

The role of the striatum in learning to orthogonalize action and valence: a combined PET and 7T MRI aging study

Valentina Perosa^{1,2}, Lieke de Boer³, Gabriel Ziegler^{1,2}, Ivayla Apostolova⁵, Ralph Buchert⁵, Coraline Metzger^{1,2}, Holger Amthauer⁶, Marc Guitart-Masip^{3,4}, *Emrah Düzel^{1,2,7}, *Matthew J. Betts^{1,2}

¹Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

²German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany

³Ageing Research Centre, Karolinska Institute, SE-11330 Stockholm, Sweden

⁴Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London WC1B 5EH, UK

⁵Department of Radiology and Nuclear Medicine, University Hospital Magdeburg, Germany

⁶Department of Nuclear Medicine, Charité–Universitätsmedizin Berlin, Germany

⁷Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK

* Both authors contributed equally to this work.

Corresponding author: Matthew Betts

Corresponding author's address: Deutsches Zentrum für Neurodegenerative Erkrankungen e.V. (DZNE) c/o Universitätsklinikum Magdeburg, Leipziger Strasse 44, Haus 64, 39120 Magdeburg, Deutschland

Corresponding author's e-mail address: matthew.betts@dzne.de

Abstract

Pavlovian biases influence instrumental learning by coupling reward seeking with action invigoration and punishment avoidance with action suppression. Using a probabilistic go/no-go task designed to orthogonalize action (go/no-go) and valence (reward/punishment), recent studies have shown that the interaction between the two is dependent on the striatum and its key neuromodulator dopamine. Using this task, we sought to identify how structural and neuromodulatory age-related differences in the striatum may influence Pavlovian biases and instrumental learning in 25 young and 31 older adults. Computational modeling revealed a significant age-related reduction in *reward* and *punishment sensitivity* and marked (albeit not significant) reduction in *learning rate* and *lapse rate (irreducible noise)*. Voxel-based morphometry analysis using 7 Tesla MRI images showed that individual differences in *learning rate* in older adults were related to the volume of the caudate nucleus. In contrast dopamine synthesis capacity in the dorsal striatum, assessed using [¹⁸F]-DOPA positron emission tomography in 22 of these older adults, was not associated with learning performance and did not moderate the relationship between caudate volume and *learning rate*. This multiparametric approach suggests that age-related differences in striatal volume may influence learning proficiency in old age.

Keywords: 7T structural MRI, aging, dopamine, instrumental learning, striatum

Introduction

Hardwired interactions between action and valence in learning tasks significantly influence choices. These behavioral tendencies can be described as Pavlovian biases and serve to accelerate the acquisition of the appropriate behavioral responses in those circumstances most commonly encountered during decision making, yet may also corrupt the flexibility of learning (Dayan et al. 2006; Gray and Mcnaughton 2000). This is in contrast to instrumental learning, which a priori should be based on the contingency between behaviour and outcome, that is the stimulus-reinforcement pair and not the response-reinforcement association (Williams 1987).

Achieving a reward by performing an action and avoiding punishment by remaining passive is typically performed more successfully than not initiating an action to gain rewards or performing actions to avoid punishment (Guitart-masip et al. 2012a, de Boer et al. 2018). This bias has been previously demonstrated using a probabilistic monetary go/no-go task that orthogonalizes, i.e. independently manipulates, action (go/no-go) and valence (win/avoid-losing) (Guitart-Masip et al. 2011). Furthermore, computational modeling has shown that this behavioral moderation can be accounted for in terms of an influence of a Pavlovian system, effectively coupling action and valence (Guitart-masip et al. 2012). However, it has been recently argued that an instrumental bias may also contribute to this learning asymmetry (Swart et al. 2017) implying a facilitated learning of go responses leading to reward, as well as an impaired unlearning of no-go responses leading to punishment. Thus both Pavlovian and instrumental mechanisms may contribute to action-valence learning whereby instrumental learning bias is implemented at the time participants incorporate the observed outcome to their action values (Swart et al. 2017), and the Pavlovian bias is implemented at the time of choice whereby the expectation of value on a given trial is added to the action value for the go choice (Guitart-masip et al. 2012a; Swart et al. 2017; de Boer et al. 2019).

Dopamine is considered to be essential in the process of instrumental learning (Dayan and Niv 2008) by encoding reward-prediction errors (RPEs) (as reviewed by Schultz, 1998) and also plays an important role in the generation and invigoration of motor responses, including actions directed towards rewards and in the avoidance of punishments (Niv et al. 2007; Salamone and Correa 2012). Furthermore, human behavioral and functional neuroimaging studies specifically designed to orthogonalize action and valence have challenged existing views that neural representations in the striatum represent valence, instead demonstrating a dominant role in the anticipation of action (for a review see (Guitart-Masip et al. 2014)). Moreover, by elevating dopamine using levodopa (L-DOPA), individuals showed a decrease in the coupling between action and valence during learning (Guitart-masip et al. 2014), but also an increase in functional activity in the striatum related to rewarding actions (Guitart-Masip et al. 2012b). These studies highlight the importance of orthogonalizing action and valence to identify cognitive and neuronal aspects of value representation and action selection. However, the exact contribution of striatal dopamine to instrumental and Pavlovian control remains unclear.

During aging, numerous cognitive domains (e.g. Grady 2013) including instrumental learning become impaired (Mell et al. 2005; Dreher et al. 2008). Age-related differences have been linked to functional activity in dopaminergic target regions such as the striatum and prefrontal cortex (Fera et al. 2005; Mell et al. 2009; Samanez-Larkin et al. 2010). Furthermore, both structural and neuromodulatory changes in the striatum have been previously shown to be associated with age-related decline in pursuing reward (Schott et al. 2007; Mohr et al. 2010; Di et al. 2014) and changes in the representation of RPE (Mell et al. 2009; Chowdhury 2013; Eppinger et al. 2013; Vink et al. 2015). As observed in previous studies designed to orthogonalize action and valence, performance in both young (Guitart-masip et al. 2012a;

Cavanagh et al. 2013; Guitart-Masip et al. 2013) and older adults (Chowdhury et al. 2013) is suboptimal in task conditions when Pavlovian and instrumental controllers conflict. As such, these two systems may be segregated and compete for behavioral control. Indeed, this notion is supported by evidence that instrumental and Pavlovian responses recruit different corticostriatal loops (Haber and Knutson 2010). The dorsal striatum is involved in learning and performance pertaining to goal-directed and habitual instrumental responding (Liljeholm and O'Doherty 2012) whereas the ventral striatum is more involved in Pavlovian learning (Reynolds and Berridge 2008; Corbit and Balleine 2011). However, no study to date has focused on how age-related structural and neuromodulatory differences in the striatum influence action learning in older age using a task that orthogonalizes action and valence.

In this study we hypothesized an age-related decrease in reinforcement learning using a probabilistic go/no-go task (Guitart-Masip et al. 2012), is associated with both an age-related decline in striatal volume and synthesis capacity of dopamine in older adults. Firstly, we assessed younger and older adults' performance in a task that orthogonalizes action and valence using a computational model with a constant Pavlovian bias parameter. In a second step, we assessed how structural age-related differences in the striatum may influence learning using high-resolution structural magnetic resonance imaging (MRI) at 7 Tesla (T) in younger and older adults. Finally, we used [^{18}F]-FDOPA-PET to assess how interindividual variability in dopamine synthesis capacity in the striatum, may impact performance in older adults.

Materials and Methods

Participants

A total of 56 participants were recruited as volunteers by our department, the German Center for Neurodegenerative Disease (DZNE), Magdeburg. We included 25 healthy young adults (mean age = 24.16, SD = 2.16; 12 females and 13 males) and 31 older adults between 62 and 78 years old (mean age = 68.58, SD = 4.50; 19 females and 12 males), who were previously screened for contraindications for 7 Tesla MRI scanning (tattoos, tinnitus, pacemaker, metallic implants, etc.). Furthermore, 22 of the older adults (mean age = 68.93, SD = 4.41; 10 females and 12 males) also underwent [¹⁸F]-FDOPA-PET. One subject that underwent FDOPA-PET declined to undergo the 7 Tesla MRI scan. All subjects performed an instrumental learning task (orthogonalized go/no-go task) and underwent a 7T MRI scan on the same day. On a separate visit, all older adults undertook a 3T MRI scan and were asked to complete a neuropsychological test battery, with the aim of excluding cognitive impairment and depression. The battery included a *Mini Mental State Test (MMSE)* (Folstein et al. 1975), (mean 29.6; SD 0.67), the *Stroop Test* (Glaser and Glaser 1982) and the *Logical Memory Test I and II* as part of the Wechsler memory scale. Furthermore, participants performed part of the *Test of Attentional Performance (TAP)*, which measured *alertness* and *divided attention* (Leclercq and Zimmermann 2004). In addition to these tests, volunteers were asked to complete the *Freiburger Personality Inventory* and *Beck Depression Inventory II* (Beck et al. 1996). One subject from the older group was excluded after showing a positive score in the depression screening (BDI II = 16 points) leaving a total of 30 older adults. All older adults were examined by a physician to exclude for any history of neurological or psychiatric disorders. All participants provided written consent according to the declaration of Helsinki and were compensated for transport costs and travel time. The study was approved by the Ethics Committee of the Faculty of Medicine, Otto-von-Guericke University of Magdeburg.

Go/no-go task

All subjects completed a probabilistic monetary go/no-go task as previously described (Guitart-Masip et al. 2012a). The goal of the task for the participant was to maximize reward and minimize punishment. The task consisted of four trial types depending on the identity of the fractal cue presented at the beginning of the trial:

1. Go to win (GW): to press a button in order to gain a reward
2. Go to avoid losing (GAL): to press a button in order to avoid punishment
3. No go to win (NGW): not pressing a button to gain reward
4. No go to avoid losing (NGAL): not pressing a button to avoid punishment

The participants wouldn't know at the beginning of the test, which image corresponded to which condition and had to learn them by trial and error. At each trial a cue appeared on the screen for 1000 ms and after an interval marked by a cross (250 – 3500 ms), the subject was presented with a circle and had to press the button (go) within 1500 ms (Figure 1), or withhold the action (no-go). The button to be pressed was an arrow on a keyboard. Participants had to press the left arrow if the circle appeared on the left side of the screen or the right arrow if it appeared on the right side of the screen. After each trial a feedback in the form of an arrow was shown: a downward red arrow signaled a loss of 1 Euro, a horizontal yellow bar represented a neutral outcome, an upward green arrow signified a reward of 1 Euro. After feedback was displayed, the cue for the following trial proceeded.

The task consisted of 240 trials and each image was shown 60 times (60 trials for each condition). The outcome was probabilistic, in win trials 80% of correct choices and 20% of incorrect choices were rewarded (the remaining 20% of correct and 80% of incorrect choices leading to no outcome), while in lose trials 80% of correct choices and 20% of incorrect choices avoided punishment. Before starting the task, the probabilistic nature of the task was explained to the participants in detail by showing a scheme of the probabilistic possibilities. A short training session was completed, in order to familiarize with the buttons and the speed of

the trials, but without showing the cues that would appear in the actual task. During the training session the participant was asked to complete 10 practice trials by pressing a button each time a target circle appeared on the left or right side of the screen. Participants were also informed that whilst their reward at the end of the task may be higher, they would receive a maximum reward of 15 Euros and a minimum of 5 Euros. Earnings were displayed at the end of the session.

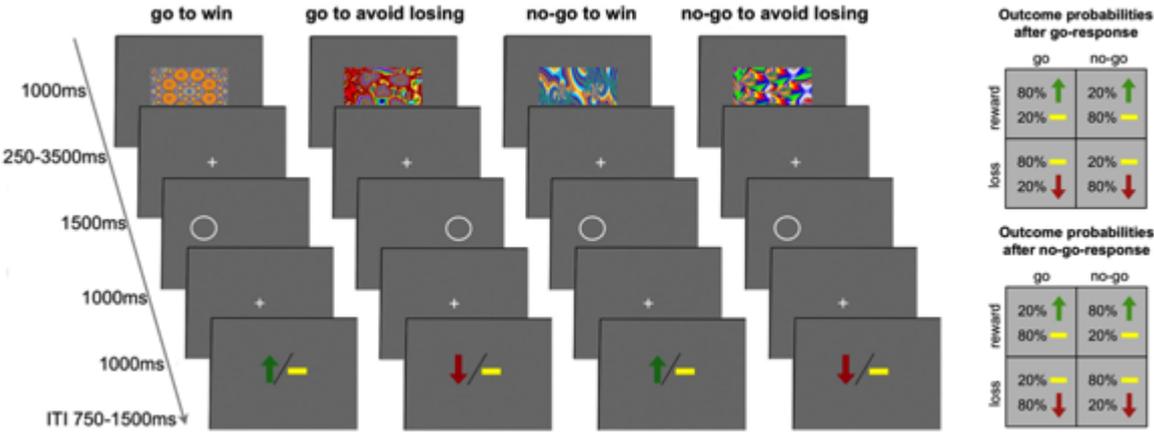


FIGURE 1 NEAR HERE

Figure 1. Paradigm of the probabilistic monetary go/no-go task. The fractal images indicate the respective four different conditions. On go trials, subjects needed to press a button according to the side where the circle appeared. On no-go trials they needed to withhold a response. An upward arrow symbolized reward (originally in green) and a downward arrow symbolized losses (originally in red). Horizontal bars (originally in yellow) corresponded to a neutral outcome. On the right-hand side, probability outcomes were displayed after go responses (go; top), and after withholding an action (no-go; bottom).

Statistical analysis of learning performance

In order to determine the differences of choice patterns between age groups, we performed a

logistic multi-level analysis using the lme4 package (Bates et al., 2014) in R 3.4.3 (“Kite-Eating Tree”) (R Development Core Team, 2015). The analysis estimated the probability of choosing a go response in the four task-conditions based on different predictors. Predictors for each trial included *valence* (win/avoid losing), *action* (go vs. no-go), as well as *time* (trial number) and *group* (young/older). Using the model with the best data fit, the fixed-effect predictors included all of the four predictors mentioned above, and their possible interactions. Subsequently, random effects for action, valence and trial and their possible interactions were included. This model was compared to other versions of the logistic regression model using the Bayesian Information Criterion using the R function *BIC*.

Computational modeling of learning behaviour using the go/no-go task

We fit choice behavior to a set of 6 nested reinforcement learning (RL) models incorporating different RL hypothesis. The base model was a Q-learning algorithm (Sutton and Barto 1998) that used a Rescorla-Wagner update rule to independently track the action value of each choice given each fractal image ($Q_i(\text{go})$ and $Q_i(\text{nogo})$) with learning rate (ϵ) as a free parameter. In the model, the probability of choosing one action on trial t was a sigmoid function of the difference between the action values scaled by a slope parameter that was parameterized as sensitivity to reward. This basic model was initially augmented with an irreducible action noise parameter also known as a lapse rate (ξ) (Talmi et al. 2009) and then further expanded by adding a static bias parameter to the value of the go action (b). The model was then augmented by adding a fixed Pavlovian value of 1 to the value of the go action as soon as the first reward was encountered for win cues, and a fixed Pavlovian value of -1 to the value of the go action as soon as the first punishment was encountered for loss cues. This fixed Pavlovian value was weighted by a further free parameter (Pavlovian parameter) into the value of the go action (π). Note that this definition of the Pavlovian value

is different from the definition in previous studies that have used this task (Guitart-masip et al. 2012a; de Boer et al. 2019), as model comparison demonstrated it a better fit than a variable Pavlovian value updated on a trial-by-trial basis. The state (action independent) values for each fractal image were updated on every trial using a Rescorla-Wagner update rule with the same learning rate as the update of the action values. Finally, the model including the static action bias and the Pavlovian bias were augmented by including different sensitivities for reward and punishment. Full equations are provided in the Supplemental Material.

Model fitting procedure and comparison

As in previous reports (Huys et al. 2011; Guitart-Masip et al. 2012a) we used a hierarchical Type II Bayesian (or random effects) procedure using maximum likelihood to fit simple parameterized distributions for higher level statistics of the parameters. Since the values of parameters for each subject are ‘hidden’, this employs the Expectation-Maximization (EM) procedure. For each iteration, the posterior distribution over the group for each parameter is used to specify the prior over the individual parameter fits on the next iteration. All six computational models were fit to the data using a single distribution for all participants. This fitting procedure was, therefore, blind to the existence of different groups with putatively different parameter values. Before inference, all parameters except the action bias were suitably transformed to enforce constraints (log and inverse sigmoid transforms). Six modeling parameters were extracted for each individual, namely *reward sensitivity*, *punishment sensitivity*, *Pavlovian bias*, *action bias*, *learning rate* and *lapse rate (irreducible noise)*.

Models were compared using the integrated Bayesian Information Criterion (iBIC) as described in detail in Huys et al. (2011) and Guitart-Masip et al. (2012a), where small iBIC values indicate a model that fits the data better after penalizing for the number of data points

associated with each parameter (Table 1). Comparing iBIC values is akin to a likelihood ratio test (Kass R. E. and Raftery A. E. 1995).

Non-parametric Mann-Whitney U tests were applied *ad-hoc* to assess for age-related differences in the modeling parameters. To correct for the effect of multiple comparisons, we applied a statistical threshold of $p < 0.008$ (i.e. Bonferroni-corrected $p < 0.05$ for 6 tests). Furthermore, we performed partial Pearson's correlations using age and sex as covariates, to investigate the relationship between modeling parameters and measures of global cognition assessed using the neuropsychology test battery. The significance threshold was set at $p < 0.05$ following Bonferroni correction for multiple comparisons as described in the Supplementary Material.

Table 1

Model no.	Model parameters	No. of parameters	Likelihood	Pseudo-R ²	iBIC
1	ε, ρ	2	-6,111	0.33	12,260
2	ε, ρ, ξ	3	-6098	0.33	12,254
3	$\varepsilon, \rho, \xi, b$	4	-5724	0.37	11,523
4	$\varepsilon, \rho_{win}, \rho_{lose}, \xi, b$	5	-5563	0.39	11,220
5	$\varepsilon, \rho_{win}, \rho_{lose}, \xi, b, \pi_{fluct}$	6	-5467	0.40	11,048
6	$\varepsilon, \rho_{win}, \rho_{lose}, \xi, b, \pi_{constant}$	6	-5431	0.41	10,975

Model comparison of the six models tested to account for the behavioral data. The winning model is highlighted in bold font. Parameters: ε , learning rate; ρ_{win} , weighting of reward on win trials; ρ_{lose} , weighting of punishments on lose trials; ξ , irreducible noise; b , action bias; π , Pavlovian bias; iBIC, integrated Bayesian information criterion.

Structural magnetic resonance imaging

All young adults ($n = 25$) and 30 older adults (after excluding 1 case for depression) underwent a T_1 -weighted high-field scan using 7 Tesla MRI, which provided high-resolution structural images of the whole brain. T_1 -weighted images at high-resolution reduce partial volume effects and allow accurate discrimination between grey matter (GM) and white matter (WM) used to delineate subcortical areas such as the striatum. The T_1 -weighted 3D-MPRAGE MRI data was acquired for each subject, using a Siemens MAGNETOM 7.0 Tesla MRI scanner and a Nova Medical (Willmington, MA) 32-channel head-coil. Each three-dimensional magnetisation-prepared rapid gradient echo (3D-MPRAGE; voxel size $0.8 \times 0.8 \times 0.8 \text{ mm}^3$) image was acquired using the following acquisition parameters: echo time (TE) was 2.09 ms, repetition time (TR) was 2000 ms and flip angle was 5° . Furthermore, inversion time was 1050 ms, receiver bandwidth was 230 Hz/pixel and echo spacing 6.1 ms. 3D matrix dimensions were $320 \times 320 \times 224$ (straight-sagittal slice orientation with 0.5-mm inter-slice gap), $7/8$ partial Fourier, and $0.8 \times 0.8 \times 0.8 \text{ mm}^3$ voxel size. GRAPPA was also enabled with acceleration factor of 2 and 32 reference lines.

In order to aid coregistration to PET images, all older participants that volunteered to undergo FDOPA PET also underwent a separate MRI session within a period of 12 months from the FDOPA PET scan. This was conducted on a Siemens Verio 3 Tesla system using a standard Siemens 32-channel phased-array head coil for reception. T_1 -weighted 3D-MPRAGE images were acquired using the following parameters: inversion time was set to 1100 ms, flip angle was 7° , time to echo was 4.37 ms, receiver bandwidth was 140 Hz/pixel, echo spacing was 11.1 ms, and repetition time was 2500 ms. The 3D matrix dimensions were $256 \times 256 \times 192$ (0.5 mm interslice gap), $7/8$ partial Fourier, and $1 \times 1 \times 1 \text{ mm}^3$ voxel size. GRAPPA was also enabled with acceleration factor of 2 and 24 reference lines

Voxel-based morphometry

Analysis and processing of 7T MPRAGE images were performed to assess the association between striatal volume and the task modeling parameters. We used the CAT12 toolbox (cat12 r938, Structural Brain Mapping group, Jena University Hospital, Jena, Germany) implemented in SPM12 (Statistical Parametric Mapping software; Wellcome Trust Centre for Neuroimaging, London, UK) run in MATLAB R2014b (Mathworks, Sherborn, MA, USA). In CAT12, all MPRAGE images were initially bias corrected and segmented into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF). Scans were visually inspected to exclude artefacts and then underwent a sample homogeneity check to identify potential outliers. Subsequently, GM and WM segments were warped to a common template using DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra) (Ashburner 2007). To account for the volume changes induced by normalisation, GM segments were multiplied by the Jacobian determinants of the deformations. Finally, GM and WM maps were smoothed using a Gaussian kernel at 6 mm full width at half maximum.

The voxel-based GM analysis was performed on all computational modeling parameters. A General linear model (GLM) was specified in SPM12, using the normalized GM segments and selected modeling parameter, as a regressor of interest, using the remaining modeling parameters, *sex* and *total intracranial volume* (TIV, calculated using CAT12) as covariates. The GM segments and the modeling parameters for each group were entered separately into a GLM to assess for age-related interactions between groups. All voxel-based hypothesized relationships were tested using t-tests generated in a GLM. All results were displayed using an uncorrected threshold of $p < 0.001$ without selecting a cluster threshold.

The initial VBM analyses were restricted to the striatum comprising the caudate, putamen and nucleus accumbens. The striatal regions of interest in 0.5 mm MNI152 space (Montreal Neurological Institute, McGill University, Canada) were obtained from FSL version 5.0 (<http://fsl.fmrib.ox.ac.uk>) and co-registered to the template image provided in CAT12. In a

further exploratory analysis we expanded the analysis to include the fronto-striatal network, comprising prefrontal cortex, cingulate cortex as well as the surrounding medial cortex, pallidum, substantia nigra/ventral tegmental area (SN/VTA) and the hippocampus. The fronto-striatal mask was created using segmentations from the built-in neuromorphometrics atlas in CAT12.

[¹⁸F]-FDOPA-PET

PET data from 22 older participants were acquired with a Biograph mCT (Siemens) PET/CT scanner in 3D mode. After a low-dose transmission CT for attenuation correction, a list-mode emission recording lasting 60 min was started simultaneously with i.v. injection of 200 MBq of F-18-FDOPA as a slow bolus. The emission data were reconstructed into 20 dynamic frames (3×20 s, 3×1 min, 3×2 min, 3×3 min, 7×5 min, 1×6 min) using the iterative reconstruction algorithm of the system software and isotropic voxel size of 2 mm.

PET data from 22 older adults were analysed fully automatically using a MATLAB/SPM8 script comprising the following steps. First, reconstructed frames were converted from DICOM to NIFTI format using the `dcm2nii` routine of the MRICron software package (<https://www.nitrc.org/projects/mricron>). Second, frames 7 - 20 (4 - 60 min p.i.) were corrected for head motion between frames using the 'Realign' tool in SPM8 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, UK). The realign transformation of frame 7 was applied to frames 1 - 6 (0 - 4 min p.i.), as these early frames did not provide sufficient anatomical information to reliably determine head motion. Third, the dynamic PET frames were coregistered to the individual 3D-MPRAGE MRI using the 'Coregister' tool of SPM8. The integral of PET frames 7 - 20 was used as source image, the 3D-MPRAGE MRI was used as target image for coregistration. The resulting rigid body transformation was applied to each of the realigned PET frames. Fourth, time activity curves

of caudate nucleus, putamen and nucleus accumbens, separately for left and right hemisphere, were determined by transferring the contours from automatic segmentation of these striatal regions in the subject's 3D-MPRAGE MRI using the FIRST algorithm (Patenaude et al. 2011) in FSL version 5.0 from the MRI to the coregistered PET frames. The time activity curve of the cerebellum (excluding vermis) was extracted using a mask from the WFU Pick Atlas (Maldjian et al. 2003). The inverse of the elastic transformation from subject space to the anatomical space of the Montreal Neurological Institute (MNI) obtained by the 'normalization' tool of SPM8 was used to map the cerebellum mask to the subject's PET frames. Fifth, dopamine synthesis capacity (K_i) was estimated for each striatal subregion by the slope of the tissue slope-intercept plot of its time activity curve (Patlak and Blasberg 1985). The cerebellum (excluding vermis) was used as reference region. Frames recorded between 5 and 60 min were used for the linear fit (Hoshi et al.).

In order to investigate the relationship between behavioral performance, modeling parameters and regional dopamine synthesis capacity (K_i), partial Pearson's correlations using age and sex as covariates, were performed using a significant threshold of $p < 0.008$ (i.e. Bonferroni-corrected $p < 0.05$ for 6 tests). All analyses were performed in Statistical Package for Social Science (IBM SPSS statistics), version 23.

Results

Behavioral Results

Overall, participants were better at learning go responses compared to no-go responses evidenced by increased responding to go vs. no-go cues over trials ($z = 61.1$; $p < 0.0001$). As expected, cue valence (win vs lose cues) also influenced responding ($z = -26.6$; $p < 0.0001$).

Participants were better at learning go and no-go responses coupled to wins compared to losses evident by the significant interaction between *action* and *valence* ($z = 22.0$; $p < 0.0001$). Furthermore a significant difference in learning go vs. no-go cues was observed between groups whereby older adults demonstrated a significant decrease in go vs. no-go learning ($z = -43.5$; $p < 0.0001$) compared to younger adults. A significant age-related difference was observed between groups for the interaction between action and valence, whereby older adults demonstrated reduced preference for learning go responses coupled to wins vs. losses ($z = -14.2$; $p < 0.0001$) compared to younger adults.

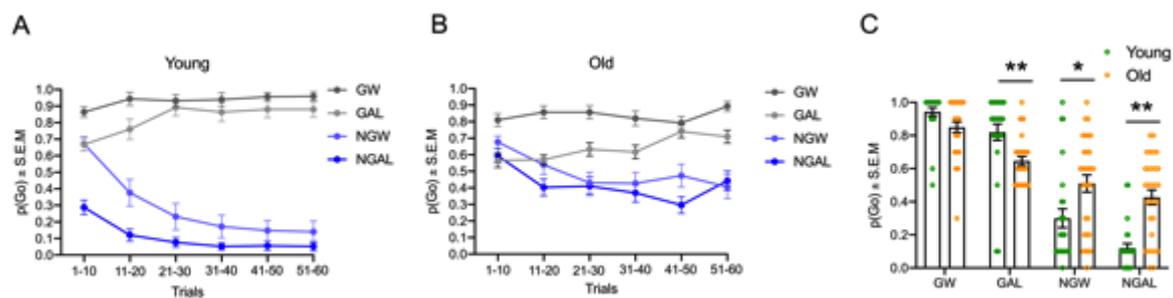


FIGURE 2 NEAR HERE

Figure 2. Go/no-go task performance in young and older adults: Proportion of “go” responses across the task is shown respectively for younger (A) and older (B) participants for each condition. Young participants learnt better to make active choices to gain rewards (go to win), than to avoid losing (go to avoid losing); the proportion of inappropriate “go” responses in “no-go” conditions was furthermore higher across the task when inaction was necessary to gain a reward (no-go to win), than to avoid punishment (no go to avoid losing) (A). Older adults also learnt better to make active choices to gain rewards; however unlike younger adults, demonstrated similar learning curves for passive (no-go) conditions

independent of valence (B). Proportion of “go” responses (+/- s.e.m.) in the four task conditions indicating significant differences between young and older adults. Values are mean \pm s.e.m.

Computational modeling parameters

The previously described reinforcement learning (Q-learning) model was used to fit performance of this task on the current data set. The parameters' medians of the two age groups for the 6 parameters of the winning model were compared using non-parametric Mann Whitney U tests (see Figure 3 and Supplementary Table 1), demonstrating a significant age-related difference in *reward sensitivity* ($p < 0.001$) and *punishment sensitivity* ($p < 0.001$). Age-related differences in *learning rate* ($p = 0.024$) and *lapse rate* ($p = 0.032$) were also observed, however these effects did not survive Bonferroni correction for multiple comparisons (corrected significant p value < 0.008). To investigate whether differences in the modeling parameters were related to more intact cognitive function in older adults, we correlated each parameter with neuropsychological measures (namely the Stroop Test, Logical Memory I and II and test of attentional performance revealing no significant results Supplementary tables 2 – 7). Furthermore no significant correlation between each modeling parameter was observed taking the whole sample together (Supplementary table 8a) or within the young and older groups separately (Supplementary table 8b and 8c)

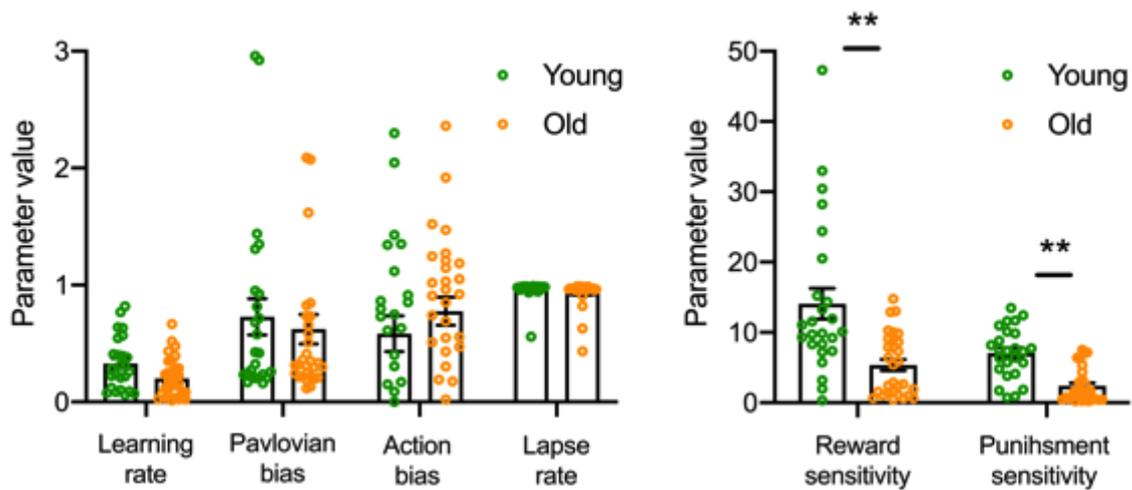


FIGURE 3 NEAR HERE

Figure 3. Comparison of modeling parameters between young and older adults. Reward and punishment sensitivity are shown on a different scale. A non-parametric analysis was performed using Mann Whitney U tests to assess for age-related differences in modeling parameters across groups. A significant age-related difference in punishment and reward sensitivity was observed between groups (** $p < 0.01$ following Bonferroni correction for multiple comparisons). Values indicate the mean and interquartile range.

Analysis of brain morphometric differences

The voxel-wise comparison between young and older adults was performed using the computational modeling parameters. Using a mask restricting the analysis to the striatum, we observed that differences caudate nucleus volume between young and older adults was related to *learning rate* (F-contrast; 57 voxels right; 127 voxels left; see Figure 4A-C; Table 2). By testing for the direction of this brain-behavioral association we demonstrated, that the volume of the caudate nucleus was associated with higher *learning rate* in older compared to younger adults (not shown). In fact, the volumes of the left ($r = 0.55$, $p = 0.002$; Figure 4D) and right (r

= 0.87, $p < 0.001$; Figure 4E) caudate nucleus in older adults positively correlated with *learning rate* whereas only a weak negative correlation between the right caudate nucleus volume ($r = -0.42$, $p = 0.04$) and *learning rate* was observed in younger adults.

We further performed a regression analysis in older adults to explore how the modeling parameters were associated with striatal volume within the older group, which revealed a significant positive correlation between *learning rate* and bilateral caudate nucleus (448 voxels caudate right; 184 voxels caudate left, see Figure 4F-H, Table 2). However, no such positive correlation was observed within the younger group, which displayed a negative correlation between the right caudate nucleus volume and *learning rate*. No significant main effect of *learning rate* with caudate volume was observed across the whole group.

As previously described, we performed an additional exploratory voxel-based analysis using a more comprehensive mask including the pallidum, insula, prefrontal cortex, cingulate cortex, the surrounding medial cortex, SN/VTA and hippocampus. The regression analysis further revealed a significant positive correlation between the left pallidum and *learning rate* in older adults (Supplementary Table 9).

Setting the *Pavlovian bias*, *action bias*, *lapse rate*, *reward* and *punishment sensitivity* as regressors of interest revealed no significant differences in the striatum both within and between groups. However, a significant cluster in the bilateral insula was observed in younger adults setting *lapse rate* as a parameter of interest (see Supplementary Table 10). Finally, no significant effects in WM were observed.

Table 2

MNI coordinates								
No. of voxels	x (mm)	y (mm)	z (mm)	Structure	F	Z	$P_{(\text{cluster})}$	$P_{(\text{peak})}$
127	-15	-9	21	Left caudate	10.74	3.55	0.008	< 0.001
58	15	21	12	Right caudate	11.09	3.61	0.057	< 0.001
No. of voxels	x (mm)	y (mm)	z (mm)	Structure	t	Z	$P_{(\text{cluster})}$	$P_{(\text{peak})}$
448	-15	-9	21	Left caudate	4.62	4.10	< 0.001*	< 0.001
184	15	21	12	Right caudate	4.61	4.09	< 0.009	< 0.001

*Voxel based morphometry results for learning rate in the striatum in MNI space. Displayed are all clusters > 30 voxels. F statistics refer to structural differences between young and older adults and T statistics refer to regression analyses in older adults with respect to learning rate. * refers to clusters that survived family wise error correction ($p_{(FWE)} < 0.05$). All displayed results were significant at peak-level ($p_{(peak)} < 0.001$). Key: MNI, Montreal Neurological Institute.*

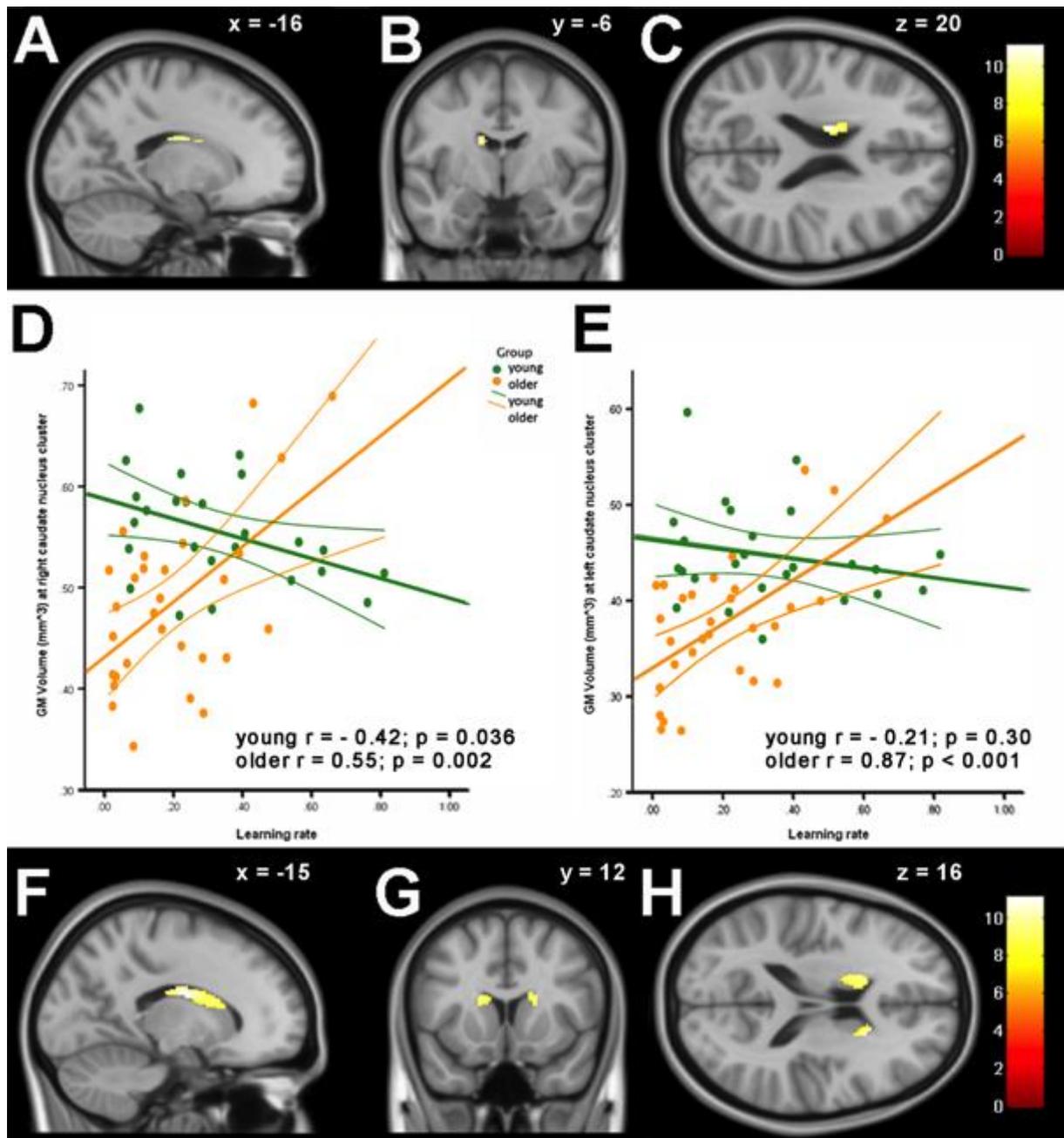


Figure 4. Voxel-based morphometry (VBM) analysis in young and older adults revealed an age group by learning rate interaction with caudate volume (sagittal (A), coronal (B) and axial (C) view), suggesting a difference in the association between volume and learning rate between both age groups. More specifically, the volume of the caudate nucleus was associated with higher learning rate in older compared to younger adults (not shown). The volumes of left (D) and right (E) caudate nucleus in older adults (in orange) positively

correlated with learning rate whereas a weak negative correlation between the right caudate nucleus volume and learning rate was observed in younger adults (in green). Thin lines represent 95% confidence intervals. A regression analysis in older adults demonstrated that individual variability in learning rate positively correlated with bilateral caudate volume. Results are presented in sagittal (F), coronal (G) and axial (H) views using MNI coordinates at $p < 0.001$, uncorrected threshold.

Dopamine synthesis capacity

We assessed the relationship between task performance, computational modeling parameters and synthesis capacity of dopamine in the dorsal (caudate nucleus and putamen) and ventral striatum (nucleus accumbens) of older adults. In contrast to the voxel based morphometry (VBM) results, we found no significant relationship between task performance or *learning rate* with synthesis capacity of dopamine (K_i) in the caudate nucleus. A correlation between synthesis capacity in right caudate nucleus and *lapse rate* was detected ($r = 0.49$, $p = 0.04$), but did not survive Bonferroni correction for multiple comparisons. Furthermore, mean cluster values of the bilateral caudate of the brain-behavioral analysis with learning rate did not correlate with dopamine synthesis capacity in this region and thus did not moderate the relationship between caudate volume and learning rate.

A positive correlation was also observed between percentage of correct answers in GW (actions pertaining to reward) and dopamine synthesis capacity in right ($r = 0.54$, $p = 0.02$) and left putamen ($r = 0.47$, $p = 0.04$), as well as between *Pavlovian bias* and dopamine synthesis capacity in the right nucleus accumbens ($r = 0.47$, $p = 0.04$). However these correlations did not survive Bonferroni correction for multiple comparisons (see supplementary Table 11).

Discussion

The aim of the study was to investigate how structural and neuromodulatory differences in the striatum modulate the impact of valence on learning action invigoration and inhibition (Guitart-Masip et al. 2011). Older adults demonstrated poorer performance in the probabilistic go/no-go task compared to younger adults. Assessment of striatal GM volume, revealed a significant group interaction between *learning rate* and volume of the caudate nucleus. Most importantly, differences in *learning rate* within the group of older adults, correlated with changes in volume but not dopamine synthesis capacity of the bilateral caudate nucleus. These findings suggest that structural age-related changes to the caudate may underlie learning deficits in aging.

Consistent with previous studies, instrumental learning of action-valence associations (Guitart-masip et al. 2012a; Chowdhury 2013) was better for non-conflicting (GW and NGAL) versus conflicting conditions (NGW and GAL) in both young and older adults. Computational modeling using a task-specific reinforcement learning-model with a constant *Pavlovian bias* revealed older adults demonstrated a significant age-related decrease in *reward* and *punishment sensitivity* and a marked (albeit non-significant) reduction in the *learning* and *lapse rate* compared to younger adults. Although the behavioral analysis did show that the probability of choosing an action was differentially contingent on valence between groups, no difference in the *Pavlovian bias* was observed between young and older adults. This suggests that the age-related differences in the impact of valence on learning are better captured by the attenuation of reward and punishment sensitivity in older adults.

The primary motivation of the study was to investigate how the structural integrity of the striatum in ageing may influence instrumental learning and Pavlovian biases in younger and

older adults. In particular, we explored the neural substrates underlying age-related differences in learning using computational modeling and VBM. In a comparison of young and older adults, we found age-related differences in bilateral caudate nucleus volume related to *learning rate*. Furthermore, we found a positive correlation between caudate nucleus volume and *learning rate* in older adults, whilst a weak and unilateral negative correlation between caudate volume and *learning rate* was observed in younger adults.

Collectively these results demonstrate that reduced learning in older adults strongly relates to the structural integrity of the caudate nucleus. According to reinforcement learning theories, the caudate nucleus, as part of the dorsal striatum (DS), has previously been linked to instrumental conditioning and action value representation (Samejima 2005; Seo et al. 2012). This contrasts with the ventral striatum VS, which has classically been linked to Pavlovian conditioning and expected value representation (O'Doherty et al. 2004; Schmidt et al. 2012). More specifically, it has been suggested that the caudate nucleus may be implicated in goal-directed behavior and thus may directly mediate instrumental learning performance (Liljeholm and O'Doherty 2012). Moreover, a previously conducted functional MRI study in young adults using a variation of this task that does not require learning (Guitart-Masip et al. 2011), demonstrated an association between the anticipation of action value and activity in the DS and suggests the DS may be crucial for evaluating the weight of an action. Thus it is conceivable that degeneration to the DS, i.e. as a result of normal aging, could impair instrumental learning performance. Furthermore previous studies have shown age-related reductions in RPE representation, and thus learning, and not reward value representation, may be responsible for poorer performance in older adults using other reward-based probabilistic tasks (Chowdhury et al., 2013; Samanez-Larkin et al. 2014). In summary, our findings are consistent with the role of the DS in action value learning and indicate structural age-related differences in this brain region impair action learning in older age. On the contrary, no clear relation between DS volume and learning was observed within the group of young adults.

The use of high-field MRI represents a novel aspect of our study and strengthens the VBM results since the increased signal to noise ratio at 7T permits superior differentiation of GM from WM in both cortical and subcortical areas (Duyn 2012; Plantinga et al. 2014). Furthermore, direct comparisons of T₁-weighted images acquired at 7T compared to lower field strength scans, have confirmed better edge detection power in the basal ganglia region and more precise GM segmentation in subcortical regions such as the striatum using ultra high-field MRI (Cho et al. 2010). Through our multimodal neuroimaging approach, we were able to assess structural differences and dopamine synthesis capacity in the striatum in the same older adults. In doing so we revealed that the positive correlation between the structural integrity of the caudate nucleus and *learning rate* was not supported by a comparable correlation with dopamine synthesis capacity. This discrepancy may suggest that differences in the rate of learning from rewards and negative outcomes may not be dependent on dopamine. However the lack of association between learning rate and dopamine synthesis capacity may have also been influenced by the limited sample size.

Considering a large body of literature stressing a role for dopamine in reinforcement learning, the findings presented might seem controversial. According to an influential RPE model, changes in phasic dopamine signals reflect RPE (Schultz 1998, 2002) that are reported to the striatum, where positive RPE reinforce rewarded actions and negative RPE extinguishes unrewarded actions (Frank et al. 2004). Moreover, there is ample genetic and pharmacological evidence supporting an involvement of dopamine in learning (Fossella et al. 2002). Furthermore, neurocognitive models of Parkinsonism imply that learning in PD patients is impaired due to reduced dynamic dopaminergic modulation required for positive and negative RPE (Frank 2005) and numerous studies suggest that dopamine mediated basal

ganglia pathways are required for reward and punishment-based learning (Kravitz et al. 2012; van der Schaaf et al. 2013). However, at least one study has shown how manipulation of dopamine levels in humans using L-DOPA or dopamine antagonists only affects learning through rewards (Pessiglione et al. 2006). Most importantly, computational modeling approaches using alternative instrumental learning tasks that provide a more subtle approach to differentiate between learning and asymptotic performance have shown that dopamine agonists or antagonists do not influence the *learning rate* latent variable (Eisenegger et al. 2014; Lee et al. 2015), despite modulating functional activation in the striatum. These findings suggest that dopamine may not always be the main influencer of *learning rate* in reinforcement learning tasks and may rather relate to exploitation of rewards (Shiner et al. 2012; Smittenaar et al. 2012; Averbeck and Costa 2017). Nonetheless, interindividual differences in baseline dopamine may also impact these pharmacological effects evident by a previous study demonstrating Bromocriptine can enhance reward-based reversal learning in young adults with relatively low striatal dopamine synthesis capacity, yet can paradoxically impair it in those with high dopamine synthesis capacity (Cools et al. 2009).

In line with the controversial role of dopamine, a recent study using a similar task in Parkinson's disease patients, revealed dopaminergic medication had no overall effect on learning, although "on" medication state influenced learning patterns promoting NGW and inhibiting GAL (van Wouwe et al. 2017). Another study delivers further proof of the uncertain role of dopamine in this context, showing L-DOPA leads to improved performance in the NGW condition and decreases the strength of the *Pavlovian bias* (Guitart-Masip et al. 2013). In this study the authors speculated that the decreased coupling between action and valence with L-DOPA, could be a result of increased working memory and executive functions as a result of increased dopamine in the prefrontal cortex. Furthermore additional

neuromodulators should also be considered. For example serotonin has been attributed a role in behavioral inhibition (Dayan and Huys 2009; Guitart-masip et al. 2014) and learning from punishment (Crockett et al. 2012, den Ouden et al. 2013), while acetylcholine has also been shown to finely tune representations of reward (Suzuki and Amaral 2004)

Some important methodological limitations should be considered on interpretation of these findings. Firstly, VBM analysis at 7T has been challenged, due to greater field inhomogeneities and GM volume estimation differences when compared to 3T (Belaroussi et al. 2006) although GM volume definition in the basal ganglia has proved to be reliable (Seiger et al. 2015). In order to reduce field inhomogeneities we applied a bias correction step and carefully inspected all GM segmentations. Second, the SN/VTA demonstrates poor contrast in T_1 -weighted images and is not reliably segmented in GM. Additional MRI sequences demonstrating superior contrast in these regions e.g. using T_2^* -weighted images or Quantitative Susceptibility Maps at 7 T (Deistung et al. 2013; Betts et al. 2016) would be desired to further identify the role of the midbrain in instrumental learning. Third, whilst VBM is thought to identify volumetric variations, the nature of GM differences identified by this method are still poorly understood and could also depend on neuronal size or density, dendritic arborization or even changes in the neurophil (Mechelli et al. 2005). Furthermore, we acknowledge our limited sample size for both VBM and PET analyses may have led to an overestimation of significance due to noisy measures (Loken and Gelman 2017), but conversely may have also limited our ability to detect an association between dopamine synthesis capacity and the modeling parameters. Finally, we acknowledge that it is difficult to interpret differences in absolute Pavlovian bias values across subjects since these values are contingent on each individual's reward/punishment sensitivity.

In the future it would be desirable to extend these findings using alternative dopaminergic tracers i.e. to assess how age-related change in postsynaptic receptor density may influence Pavlovian learning in ageing. Indeed negative correlations with age have been previously reported for D₁ (Wang et al. 2008; Bäckman et al. 2009) and D₂-like receptor density in the striatum (Antonini et al. 1993), as well as for presynaptic dopamine transporter (DAT) (van Dyck et al. 1995; Erixon-Lindroth et al. 2005). More specifically, recent work identified an association between the variability of striatal D1 receptor density, in particular in the DS, and behavioral measures using a variant of the same go/no-go task (de Boer et al. 2019). However these effects did not appear to be influenced by age. In future, pharmacological intervention studies targeting additional neuromodulatory systems would also be desirable in order to determine the role of acetylcholine and serotonin in Pavlovian learning.

Using a combination of structural MRI, f-DOPA PET and computational modeling, we identified a dissociation between structural and neuromodulatory influences on Pavlovian learning using a go/no-go task that orthogonalizes action and valence. Our study revealed that age-related differences in the caudate nucleus were associated with *learning rate*, demonstrating the structural integrity of the dorsal striatum may be an important neural substrate underlying learning deficits in old age. No significant relationship between striatal dopamine synthesis capacity and *learning rate* was observed suggesting that learning in older age might depend more on structural compared to neuromodulatory age-related differences in the striatum.

Acknowledgements

We are very grateful to all participants who volunteered to take part in the study. We thank Iris Mann, Anne Hochkepler and Melanie Eggestein for help in subject recruitment and data collection. We also wish to thank our local radiographers Kerstin Möhring, Ilona Wiedenhöft,

Denise Scheermann and Renate Blobel-Lüer for their dedicated efforts in scanning standardisation and quality assurance.

Funding

This project received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 720270 (HBP SGA1) and was supported by the SFB 779-TP A07 awarded to E.D. M G-M was supported by a research grant from the Swedish Research Council (VR521-2013-2589).

Notes

No conflicts of interest to declare.

References

- Antonini A, Leenders KL, Reist H, Thomann R, Beer HF, Locher J. 1993. Effect of age on D2 dopamine receptors in normal human brain measured by Positron Emission Tomography and 11C-raclopride. *Ann Neurol.* 50:474–480.
- Ashburner J. 2007. A fast diffeomorphic image registration algorithm. *Neuroimage.* 38:95–113.
- Averbeck BB, Costa VD. 2017. Motivational neural circuits underlying reinforcement learning. *Nat Neurosci.* 20:505–512.
- Bäckman L, Karlsson S, Fischer H, Karlsson P, Brehmer Y, Rieckmann A, MacDonald SWS, Farde L, Nyberg L. 2009. Dopamine D1 receptors and age differences in brain activation during working memory. *Neurobiol Aging.* 32:1849–1856.
- Beck a T, Steer R a, Ball R, Ranieri W. 1996. Comparison of Beck Depression Inventories - IA and -II in psychiatric outpatients. *J Pers Assess.*
- Belaroussi B, Milles J, Carme S, Zhu YM, Benoit-Cattin H. 2006. Intensity non-uniformity correction in MRI: Existing methods and their validation. *Med Image Anal.*
- Betts MJ, Acosta-cabronero J, Cardenas-blanco A, Nestor PJ, Düzel E. 2016. NeuroImage High-resolution characterisation of the aging brain using simultaneous quantitative susceptibility mapping (QSM) and R^2 * measurements at 7 T. *Neuroimage.* 138:43–63.
- Cavanagh JF, Eisenberg I, Guitart-masip M, Huys Q, Frank MJ. 2013. Frontal Theta Overrides Pavlovian Learning Biases. *J Neurosci.* 33:8541–8548.
- Cho Z-H, Min H-K, Oh S-H, Han J-Y, Park C-W, Chi J-G, Kim Y-B, Paek SH, Lozano AM,

- Lee KH. 2010. Direct visualization of deep brain stimulation targets in Parkinson disease with the use of 7-tesla magnetic resonance imaging. *J Neurosurg.* 113:639–647.
- Chowdhury R et A. 2013. Dopamine restores reward prediction errors in old age. *Nat Neurosci.* 16:648–653.
- Chowdhury R, Guitart-masip M, Lambert C, Dolan RJ, Düzel E. 2013. Structural integrity of the substantia nigra and subthalamic nucleus predicts flexibility of instrumental learning in older-age individuals. *Neurobiol Aging.*
- Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, Esposito MD. 2009. Striatal Dopamine Predicts Outcome-Specific Reversal Learning and Its Sensitivity to Dopaminergic Drug Administration. 29:1538–1543.
- Corbit LH, Balleine BW. 2011. The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J Neurosci.* 31:11786–11794.
- Dayan P, Huys QJM. 2009. Serotonin in Affective Control. *Annu Rev Neurosci.* 32:95–126.
- Dayan P, Niv Y. 2008. Reinforcement learning: The Good, The Bad and The Ugly. *Curr Opin Neurobiol.* 18:185–196.
- Dayan P, Niv Y, Seymour B, Daw N. 2006. The misbehavior of value and the discipline of the will. *Neural Networks.* 19:1153–1160.
- de Boer L, Axelsson J, Chowdhury R, Riklund K, Dolan RJ, Nyberg L, Bäckman L, Guitart-Masip M. 2019. Dorsal striatal dopamine D1 receptor availability predicts an instrumental bias in action learning. *Proc Natl Acad Sci U S A.* 116:261–270.
- Deistung A, Schäfer A, Schweser F, Biedermann U, Turner R, Reichenbach JR. 2013. Toward in vivo histology: A comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R^{2*}-imaging at ultra-high magnetic field strength.

- Neuroimage. 65:299–314.
- den Ouden HEM, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, Franke B, Cools R. 2013. Dissociable Effects of Dopamine and Serotonin on Reversal Learning. *Neuron*. 80:1090–1100.
- Di X, Rypma B, Biswal BB. 2014. Correspondence of Executive Function Related Functional and Anatomical Alterations in Aging Brain. *Prog Neuropsychopharmacol Biol Psychiatry*. 6:2166–2171.
- Dreher J-C, Meyer-Lindenberg A, Kohn P, Berman KF. 2008. Age-related changes in midbrain dopaminergic regulation of the human reward system. *Proc Natl Acad Sci U S A*. 105:15106–15111.
- Duyn JH. 2012. The future of ultra-high field MRI and fMRI for study of the human brain. *Neuroimage*. 62:1241–1248.
- Eisenegger C, Naef M, Linssen A, Clark L, Gandamaneni PK, Mu U. 2014. Role of Dopamine D2 Receptors in Human Reinforcement Learning. 2366–2375.
- Eppinger B, Schuck NW, Nystrom LE, Cohen JD. 2013. Reduced striatal responses to reward prediction errors in older compared with younger adults. *J Neurosci*. 33:9905–9912.
- Erixon-Lindroth N, Farde L, Wahlin T-BR, Sovago J, Halldin C, Bäckman L. 2005. The role of the striatal dopamine transporter in cognitive aging. *Psychiatry Res*. 138:1–12.
- Fera F, Weickert TW, Goldberg TE, Tessitore A, Hariri A, Das S, Lee S, Zolnick B, Meeter M, Myers CE, Gluck M a, Weinberger DR, Mattay VS. 2005. Neural mechanisms underlying probabilistic category learning in normal aging. *J Neurosci*. 25:11340–11348.
- Folstein MF, Folstein SE, McHugh PR. 1975. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 12:189–198.
- Fossella J, Sommer T, Fan J, Posner MI, Swanson JM, Pfaff DW, Wu Y, Fossella J. 2002.

- Assessing the molecular genetics of attention networks. *BMC Neurosci.* 3:14.
- Frank MJ. 2005. Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism. *J Cogn Neurosci.* 17:51–72.
- Frank MJ, O'Reilly RC, Seeberger LC. 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science.* 306:1940–1943.
- Glaser MO, Glaser WR. 1982. Time course analysis of the Stroop phenomenon. *J Exp Psychol Hum Percept Perform.*
- Grady C. 2013. Trends in Neurocognitive Aging. *Nat Rev Neurosci.* 123:106–116.
- Guitart-Masip M, Chowdhury R, Sharot T, Dayan P, Duzel E, Dolan R. 2012. Action controls dopaminergic enhancement of reward representations. *Proc Natl Acad Sci.* 109:7511–7516.
- Guitart-Masip M, Duzel E, Dolan R, Dayan P. 2014. Action versus valence in decision making. *Trends Cogn Sci.* 18:194–202.
- Guitart-masip M, Economides M, Huys QJM, Dayan P, Dolan RJ. 2014. Differential, but not opponent, effects of L -DOPA and citalopram on action learning with reward and punishment. *Psychopharmacology (Berl).* 231:955–966.
- Guitart-Masip M, Economides M, Huys QJM, Frank MJ, Chowdhury R, Duzel E, Dayan P, Dolan RJ. 2013. Differential, but not opponent, effects of l-DOPA and citalopram on action learning with reward and punishment. *Psychopharmacology (Berl).* 231:955–966.
- Guitart-Masip M, Fuentemilla L, Bach DR, Huys QJM, Dayan P, Dolan RJ, Duzel E. 2011. Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J Neurosci.* 31:7867–7875.
- Guitart-masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ. 2012. Go and no-

- go learning in reward and punishment: Interactions between affect and effect. *Neuroimage*. 62:154–166.
- Haber SN, Knutson B. 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*. 35:4–26.
- Huys QJM, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, Dayan P. 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput Biol*. 7:e1002028.
- Kass R. E. and Raftery A. E. 1995. Bayes Factors. *J Am Stat Assoc*. Vol. 90:773–795.
- Kravitz A V., Tye LD, Kreitzer AC. 2012. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat Neurosci*. 15:816–818.
- Leclercq M, Zimmermann P. 2004. *Applied Neuropsychology of attention: Theory, Diagnosis and Rehabilitation*. Psychology Press, 2004.
- Lee E, Seo M, Dal Monte O, Averbeck BB. 2015. Injection of a Dopamine Type 2 Receptor Antagonist into the Dorsal Striatum Disrupts Choices Driven by Previous Outcomes, But Not Perceptual Inference. *J Neurosci*. 35:6298–6306.
- Liljeholm M, O’Doherty JP. 2012. contributions of the striatum to learning, motivation, and performance: An associative account. *Trends Cogn Sci*.
- Loken E, Gelman A. 2017. Measurement error and the replication crisis. *Science* (80-). 355:584–585.
- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*. 19:1233–1239.
- Mechelli A, Price CJ, Friston KJ, Ashburner J. 2005. *Voxel-Based Morphometry Applications of the Human Brain : Methods and*. 1–9.

- Mell T, Heekeren HR, Marschner A, Wartenburger I, Villringer A, Reischies FM. 2005. Effect of aging on stimulus-reward association learning. *Neuropsychologia*. 43:554–563.
- Mell T, Wartenburger I, Marschner A, Villringer A, Reischies FM, Heekeren HR. 2009. Altered function of ventral striatum during reward-based decision making in old age. *Front Hum Neurosci*. 3:34.
- Mohr PNC, Li SC, Heekeren HR. 2010. Neuroeconomics and aging: Neuromodulation of economic decision making in old age. *Neurosci Biobehav Rev*. 34:678–688.
- Niv Y, Daw ND, Joel D, Dayan P. 2007. Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. 191:507–520.
- O’Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ, Instrumental. 2004. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental conditioning. *Science (80-)*. 452:1604–1563.
- O’Doherty ML and JP. 2012. Contributions of the striatum to learning, motivation, and performance: an associative account. *Trends Cogn Sci*. 16:467–475.
- Patenaude B, Smith SM, Kennedy DN, Jenkinson M. 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage*. 56:907–922.
- Patlak CS, Blasberg RG. 1985. Graphical Evaluation of Blood-to-Brain Transfer Constants from Multiple-Time Uptake Data. Generalizations. *J Cereb Blood Flow Metab*. 5:584–590.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 442:1042–1045.
- Plantinga BR, Temel Y, Roebroek A, Uludağ K, Ivanov D, Kuijf ML, ter Haar Romenij BM. 2014. Ultra-high field magnetic resonance imaging of the basal ganglia and related structures. *Front Hum Neurosci*. 8:1–22.

- Reynolds SM, Berridge KC. 2008. Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nat Neurosci.* 11:423–425.
- Salamone JD, Correa M. 2012. The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron.* 76:470–485.
- Samanez-Larkin GR, Kuhnen CM, Yoo DJ, Knutson B. 2010. Variability in Nucleus Accumbens Activity Mediates Age-Related Suboptimal Financial Risk Taking. *J Neurosci.* 30:1426–1434.
- Samanez-Larkin GR, Worthy D a, Mata R, McClure SM, Knutson B. 2014. Adult age differences in frontostriatal representation of prediction error but not reward outcome. *Cogn Affect Behav Neurosci.* 14:672–682.
- Samejima K. 2005. Representation of Action-Specific Reward Values in the Striatum. *Science (80-).* 310:1337–1340.
- Schmidt L, Lebreton M, Cléry-Melin ML, Daunizeau J, Pessiglione M. 2012. Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biol.* 10.
- Schott H, Niehaus L, Wittmann BC, Schu H, Seidenbecher CI, Heinze H, Du E. 2007. Ageing and early-stage Parkinson ' s disease affect separable neural mechanisms of mesolimbic reward processing. *Brain.* 2412–2424.
- Schultz W. 1998. Predictive Reward Signal of Dopamine Neurons. *J Neurophysiol.* 80:1–27.
- Schultz W. 2002. Getting Formal with Dopamine and Reward. *Neuron.* 36:241–263.
- Seiger R, Hahn A, Hummer A, Kranz GS, Ganger S, K??blb??ck M, Kraus C, Sladky R, Kasper S, Windischberger C, Lanzenberger R. 2015. Voxel-based morphometry at ultra-high fields. A comparison of 7T and 3T MRI data. *Neuroimage.* 113:207–216.
- Seo M, Lee E, Averbeck BB. 2012. Action selection and action value in frontal-striatal circuits. *Neuron.* 74:947–960.

- Shiner T, Seymour B, Wunderlich K, Hill C, Bhatia KP, Dayan P, Dolan RJ. 2012. Dopamine and performance in a reinforcement learning task : evidence from Parkinson ' s disease. *Brain*. 1–13.
- Smittenaar P, Chase HW, Aarts E, Nusslein B, Bloem BR, Cools R. 2012. Decomposing effects of dopaminergic medication in Parkinson ' s disease on probabilistic action selection – learning or performance ? *Eur J Neurosci*. 35:1144–1151.
- Sutton RS, Barto AG. 1998. Sutton & Barto Book: Reinforcement Learning: An Introduction [WWW Document]. MIT Press Cambridge, MA, A Bradford B. URL <http://webdocs.cs.ualberta.ca/~sutton/book/the-book.html>
- Suzuki WA, Amaral DG. 2004. Functional neuroanatomy of the medial temporal lobe memory system. *Cortex*.
- Swart JC, Froböse MI, Cook JL, Geurts DE, Frank MJ, Cools R, den Ouden HE. 2017. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife*. 6:1–36.
- Talmi D, Dayan P, Kiebel SJ, Frith CD, Dolan RJ. 2009. How Humans Integrate the Prospects of Pain and Reward during Choice. *J Neurosci*. 29:14617–14626.
- Ueda Y, Yamanaka K, Noritake A, Enomoto K, Matsumoto N, Yamada H, Samejima K, Inokawa H, Hori Y, Nakamura K, Kimura M. 2017. Distinct Functions of the Primate Putamen Direct and Indirect Pathways in Adaptive Outcome-Based Action Selection. *Front Neuroanat*. 11:1–8.
- van der Schaaf ME, Zwiers MP, van Schouwenburg MR, Geurts DEM, Schellekens AFA, Buitelaar JK, Verkes RJ, Cools R. 2013. Dopaminergic drug effects during reversal learning depend on anatomical connections between the orbitofrontal cortex and the amygdala. *Front Neurosci*. 7:1–10.

- van Dyck CH, Seibyl JP, Malison RT, Laruelle M, Wallace E, Zoghbi SS, Zea-Ponce Y, Baldwin RM, Charney DS, Hoffer PB. 1995. Age-related decline in striatal dopamine transporter binding with iodine-123- β -CITSPECT. *J Nucl Med.* 36:1175–1181.
- Van Wouwe NC, Claassen DO, Neimat JS, Kanoff KE, Wylie SA. 2017. Dopamine Selectively Modulates the Outcome of Learning Unnatural Action–Valence Associations. *J Cogn Neurosci.* 29:816–826.
- Vink M, Kleerekooper I, van den Wildenberg WPM, Kahn RS. 2015. Impact of aging on frontostriatal reward processing. *Hum Brain Mapp.* 36:2305–2317.
- Wang Z, Aguirre GK, Rao H, Wang J, Fernández- MA, Childress AR, Detre JA. 2008. Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. *Magn Reson Imaging.* 26:261–269.
- Williams A. 1987. The other psychology of animal learning: a review of mackintosh's conditioning and associative learning. *J Exp Anal Behav.* 1:175–186.