Dopamine D_2 /D_3 PET Radioligands ^{18}F-Fallypride and ^{11}C-FLB457: A Comparison of Kinetics in Extrastriatal Regions Using a Multiple-Injection Protocol. *Journal of Cerebral Blood Flow & Metabolism*, *30*(5), 994–1007. https://doi.org/10.1038/jcbfm.2009.270

Volkow, N. D., Gur, R. C., Wang, G. J., Fowler, J. S., Moberg, P. J., Ding, Y. S., ... & Logan, J. (1998). Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. American Journal of psychiatry, 155(3), 344-349.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: *Cognitive, Affective, & Behavioral Neuroscience*, *3*(4), 255–274. https://doi.org/10.3758/CABN.3.4.255

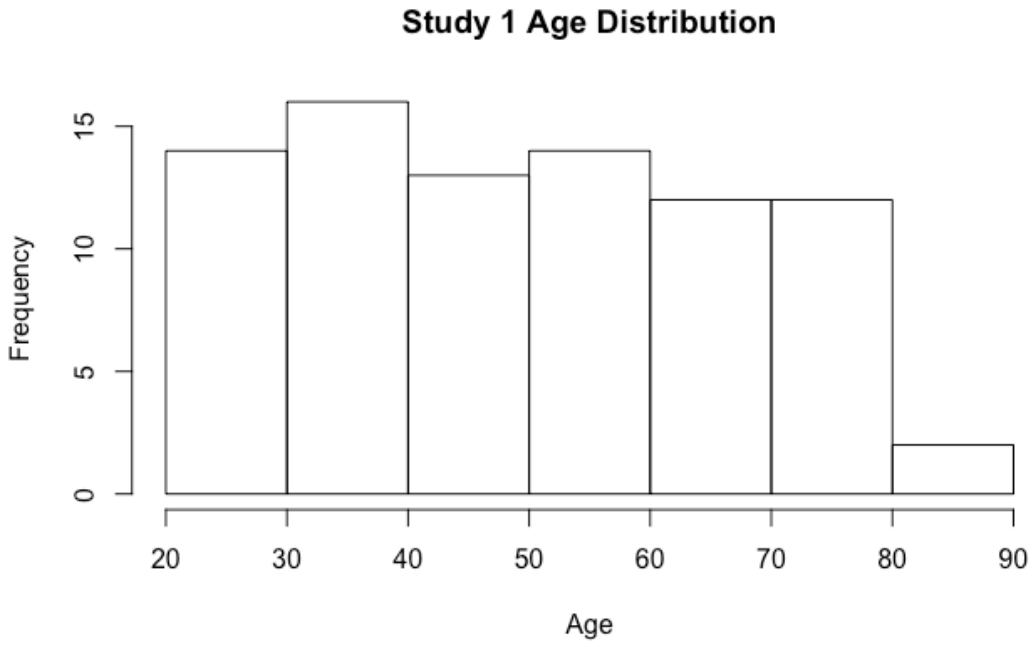
Wang, Y., Chan, G. L. Y., Holden, J. E., Dobko, T., Mak, E., Schulzer, M., … Stoessl, A. J. (1998). Age-dependent decline of dopamine D1 receptors in human brain: A PET study. *Synapse*, *30*(1), 56–61. https://doi.org/10.1002/(SICI)1098-2396(199809)30:1<56::AID-SYN7>3.0.CO;2-J

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292. https://doi.org/10.1037/0033-2909.120.2.272

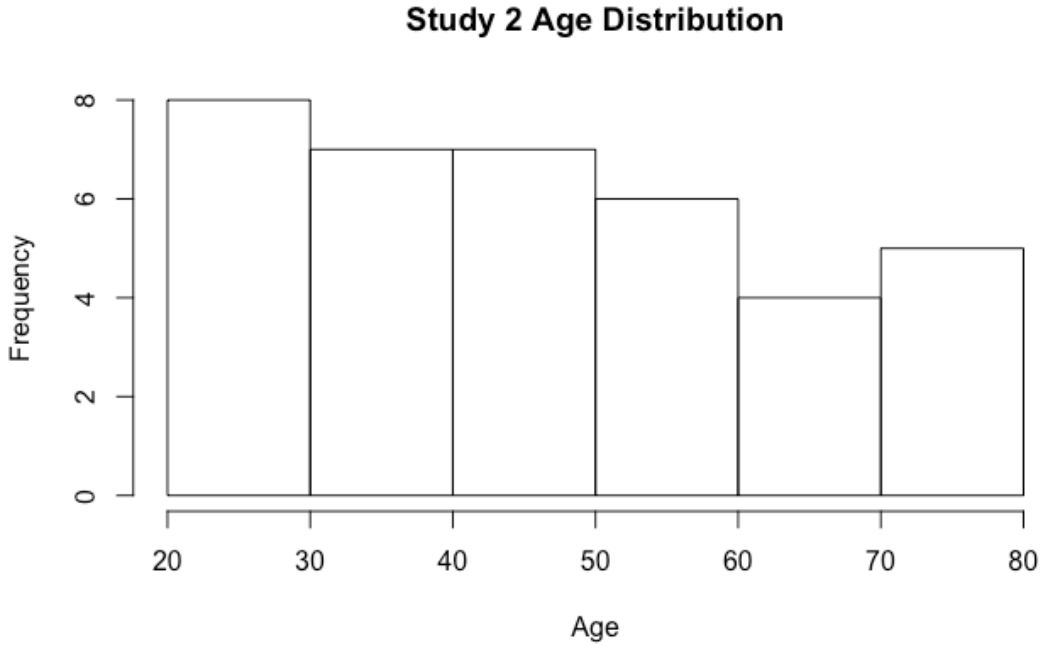
Zald, D. H., Woodward, N. D., Cowan, R. L., Riccardi, P., Ansari, M. S., Baldwin, R. M., … Kessler, R. M. (2010). The interrelationship of dopamine D2-like receptor availability in striatal and extrastriatal brain regions in healthy humans: A principal component analysis of [18F]fallypride binding. *NeuroImage*, *51*(1), 53–62. https://doi.org/10.1016/j.neuroimage.2010.02.006

SUPPLEMENTARY MATERIALS
Juarez et al

Supplementary Figure 1. Age distribution for Study 1, N=83.



Supplementary Figure 2. Age distribution for Study 2, N=37.



Supplementary Table 1. Non-PVC region of interest analyses for Study 1. (Note: midbrain ROI reported within the paper is already non-PVC.) Correlations between cognitive test scores and D2R BP_{ND} controlling for age (N=83). * p < .05 uncorrected

	Trails A	Trails B-A	Digit Span	Letter-Number Sequencing	Verbal Paired Associates Delayed Recall
Striatum	**.277* [0.059,0.495]**	–.131 [–0.378,0.116]	**.240* [0.001,0.480]**	.033 [–0.179,0.246]	.141 [–0.061,0.343]
Anterior cingulate	**.221* [0.014,0.428]**	–.154[–0.385,0.078]	.007 [–0.224,0.238]	–.124 [–0.322,0.073]	.002 [–0.191,0.195]
Thalamus	.105 [–0.100,0.310]	–.132 [–0.358,0.093]	.044 [–0.179,0.268]	–.063 [–0.256,0.131]	–.016 [–0.204,0.173]
Amygdala	.087 [–0.122,0.296]	.101 [–0.129,0.330]	.006 [–0.234,0.222]	–.107[–0.302,0.089]	.008 [–0.183,0.199]
Hippocampus	.064 [–0.136,0.264]	.006 [–0.215,0.226]	–.003 [–0.243,0.191]	–.114 [–0.301,0.072]	–.072 [–0.254,0.109]
Insula	**.239* [0.023,0.454]**	.018 [–0.226,0.262]	.091 [–0.149, 0.331]	–.065 [–0.273, 0.143]	–.062 [–0.264,0.139]

Supplementary Table 2. Non-PVC region of interest analyses for Study 1. (Note: midbrain ROI reported within the paper is already non-PVC.) Correlations between age and D2R BP_{ND} (N=83). * p < .05 uncorrected

	Age
Striatum	**–0.511*, [–0.698, –0.323], p < .0001**
Anterior cingulate	**–0.407*, [–0.606, –0.208], p = .0001**
Thalamus	**–0.339*, [–0.544, –0.134], p = .002**
Amygdala	**–0.378*, [–0.580, –0.177], p = .0004**
Hippocampus	**–0.247*, [–0.458, –0.036], p = .024**
Insula	**–0.482*, [–0.673, –0.292], p <.0001**

Supplementary Table 3. Correlations between age and cognitive tests before (first row after *Simple effect of age*) and after controlling for D2R BP_{ND} in Study 1 (N=83) *p<.05 uncorrected

	Trails A	Trails B-A	Digit Span	Letter-Number Sequencing	Verbal Paired Associates Delayed Recall
Simple effect of age	**.461* [0.268, 0.665]**	**.222* [0.009,0.434]**	**–.274* [–0.483,–0.064]**	**–.554* [–0.735,–0.372]**	**–.614* [–0.790,–0.438]**
Age effect controlling for regional D2 BP_{ND} in…					
Midbrain	**.561* [0.340,0.782]**	.238 [–0.009,0.485]	–.215 [–0.458,0.028]	**–.606* [–0.816,–0.396]**	**–.618* [–0.824,–0.411]**
Striatum	**.583* [0.373,0.792]**	.167 [–0.071,0.404]	–.165 [–0.395,0.065]	**–.536* [–0.740,–0.332]**	**–.552* [–0.747,–0.358]**
Anterior cingulate	**.491* [0.294,0.688]**	**0.225* [0.007, 0.444]**	**–.275* [–0.491,–0.059]**	**–.572* [–0.758,–0.387]**	**–.590* [–0.770,–0.411]**
Thalamus	**.490* [0.291,0.690]**	.186 [–0.034,0.405]	**–.263* [–0.481,–0.045]**	**–.570* [–0.758,–0.381]**	**–.619* [–0.803,–0.436]**
Amygdala	**.486* [0.284,0.688]**	**.250* [0.028,0.473]**	**–.027* [–0.487,–0.046]**	**–.597* [–0.785,–0.409]**	**–.603* [–0.788,–0.418]**
Hippocampus	**.467* [0.269,0.666]**	**.227* [0.009,0.445]**	**–.276* [–0.491,–0.060]**	**–.572* [–0.757,–0.387]**	**–.624* [–0.804,–0.443]**
Insula	**.547* [0.346,0.749]**	.232 [0.003,0.461]	**–.254* [–0.479, –0.029]**	**–.587* [–0.782, –0.393]**	**–.641* [–0.830,–0.452]**

Supplementary Table 4. Non-PVC region of interest analyses for Study 2. (Note: midbrain ROI reported within the paper is already non-PVC.) Correlations between cognitive test scores and D2R BP_{ND} controlling for age (N=37). * p < .05 uncorrected

	Trails A	Trails B-A	Digit Span	Letter-Number Sequencing	Verbal Paired Associates Delayed Recall
Anterior cingulate	–.008 [–0.340,0.355]	.172 [–0.159,0.504]	–.135[–0.520,0.249]	–.285 [–0.629,0.059]	–.043 [–0.374,0.288]
Thalamus	–.096 [–0.405,0.214]	.214 [–0.078,0.507]	–.139 [–0.482,0.205]	–.227 [–0.538,0.083]	.133 [–0.160,0.426]
Amygdala	–.214 [–0.531,0.103]	.198 [–0.110,0.506]	–.103 [–0.463,0.258]	**–.347* [–0.660,–0.034]**	–.148 [–0.454,0.158]
Hippocampus	.202 [–0.093,0.498]	.210 [–0.074,0.495]	–.029 [–0.366,0.308]	–.223 [–0.525,–0.079]	.148 [–0.136,0.433]
Insula	–.003 [–0.346,0.340]	.309 [–0.006, 0.624]	–.195 [–0.571,0.181]	–.198 [–0.543,0.148]	–.082 [–0.407,0.244]

Supplementary Table 5. Non-PVC region of interest analyses for Study 2. (Note: midbrain ROI reported within the paper is already non-PVC.) Correlations between age and D2R BP_{ND} (N=37). * p < .05 uncorrected

	Age
Anterior cingulate	**–0.500*, [–0.787, –0.213], p = .002**
Thalamus	–0.253, [–0.573, 0.068], p = .132
Amygdala	**–0.378*, [–0.685, –0.072], p = .021**
Hippocampus	–0.112, [–0.441, 0.218], p = .511
Insula	**–0.479*, [–0.770, –0.188], p = .003**

Supplementary Table 6. Correlations between age and cognitive tests before (first row after *Simple effect of age*) and after controlling for D2R BP_{ND} in Study 2 (N=37) *p<.05 uncorrected

	Trails A	Trails B-A	Digit Span	Letter-Number Sequencing	Verbal Paired Associates Delayed Recall
Simple effect of age	**.445* [0.149, 0.742]**	**.499* [0.211,0.786]**	–.070 [–0.401,0.260]	**–.393* [–0.697,–0.088]**	**–.521* [–0.803,–0.238]**
Age effect controlling for regional D2 BP_{ND} in…					
Midbrain	.260 [–0.120,0.640]	**.748* [0.392,1.104]**	–.311 [–0.729,0.106]	**–.683* [–1.055,–0.310]**	**–.629* [–0.998,–0.260]**
Anterior cingulate	**.442* [0.099,0.784]**	**0.589* [0.263, 0.914]**	–.107 [–0.488,0.274]	**–.480* [–0.827,–0.134]**	**–.544* [–0.870,–0.217]**
Thalamus	**.425* [0.118,0.732]**	**.545* [0.255,0.835]**	–.102 [–0.442,0.238]	**–.442* [–0.749,–0.134]**	**–.491* [–0.782,–0.201]**
Amygdala	**.380* [0.067,0.694]**	**.566* [0.264,0.869]**	–.098 [–0.453,0.258]	**–.499* [–0.810,–0.188]**	**–.573* [–0.874,–0.273]**
Hippocampus	**.447* [0.142,0.752]**	**.521* [0.230,0.813]**	–.094 [–0.430,0.243]	**–.445* [–0.738,–0.151]**	**–.507* [–0.800,–0.218]**
Insula	**.452* [0.117,0.787]**	**.634* [0.327,0.941]**	–.151 [–0.519, 0.217]	**–.474* [–0.812, –0.136]**	**–.556* [–0.874,–0.238]**