Dopamine restores reward prediction errors in old age

Rumana Chowdhury[1,2,*], **Marc Guitart-Masip**[2,3,*], **Christian Lambert**[2,4], **Peter Dayan**[5], **Quentin Huys**[2,5,6], **Emrah Düzel**[1,7,8], and **Raymond J Dolan**[2]

[1]Institute of Cognitive Neuroscience, University College London, UK, WC1N 3AR

[2]Wellcome Trust Centre for Neuroimaging, University College London, UK, WC1N 3BG

[3]Aging Research Center, Karolinska Institute, 113 30 Stockholm, Sweden

[4]St George's University of London, Cranmer Terrace, London, UK, SW17 0RE

[5]Gatsby Computational Neuroscience Unit, University College London, UK, WC1N 3AR

[6]Translational Neuroimaging Unit, Department of Biological Engineering, ETH and University of Zurich, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry Zurich

[7]Otto-von-Guericke-University Magdeburg, Institute of Cognitive Neurology and Dementia Research, Leipziger Str. 44, 39120 Magdeburg, Germany

[8]German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany

Abstract

Senescence affects the ability to utilize information about the likelihood of rewards for optimal decision-making. In a human functional magnetic resonance imaging (fMRI) study, we show that healthy older adults have an abnormal signature of expected value resulting in an incomplete reward prediction error signal in the nucleus accumbens, a brain region receiving rich input projections from substantia nigra/ventral tegmental area (SN/VTA) dopaminergic neurons. Structural connectivity between SN/VTA and striatum measured with diffusion tensor imaging (DTI) was tightly coupled to inter-individual differences in the expression of this expected reward value signal. The dopamine precursor levodopa (L-DOPA) increased the task-based learning rate and task performance in some older adults to a level shown by young adults. Critically this drug-effect was linked to restoration of a canonical neural reward prediction error. Thus we identify a neurochemical signature underlying abnormal reward processing in older adults and show this can be modulated by L-DOPA.

Introduction

Aging in humans is associated with a range of changes in cognition. For example, older adults are particularly poor at making decisions when faced with probabilistic rewards, possibly due to impaired learning of stimulus-outcome contingencies [1] [2]. Such findings raise

two fundamental questions, namely what are the substrates for learning in these circumstances and what accounts for this aberrant decision-making.

One function critical for decision-making is learning to predict rewards. There is ample evidence from animal experiments that the neuromodulator dopamine encodes the difference between actual and expected rewards (so-called 'reward prediction errors') [3, 4]. In humans there is now compelling evidence that functional activation patterns in the nucleus accumbens, a major target region of dopamine neurons [5], report rewarding outcomes and associated prediction errors [6-9]. A more direct link to dopamine is seen using pharmacological challenge with dopaminergic agents [10, 11].

In terms of what might go wrong during aging, one important clue is the well-described age-related loss of dopamine neurons within the substantia nigra/ventral tegmental area (SN/VTA) [12] [13], evident both in histology and when using diffusion tensor imaging (DTI) as an indirect marker of structural degeneration [14] [15]. However, the consequences for decision-making of this decline in dopamine are unclear because of functional interactions among the triplet of reward representations, representations of prediction errors associated with that reward, and the learning of predictions that underpins the expression of these prediction errors. In older age, abnormal activity in the nucleus accumbens has been associated with suboptimal decision-making and reduced reward anticipation, but with normal responses to rewarding outcomes [16] [17] [18]. This has led to the suggestion that although older adults may maintain adequate representations of reward, they are unable to learn correctly from these representations.

We studied the effect of probabilistic rewarding outcomes on the separate reward and prediction components of a prediction error signal [19] in healthy older adults. To that end, we employed a simple probabilistic instrumental conditioning problem - the two armed bandit choice task (Figure 1a). Older adults underwent DTI and fMRI in combination with a pharmacological manipulation using the dopamine precursor levodopa (L-DOPA), using a within-subject double-blind placebo-controlled study. We collected behavioural data in a group of young adults to contextualise the effects of age on performance. We did not administer L-DOPA to these young controls, implying that the effects of L-DOPA could not be compared across age-groups. By exploiting a reinforcement learning model we could determine which component of the prediction error (the actual and/or expected reward representation) was impaired in older age. DTI enabled us to examine nigro-striatal structural connectivity strength, based on a hypothesis that individual differences in this structural measure would predict inter-individual differences in baseline functional reward prediction error signalling. A crucial observation here is the fact that L-DOPA administration has been associated with greater prediction errors in young adults [10] and higher learning rates in patients with Parkinson's disease [11]. We therefore predicted that L-DOPA would increase the learning rate evident in behaviour as well as boost the representation of a reward prediction error in the nucleus accumbens of healthy older adults, specifically by increasing the component associated with the expected value.

Results

Behavioural performance in young and older adults

32 older adults (mean age 70.00 years, SD 3.24; Supplementary Table 1) on placebo and L-DOPA and 22 young adults (mean age 25.18 years, SD 3.85) performed a two armed bandit choice task (Figure 1a). Older adults completed a similar number of trials under both conditions (placebo: mean 218.16, SD 1.94; L-DOPA: mean 218.47, SD 1.74) as young adults (mean 218.50, SD 2.44) (all $p > 0.4$). Older adults had similar choice reaction times on placebo (mean 796.81 ms, SD 152.89) and L-DOPA (mean 781.49 ms, SD 140.17) (paired

t-test, t (31) = 1.01, p = .321), but overall were slower under both conditions compared to young adults (mean 629.69 ms, SD 156.41) (independent t-tests, young vs. old-placebo: t (52) = 3.91; young vs. old-L-DOPA: t (52) = 3.73; both p <0.0005).

Overall the amount of money won by older adults performing the task did not differ under L-DOPA (mean £12.94, SD 0.81) compared to placebo (mean £12.64, SD 0.89) (paired t test: t(31) = 1.53, p=.137). However, older adults on placebo won significantly less money than young adults (mean £13.17, SD 1.00) (independent samples t test: t(52) = 2.05, p = .045) whereas there was no difference in the amount won between older adults on L-DOPA and young adults (t(52) = 0.971, p = .336) (Figure 1b).

A more detailed examination of the behavioural data showed that only a proportion of older adults won more money on the task under L-DOPA compared to placebo. To examine this further, we performed a median split according to drug-induced changes in performance (see Methods for details), creating a group who 'win less on L-DOPA' (total won L-DOPA < placebo, n = 17) and a group who 'win more on L-DOPA' (total won L-DOPA > placebo, n = 15). This analysis revealed that performance in older adults was consistent with an 'inverted U-shape' whereby those with high baseline levels of performance on placebo performed less well on L-DOPA while conversely those with low baseline levels of performance improved on L-DOPA (Supplementary Fig 1). Performance in the 'win less on L-DOPA' group on placebo and in the 'win more on L-DOPA' group on L-DOPA was at a similar level to performance in young adults (young adults vs. 'win less on L-DOPA' group on placebo, t(37) = 0.19, p = .854; young adults vs. 'win more on L-DOPA' group on L-DOPA, t(35) = −0.40, p = .690) whereas performance in the 'win less on L-DOPA' group on L-DOPA and in the 'win more on L-DOPA' group on placebo was worse than performance in young adults (young adults vs. 'win less on L-DOPA' group on L-DOPA t(37) = 2.07, p = .045; young adults vs. 'win more on L-DOPA' on placebo, t(35) = 3.53, p = .001). This inverted U-shape pattern of performance is in line with previous reports of the effects of dopamine on cognition [20] and suggests that variable performance across older adults is linked to individual differences in baseline dopamine status.

Reinforcement learning behaviour

We analysed trial-by-trial choice behaviour using a standard reinforcement learning model with a fixed β parameter (Figure 2a). Note that by using this methodological approach, the learning rate reflects a summary measure of reinforcement learning strength (see Methods). A model with a single fixed β = 1.27 across drug and placebo conditions, one single learning rate and one choice perseveration parameter provided the best model fit of older participants' choices among the range of models that we compared, indexed by the lowest BIC values (Supplementary Table 2). When calculating the BIC, the log evidence was penalized using the number of data points associated with each parameter.

In order to further examine the effects of L-DOPA on older participants' behaviour in the task, we used the Wilcoxon Signed Ranks Tests to determine whether the learning rates (fitted using a single prior distribution including the drug and the placebo) differed between L-DOPA and placebo. We found that participants had a significantly higher learning rate under L-DOPA compared to placebo (Z = −3.03, p = .002; Figure 2b). Crucially, this effect was significant in the group of older adults who performed better under L-DOPA ('win more on L-DOPA' group of older adults, placebo vs. L-DOPA: Z = −2.90, p = .004) but not amongst older adults who performed worse on L-DOPA ('win less on L-DOPA' group of older adults, placebo vs. L-DOPA: Z = −0.97, p = .332), providing a direct link between the effects of L-DOPA and task performance (Figure 2c). In contrast, choice perseveration was unaffected by L-DOPA (Z = −0.58, p = .562). In young adults, a model with a fixed β = 1.13 and single learning rate provided a better fit to participants' choices than when a choice

perseveration parameter was added to the model (BIC 4348.15 and 4361.01 respectively). The learning rate in young adults (median $\alpha = 0.62$, range $0.01 – 0.94$) was intermediate between, and not significantly different from older adults either under the condition of placebo ($Z = -1.32$, $p = .187$) or L-DOPA ($Z = -1.25$, $p = .211$) (Figure 2b).

L-DOPA and striatal prediction errors in older adults

We focussed our imaging analysis on within-subject comparisons of reward predictions errors in the nucleus accumbens (n = 32 older adults). Using a functional ROI approach (Supplementary Fig 2) we first defined voxels in the nucleus accumbens that signalled a 'putative' prediction error, namely voxels where there was an enhanced response at the time of outcome to actual rewards that was greater than that to expected rewards ($R(t) > Q_{a(t)}(t)$, see methods). Using this approach we identified a cluster in the right nucleus accumbens [peak voxel MNI co-ordinates x,y,z = 15, 11, −8; peak Z = 4.45, p<0.001 uncorrected; 34 voxels] (Figure 3a). Note this is a liberal definition of reward prediction errors, as voxels showing a significant effect with this contrast may not satisfy all the criteria to be considered for a 'canonical' reward prediction error, namely both a positive effect of reward and a negative effect of expected value[21][19]. We adopted this approach to test the hypothesis that canonical reward prediction errors are not fully represented in older age and critically, to test for the orthogonal effects of L-DOPA on the separate reward and expected value components of the prediction error signal.

We used this anatomically-constrained functional ROI to extract the parameters estimates for $R(t)$ and $Q_{a(t)}(t)$ separately within these activated voxels. Our two (placebo/L-DOPA) by two ($R(t)/Q_{a(t)}(t)$) repeated measures ANOVA revealed a main effect of L-DOPA ($F(1,31) = 5.712$, $p = .023$), suggesting administration of L-DOPA had an impact on the representations associated with the components of the reward prediction error (Figure 3a). Importantly, BOLD responses were only compatible with a canonical prediction error signal (positive correlation between BOLD and $R(t)$ along with a negative correlation between BOLD and $Q_{a(t)}(t)$) when participants were under L-DOPA (one-tailed one-sample t-test: $R(t)$ L-DOPA $t = 1.92$, $p = .033$; $Q_{a(t)}(t)$ L-DOPA: $t = -1.73$, $p = .047$ $R(t)$ placebo $t = 3.72$, p<0.001; $Q_{a(t)}(t)$ placebo $t = -0.11$, p= .455). This was due to a more negative representation of expected value $Q_{a(t)}(t)$ on L-DOPA compared to placebo (paired t-test, $t(31) = 2.37$, $p = .024$) whereas there was no difference in actual reward representation $R(t)$ between L-DOPA and placebo ($t(31) = 1.38$, $p = .179$). These results highlight that canonical reward prediction errors are not fully represented in older adults at baseline, whereby under placebo the nucleus accumbens responds only to reward and not to expected value. Only after receipt of L-DOPA was a canonical reward prediction error signal observed.

Under placebo, individual differences in the total amount won on the task correlated positively with the learning rate (Spearman's Rho = 0.39, $p = .027$), and task performance correlated negatively with the BOLD representation of expected value ($Q_{a(t)}(t)$) (Pearson's r = −0.42, $p = .016$), though this was not the case with reward ($R(t)$) (Pearson's r = −0.07, $p = .707$). Thus, better baseline performance was associated with a higher learning rate and more negative expected value representations in the nucleus accumbens. Across all 32 older participants, task performance on L-DOPA did not correlate with the learning rate or BOLD representations of reward or expected value (all p>0.15; Supplementary Table 3).

Critically however, subsequent analysis based on a median split for the effects of drug on performance showed that expected value ($Q_{a(t)}(t)$) parameter estimates in older adults who performed better on L-DOPA were significantly more negative on L-DOPA compared to placebo ('win more on L-DOPA' group, $Q_{a(t)}(t)$) placebo vs. L-DOPA $t(14) = 2.26$, $p = .040$; Figure 3b). In contrast, L-DOPA did not affect expected value representation in the 'win less on L-DOPA' group ($t(16) = 1.18$, $p = .257$), or reward representation in either the

'win less on L-DOPA' (t(16) = 1.56, p = .137) or 'win more on L-DOPA' group (t(14) = 0.48, p = .637). These results show that the restoration of a canonical prediction error signal, mediated by a more negative representation of expected value under L-DOPA, was associated with better task performance.

Although L-DOPA did not affect reward or expected value parameter estimates of those older participants in the 'win less on L-DOPA' group, these participants continued to show a negative BOLD correlate of $Q_{a(t)}(t)$ under L-DOPA even though their performance was worse on L-DOPA (Figure 3b One possibility was a differential effect of L-DOPA on the noise in the representations of R(t) and $Q_{a(t)}(t)$ for these participants. To address this, we measured individuals' standard error of the parameter estimates on L-DOPA and placebo. We found a significant negative correlation between the drug-induced change in the standard error of R(t) and $Q_{a(t)}(t)$ and the drug-induced change in total won on the task only in the 'win less on L-DOPA' group (Supplementary Fig 3). This suggests that for participants with high baseline levels of performance, L-DOPA increased noise in their reward and expected value representations and this was associated with a worsening in performance. Importantly, the increase in noise in the BOLD responses were not related to worse fits of the reinforcement learning models as mean model likelihood did not differ between groups and did not correlate with standard error of R(t) and $Q_{a(t)}(t)$ (Supplementary Table 4).

To visualise the effects of L-DOPA on reward prediction over the course of a trial, we extracted the BOLD time course from the nucleus accumbens functional ROI and performed a regression of this fMRI signal against R(t) and $Q_{a(t)}(t)$. Typically, we would expect to see a pattern of a reward 'prediction' (i.e. anticipation) at the time of the choice indicated by a positive effect of $Q_{a(t)}(t)$ and a reward 'prediction error' at the time of the outcome, indicated by both a positive effect of R(t) and negative effect of $Q_{a(t)}(t)$. As shown in Figure 3c, our time course analysis revealed exactly this expected pattern but only in the L-DOPA condition. Hence the abnormal response to the expected value observed amongst older adults on placebo (lack of reward anticipation at the time of the choice and absent negative expectation at the time of the outcome) was 'restored' when dopamine levels were enhanced. This analysis complements the aforementioned fMRI SPM analysis which showed that a canonical reward prediction error was only present on L-DOPA, by revealing abnormal expected value representations throughout the course of a trial under placebo.

Additionally we performed further multiple regression analyses across all older adults to identify regions in the brain where reward, expected value and putative reward prediction errors correlated with task performance (total money won), separately for L-DOPA and placebo conditions. Of note, only a model examining negative correlations between expected value and performance identified regions that survived FWE whole-brain correction (Supplementary Fig 4). Here we found a left superior parietal cluster in the placebo condition (Z = 5.06, peak voxel MNI co-ordinates: −26, −78, 50), and left inferior parietal (Z = 5.34, peak voxel MNI co-ordinates: −48, −49, 48) and right precuneus clusters (Z = 5.02, peak voxel MNI co-ordinates: 12, −72, 59) in the L-DOPA condition. This suggests that extra-striatal regions also influenced task performance, whereby individuals with a more negative representation of expected value in parietal regions won more money on the task.

Anatomical connectivity and reward prediction errors

Our analysis identified substantial inter-individual variability amongst older adults for both reward and expected value representations in the nucleus accumbens at baseline (i.e. under placebo) (Supplementary Fig 5), whereby the latter was associated with task performance. We hypothesised that this might be associated with the known variability in the age-related

decline of dopamine neurons from the SN/VTA, which may, in principle, be indexed through anatomical nigro-striatal connectivity. Using DTI and probabilistic tractography (n = 30 older adults) we defined a measure of connection strength between the right SN/VTA and right striatum (see Methods & Supplementary Fig 6). Nigro-striatal tract connectivity strength measured with DTI correlated with the fMRI parameter estimate under placebo associated with expected value ($Q_{a(t)}(t)$) (Spearman's rho = −.46, p = .010) but not with that associated with reward (R(t)) (Spearman's rho = .12, p = .54) (Figure 4). These correlations were significantly different from each other suggesting that individual functional activation differences of the representation of expected value but not reward were linked to anatomical connectivity strength between the SN/VTA and striatum (Fishers r-to-z transformation, z = −2.32, p = .002). This relationship between greater tract connectivity strength and more negative expected value parameter estimates remained significant after controlling for age, gender, total intracranial volume, size of the seed region from which tractography was performed and global white matter integrity indexed by fractional anisotropy (FA) (partial Spearman's rho = −0.44, p = .027). There was no difference in this correlation between subgroups of older adults ('win more on L-DOPA' group, n = 14: Rho = −0.54, p = .047; 'win less on L-DOPA' group, n = 16: Rho = −0.37, p = .154; Fisher's r-to-z transformation comparing both groups, z = 0.53, p = .596; Supplementary Fig 7). Neither FA values of SN/VTA nor nucleus accumbens functional ROI correlated with expected value (Pearson's r = .26 and r = .17, p = .16 and p = .38 respectively), suggesting that this correlation was related to circuit strength rather than local structural integrity as determined by FA.

Older participants with equivalent baseline performance levels to young adults ('win less on L-DOPA' group on placebo) had stronger connectivity between SN/VTA and the striatum than older participants with worse baseline performance than young adults ('win more on L-DOPA' group on placebo; between groups comparison t(29) = 2.40, p = .023). This suggests these older individuals had higher baseline integrity of the nigro-striatal dopamine circuit than older adults with lower baseline levels of performance.

Discussion

We used a probabilistic reinforcement learning task in combination with a pharmacological manipulation of dopamine, as well as structural and functional imaging, to probe reward-based decision-making in older age. Overall, older adults had an incomplete reward prediction error signal in the nucleus accumbens consequent upon a lack of a neuronal response to expected reward value. Baseline inter-individual differences of the expression of expected value were linked to performance and tightly coupled to nigro-striatal structural connectivity strength, determined using DTI. L-DOPA increased the task-based learning rate and modified the BOLD representation of expected value in the nucleus accumbens. Importantly this effect was only observed for those participants that showed a substantial drug-induced improvement on task performance.

Previous studies have shown that older adults perform worse on probabilistic learning tasks than their younger counterparts [2 22 23]. Since it is widely held that dopamine neurons encode a reward prediction error signal, it is conceivable that dopamine decline occurring as part of the normal aging process could account for these behavioural deficits. Indeed this was a prime motivation for our use of a pharmacological manipulation with L-DOPA. Whilst there was no significant difference in task performance in older adults as a group on placebo versus L-DOPA, we found that older adults with low baseline levels of performance improved on L-DOPA. Using a reinforcement-learning model, we showed that those older adults who performed better under L-DOPA had a higher learning rate on L-DOPA compared to placebo. This is consistent with findings in patients with Parkinson's Disease (a dopamine deficit disorder) whose learning rates when on dopaminergic medication were

higher than when off their medication, albeit in that instance without any significant difference in overall performance [11]. As in that study[11] it is impossible to make a definitive distinction between learning rate, the magnitude of the prediction error that arises from learning, and the stochastic way that learning leads to choice.

There are two important points in each trial at which a temporal difference (TD) error type signal can be anticipated namely at choice, when the TD error is the expected value of the chosen option, and at the time of outcome when the TD error is the difference between the reward actually provided and the expected value. Decomposing the outcome signal into these separate positive and negative components is important because the response to reward is highly correlated with the full prediction error, potentially readily confusing the two [19] [21] [24]. Overall in our experiment, under placebo, although the representation of the actual reward appeared normal, neither of the components of the expected value signal at choice or outcome was present in nucleus accumbens BOLD signal. This absence is consistent with the few behavioural [23] and neuroimaging studies [16, 17] that have suggested that older adults, on average, have abnormal expected value representations, although it is important to note that we did not find a substantial behavioural impairment. Most critically, we show that under L-DOPA, both components of the expected value signal were restored. However, closer inspection, taking account of individual differences in drug-induced effects on performance, revealed that this was only the case for those older adults whose performance improved under L-DOPA.

There are at least two possible explanations for the absence of the expected value signal. One is that a putative model-free decision-making system, most closely associated with neuromodulatory effects [3] [25] is impaired. This would render reward-based behaviour subject to the operation of a model-based system, which is thought to be less dependent on dopaminergic transmission [26]. This possibility is supported by evidence that older adults perform better than younger adults in tasks requiring a model of the environment (e.g. where future outcomes are dependent on previous choices) [27]. Reconciling it with the observations that suppressing [28] or boosting [29] dopamine in healthy young volunteers respectively suppresses or boosts model-based over model-free control is more of a challenge. In relation to this point, we identified two parietal clusters where expected value representations correlated with task performance in the L-DOPA condition. Interestingly, these clusters overlap with regions purported to signal 'state prediction errors' [30]. One possibility is that these regions may be a neural signature of model-based calculations, which have also recently been shown to be enhanced by L-DOPA in young participants [29]. Although previous studies have shown dopaminergic modulation of value representations in the prefrontal cortex [31], we did not find strong evidence for the involvement of any other extra-striatal regions implicated in the effects of L-DOPA on reward processing in our sample of older adults. However, L-DOPA may have also influenced other extra-striatal learning mechanisms in our task. For example, episodic learning mediated by the hippocampus has also been linked to the dopaminergic system [32], and could support aspects of rapid learning when it occurs.

Another possibility for the absence of a model-free expected value signal is that it is still calculated normally, but that when dopamine levels are low, it is not manifest in nucleus accumbens BOLD signal. One can reasonably expect that dopamine levels will impact the state of striatal neurons [33], but its impact on the BOLD signal of cortical and dopaminergic input to, and local activity within the striatum, remain unclear. In future studies, it would be interesting to use paradigms based on recent reports (e.g. [34, 35]) in older participants with and without L-DOPA to investigate the balance of model-free and model-based control.

Enriching the above picture are recent studies in healthy young participants showing that at least some aspects of the representation in striatal BOLD of the expected value component of the TD error are conditional on a requirement for action. In one such study, the representation of expected value in young adults was not modulated by L-DOPA [36]. However, it is not clear whether this is an effect of the more extensive training provided there (which can render behaviours insensitive to dopamine manipulations [37]), or the fact that the expected value did not fluctuate in a way that was relevant for choice. Together with recent findings [38], the current study raises an interesting possibility that dopamine might only modulate the neural representation of expected value when it is behaviourally relevant for the task at hand.

Our DTI connectivity analysis supports the notion that neuronal representations of expected value, and hence appropriate reward prediction error signalling, rely on the integrity of the dopaminergic system. The connectivity strength of tracts is one DTI metric reported to predict age-related performance differences [39, 40]. Importantly, older adults who under placebo performed the task as well as young controls had higher nigro-striatal connectivity strength than older adults with lower baseline levels of task performance. Furthermore, older individuals with stronger connectivity between SN/VTA and striatum had more robust value representations in the nucleus accumbens. Although our findings can be interpreted within the context of a well-defined decline of nigro-striatal dopamine neurons with increasing age [12] [13], we acknowledge that DTI measures of connectivity are not a direct mapping of dopamine neurons, but reflect white matter tract strength between the SN/VTA and striatum. Also, the direction of information flow cannot be inferred from DTI-based tractography [41]. Notably, we did not observe a relationship between FA of either the SN/VTA or striatum with functional activity in the accumbens. FA values characterise the extent of water diffusion, so providing an indirect measure of myelin, axons and the structural organisation of both grey and white matter [15] [42]. Our results are therefore an indication that inter-individual anatomical differences at the level of nigro-striatal circuit-strength rather than local grey-matter integrity within SN/VTA or striatum determine the success of prediction error signalling in healthy older adults.

In summary, the picture emerging from our study is that a subgroup of older adults who underperform at baseline can show a drug-induced improvement in task performance. Critically for these older adults, L-DOPA increased a task-based learning rate and led to a canonical reward prediction error signal by restoring the representation of expected value in the nucleus accumbens. On the other hand, participants that perform better on the task under placebo (i.e. on a par with young controls) have a greater representation of expected value in the striatum and stronger nigro-striatal connectivity, suggesting higher baseline dopamine status. After receiving L-DOPA, their performance decreases, perhaps because of increased noise in the representations of reward prediction errors. One possibility is that, in these participants, the administration of L-DOPA 'overdoses' the system, an interpretation in line both with a previously described 'inverted U-shape' (i.e. non-linear dose-dependent) impact of dopamine on cognition [20] [43] and a variable dopamine decline amongst older adults. By establishing a link between dopaminergic signalling in the nucleus accumbens and the representations of expected value in the brain our results provide a potential therapeutic route towards tackling age-related impairments in decision-making.

Materials and methods

Older subjects

32 healthy older adults aged 65 – 75 years participated in our study (see Supplementary Table 1 for details). Written informed consent was obtained from all participants. The study received ethical approval from the North West London Research Ethics Committee 2. Four

participants experienced side-effects (emesis) from L-DOPA administration. These participants remained in all analyses as they vomited more than 2.5 hours after L-DOPA ingestion, well after completion of the task and they did not feel unwell when performing the task in the scanner.

Young subjects

22 healthy young adults (mean age 25.18 yrs, SD 3.85; 12 females) were recruited via the University College London subject pool and word of mouth. Participants were screened to ensure they were healthy with no history of neurological or psychiatric disorders, no medications, no recent illicit drug use and no recent participation in other research studies involving medication. These subjects performed the task on a laptop and did not undergo MRI scanning or pharmacological manipulation.

Study procedure

This was a double-blind within-subject placebo controlled study. Older participants attended on two occasions, one week apart and performed the same task on both days, 60 minutes after ingestion of either levodopa (150mg levodopa + 37.5mg benserazide mixed in orange juice; L-DOPA) or placebo (orange juice alone), the order of which was randomised and counterbalanced. Benserazide promotes higher levels of dopamine in the brain whilst minimising peripheral side-effects such as nausea and vomiting. To achieve comparable drug absorption across individuals, subjects were instructed not to eat for up to two hours before commencing the study. Repeated physiological measurements (blood pressure and heart rate) and subjective mood rating scales were recorded under placebo and L-DOPA (Supplementary Table 5). After completing the task, on both days participants performed an unrelated episodic memory task and on one day had DTI scanning.

Two armed bandit task

All participants performed a two armed bandit ask (Figure 1a). Participants were given written and verbal instructions and undertook five practice trials before pharmacological manipulation. The probabilities of obtaining a reward for each stimulus were independent of each other and varied on a trial-to-trial basis according to a Gaussian random walk, generated using an identical procedure to Daw et al., (2006) [8]. Different pairs of fractal images were used on the two days of testing and randomly assigned amongst participants.

Reinforcement learning models

We fitted choice behaviour to a standard reinforcement learning model on a trial-by-trial basis. This involves $Q_a(t)$-values for each action $a \in \{0,1\}$ on trial t, which are updated if the subject chooses action a(t) as:

$$\begin{align} Q_{a(t)}(t+1) &= Q_{a(t)}(t) + \alpha\delta(t) \\ \delta(t) &= R(t) - Q_{a(t)}(t) \end{align}$$

Here, $Q_{a(t)}(t)$ is the expected value of the chosen option, which was set to zero at the beginning of the experiment. $\delta(t)$ is the reward prediction error (RPE) which represents the difference between the actual outcome R(t) and the expected outcome $Q_{a(t)}(t)$, where R(t) was one (win) or zero (no win). The free parameter α defined subjects' learning rate, with higher values reflecting greater weight being given to more recent outcomes and leading to a more rapid updating of expected value.

As standard, we used a softmax rule to determine the probability of choosing between the two stimuli on trial t. If $m_a(t)$ are the propensities for doing action a on trial t, this uses

$$P(a(t)=a) = \frac{\exp(\beta m_a(t))}{\exp(\beta m_0(t)) + \exp(\beta m_1(t))}$$

in which the inverse temperature parameter β indexes how deterministic choices were.

We consider two cases for $m_a(t)$. The simplest makes $m_a(t) = Q_a(t)$. However, it is often found that subjects have a tendency either to repeat or avoid doing the same action twice [44]. To account for this, we also considered a model in which $m_a(t) = Q_a(t) + b\chi_{a=a(t-1)}$ allowing an extra boost or suppression b associated with the action performed on the previous trial. We fit all sessions (L-DOPA and placebo) for each participant using expectation-maximization in a hierarchical random effects model.

It has previously been noted that it can be hard to infer both α and β independently of each other [44] [11], since it is their product that dominates behaviour in certain regimes of learning. We therefore adopted the strategy of first fitting a full random effects model as if they are independent, and then clamping β to the mean of its posterior distribution and re-inferring α using the random effects model. Amongst other things, this limits any strong claims about having inferred differences in true learning rates. Thus in the paper when we describe differences in the learning rate, we acknowledge the possible contribution of both α and β.

In a second step, we used the mean posterior β parameter at the group level obtained on the preceding step (single fixed $\beta = 1.27$ for older adults; single fixed $\beta = 1.13$ for young adults; note that data for young and older adults were analysed separately) as a fixed parameter in two, nested, RL models reflecting the two possibilities for $m_a(t)$. The first has one parameter, the learning rate α. The second has the learning rate α and the perseveration/alternation parameter b.

For older adults only, we then repeated the two steps described above but instead estimated two separate β terms for the L-DOPA and placebo conditions. We then fixed each β at their respective posterior group means ($\beta = 1.43$; $\beta = 1.10$ for the L-DOPA and placebo respectively) and proceeded as before to test the two models outlined above.

Model fitting procedure and comparison

For older adults we first compared the two full random effects model in order to choose between a model with one or with two separate β terms, and then we compared the two nested RL models described above with a single fixed beta. For completeness we also compared the same two RL models with two fixed betas (Supplementary Table 2). For young adults we compared the two RL models with a single fixed beta since young adults did not undergo pharmacological manipulation. Procedures for fitting the models were identical to those used by Huys et al. [45] and by Guitart-Masip et al. [24] and are fully described there.

Behavioural analysis

We analysed task performance (amount of money won) using two-tailed paired t-tests (L-DOPA vs. placebo in older adults) and independent t-tests (young vs. old). Reinforcement learning model parameters (learning rate and perseveration) were not normally distributed (Shapiro-Wilk and Kolmogorov-Smirnov p<0.05). Therefore we used two-tailed Wilcoxon Signed Ranks Tests to compare these parameters between the L-DOPA and placebo conditions. Two-tailed Pearson's and Spearman's correlations were used to analyse normally distributed and non-normally distributed data respectively.

We performed a median split according to difference in performance (total won on L-DOPA minus total won on placebo) amongst older adults (Supplementary Fig 1). This resulted in a group who 'win less on L-DOPA' (total won L-DOPA < placebo, n = 17) and a group who 'win more on L-DOPA' (total won L-DOPA > placebo, n = 15) .

Image acquisition

All MRI images were acquired using a 3.0T Trio MRI scanner (Siemens) using a 32-channel head coil. Functional data using echo-planar imaging was acquired on two days. On each day, scanning consisted of two runs each containing 194 volumes (matrix 64×74; 48 slices per volume; image resolution= $3 \times 3 \times 3$mm; TR 70ms, TE 30ms). Six additional volumes at the beginning of each series were acquired to allow for steady state magnetization and were subsequently discarded. Individual field maps were recorded using a double echo FLASH sequence (matrix size = 64×64; 64 slices; spatial resolution = $3 \times 3 \times 2$ mm; gap = 1 mm; short TE = 10 ms; long TE = 12.46 ms; TR = 1020 ms) and estimated using the FieldMap toolbox for distortion correction of the acquired EPI images. A structural multi-parameter map protocol employing a 3D multi-echo fast low angle shot (FLASH) sequence at 1mm isotropic resolution was used to acquire magnetization transfer (MT) weighted (echo time, TE, 2.2–14.70ms, repetition time, TR, 23.7ms, flip angle, FA, 6 degrees) and T1 weighted (TE 2.2–14.7ms, TR 18.7ms, FA 20 degrees) images. A double-echo FLASH sequence (TE1 10ms, TE2 12.46ms, $3 \times 3 \times 2$ mm resolution and 1mm gap) and B1 mapping (TE 37.06 and 55.59ms, TR 500ms, FA 230:–10:130 degrees, 4mm3 isotropic resolution) were acquired [46]. Diffusion weighted images were acquired using spin-echo echoplanar imaging, with twice refocused diffusion-encoding to reduce eddy-current-induced distortions [47]. We acquired 75 axial slices (whole brain to mid-pons) in an interleaved order [1.7 mm isotropic resolution; image matrix = 96×96, field of view = 220×220 mm2, slice thickness = 1.7 mm with no gap between slices, TR = 170 ms, TE = 103 ms, asymmetric echo shifted forward by 24 phase-encoding lines, readout bandwidth = 2003 Hz/pixel] for 61 images with unique diffusion encoding directions. The first seven reference images were acquired with a b-value of 100 s/mm^2 ('low b images'), the remaining 61 images with a b-value of 1000 s/mm^2 [48]. Two DTI sets were acquired with identical parameters except that the second was acquired with a reversed k-space readout direction allowing removal of susceptibility artefacts post-processing [49]. Since the SN/VTA was a major region of interest, we optimised the quality of our images by using pulse-gating to minimize pulsation artefact within the brainstem.

fMRI data analysis

Data were analysed using SPM8 (Wellcome Trust Centre for Neuroimaging, UCL, London). Pre-processing included bias correction, realignment, unwarping using individual fieldmaps, co-registration and spatial normalization to the Montreal Neurology Institute (MNI) space using diffeomorphic registration algorithm (DARTEL)[50] with spatial resolution after normalization of $2 \times 2 \times 2$mm. Data were smoothed with a 6mm FWHM Gaussian kernel. The fMRI time series data were high-pass filtered (cutoff = 128 s) and whitened using an AR(1)-model. For each subject a statistical model was computed by applying a canonical hemodynamic response function (HRF) combined with time and dispersion derivatives.

Our study builds on a wealth of pre-existing literature describing RPEs in healthy young adults performing probabilistic learning tasks, whereby the striatum signals the difference between actual and expected rewards (e.g. [7] [51]). A region showing a 'canonical' RPE should show a positive correlation with reward (R(t)) and a negative correlation with expected value ($Q_{a(t)}(t)$). This is different from a region showing a 'putative' prediction error which shows a correlation with (R(t)– $Q_{a(t)}(t)$). Studies in young participants have shown a canonical RPE signal with both a positive response to reward and negative response to

expected value in value learning tasks [19]. Separating the RPE into its components has often not been done, although recent studies have shown that since R(t) is highly correlated with the RPE, a correlation between BOLD and $(R(t) - Q_{a(t)}(t))$ may lead to false positive results suggesting that areas whose BOLD only correlates with R(t) may be thought as representing a RPE. This is an important distinction that allows us to determine whether, as a group, healthy older adults do not represent all components of a canonical RPE signal, and the extent to which the two components $(R(t); Q_{a(t)}(t))$ may be related to (i) task performance (ii) the integrity of anatomical connectivity between the dopaminergic midbrain and the striatum, and (iii) modulation by L-DOPA.

The general linear model for each subject at the 1st level consisted of regressors at the time of stimulus display separately for when a choice was made, when no choice was made and at the time of stimulus outcome. All trials in which participants made a response and obtained an outcome were included in the analysis. BOLD responses to outcomes were parametrically modelled with two separate parametric regressors (R(t) and $(Q_{a(t)}(t))$): R(t) was a binary regressor with value 0 when no reward was obtained and value 1 when a reward was obtained; $Q_{a(t)}(t)$ included the expected value on that trial and was built using the group posterior mean α distribution from the winning model). Therefore, a brain region that correlates with R(t) indicates that BOLD responses are higher after a reward is obtained when compared to no reward. Separate design matrices were calculated for the LDOPA and placebo conditions. To capture residual movement-related artefacts, six covariates were included (the three rigid-body translation and three rotations resulting from realignment) as regressors of no interest. Finally we also included 18 regressors for cardiac and respiratory phases in order to correct for physiological noise.

At the first level, we implemented the contrasts R(t), $Q_{a(t)}(t)$, which are the individual components of the RPE; and the putative RPE $R(t) > Q_{a(t)}(t)$. At the second level, we first defined a functional ROI with the contrast $R(t) > Q_{a(t)}(t)$ collapsed across L-DOPA and placebo conditions. We used an uncorrected threshold of p<0.001 to produce a whole-brain statistical parametric map of regions encoding putative RPEs from which we identified a region in the right nucleus accumbens. We used a subject-derived anatomical mask (Supplementary Fig 2) to constrain this functional ROI.

We examined the effects of placebo and L-DOPA on the reward (R(t)) and expected value $(Q_{a(t)}(t))$ components of the canonical RPE signal within the functional ROI. Here we used the Marsbar toolbox [52] to extract the parameter estimates from the functional ROI to enter into a two $(R/Q_{a(t)})$ by two (L-DOPA/placebo) repeated measures ANOVA. We conducted post hoc tests to characterise the impairment in expected value representation (one-tailed one-sample t-tests for each condition to test the null hypothesis that they are not different from zero, and two-tailed paired t-tests to compare the effect of L-DOPA to placebo). Furthermore, for each participant we measured the standard error of the parameter estimates for reward (R(t)) and expected value $(Q_{a(t)}(t))$ on L-DOPA and placebo and calculated the drug-induced change (L-DOPA minus placebo). We performed two-tailed Spearman correlations between these measures and the drug-induced change in task performance (total won L-DOPA minus placebo) (Supplementary Fig 3).

For completeness, we also performed a whole-brain voxel-based analysis of drug effects (L-DOPA > placebo) across participants using the contrasts defined at the first level for reward, expected value and the putative RPE (no regions survived whole-brain FWE correction; Supplementary Table 6). We also performed separate multiple regression analyses for the same contrasts on the placebo and L-DOPA conditions separately using task performance (total won) as a between-subjects regressor .

Time course extraction

The main aim of this analysis was to visualise the effect of reward and expected value on the BOLD signal, at the time of the choice and at the time of the outcome, from the nucleus accumbens functional ROI over the course of a trial. In the fMRI SPM analysis it was not possible to simultaneously test for the effects of value expectation on the choice and the outcome phases. This is because the time of the choice and the time of the outcome were very close together in time (3s apart) and including the same parametric modulator on both time points would have resulted in highly correlated regressors. Thus, although the SPM model included regressors at the time of the choice and time of the outcome, we only included parametric modulators at the time of the outcome, so focussing on outcome prediction errors only.

Time courses were extracted from preprocessed data in MNI space. We upsampled the extracted BOLD signal to 100 ms. The signal was divided into trials and resampled to a duration of 15 s with the onset (presentation of the stimuli) occurring at 0s, the time of the choice occurring between 0–2s and the time of the outcome at 3s. We then estimated a general linear model across trials at every time point in each subject independently, where reward and expected value were the regressors of interest. These regressors were not orthogonalised and therefore competed for variance which is a particularly stringent test [19]. We calculated group mean effect sizes at each time point and their standard errors, plotted separately for the placebo and L-DOPA conditions.

Diffusion tensor imaging (DTI) connectivity strength analysis

One participant was unable to tolerate scanning therefore DTI data was collected from 31 older individuals. DTI tractography and generation of 'relative connectivity strength' maps was performed as per Forstmann et al. [53] and summarised in Supplementary Figure 6. The aim of our tractography analysis was to determine if inter-individual differences in nigrostriatal connectivity influenced the observed baseline variability in functional prediction error signalling (Supplementary Figure 5). We used Spearman's correlations to relate connectivity strength to prediction error signalling (R(t) and $Q_{a(t)}(t)$ parameter estimates from the functional nucleus accumbens ROI) [54]. To identify outliers we converted connectivity strength to z-scores (conventionally defined as $z < –3$ or $z > 3$). Although none of our participants were outside this range, one participant had a z-score of 2.83 (equivalent to connectivity strength = 0.006) and was therefore excluded from the reported results (therefore reported DTI results are for n = 30). Even so, including this potential outlier in the analysis did not change the results. We performed partial Spearman's correlations with the following covariates: age, gender, total intracranial volume, size of the manually defined seed (right SN/VTA) region and global white matter integrity. Note that the size of the target region was not included since this was the same for all participants. Global white matter integrity was measured by segmenting FA maps and calculating mean FA values of the white matter FA maps. As a control we performed additional Pearson's correlations between FA values of the SN/VTA seed or nucleus accumbens functional ROI (FA values were normally distributed) and $Q_{a(t)}(t)$ on placebo.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Eppinger B, Hämmerer D, Li S-C. Neuromodulation of reward-based learning and decision making in human aging. Annals of the New York Academy of Sciences. 2011; 1235:1–17. [PubMed: 22023564]

2. Mell T, et al. Effect of aging on stimulus-reward association learning. Neuropsychologia. 2005; 43:554–563. [PubMed: 15716145]

3. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–1599. [PubMed: 9054347]

4. Salamone JD, Correa M, Mingote SM, Weber SM. Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. Current Opinion in Pharmacology. 2005; 5:34–41. [PubMed: 15661623]

5. Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal Pathways in Primates Form an Ascending Spiral from the Shell to the Dorsolateral Striatum. J. Neurosci. 2000; 20:2369–2382. [PubMed: 10704511]

6. O'Doherty J, et al. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. Science. 2004; 304:452–454. [PubMed: 15087550]

7. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal Difference Models and Reward-Related Learning in the Human Brain. Neuron. 2003; 38:329–337. [PubMed: 12718865]

8. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006; 441:876–879. [PubMed: 16778890]

9. Knutson B, Gibbs S. Linking nucleus accumbens dopamine and blood oxygenation. Psychopharmacology. 2007; 191:813–822. [PubMed: 17279377]

10. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature. 2006; 442:1042–1045. [PubMed: 16929307]

11. Rutledge RB, et al. Dopaminergic Drugs Modulate Learning Rates and Perseveration in Parkinson's Patients in a Dynamic Foraging Task. The Journal of Neuroscience. 2009; 29:15104–15114. [PubMed: 19955362]

12. Bäckman L, Nyberg L, Lindenberger U, Li S-C, Farde L. The correlative triad among aging, dopamine, and cognition: Current status and future prospects. Neuroscience & Biobehavioral Reviews. 2006; 30:791–807. [PubMed: 16901542]

13. Duzel E, Bunzeck N, Guitart-Masip M, Duzel S. NOvelty-related Motivation of Anticipation and exploration by Dopamine (NOMAD): Implications for healthy aging. Neuroscience & Biobehavioral Reviews. 2010; 34:660–669. [PubMed: 19715723]

14. Fearnley JM, Lees AJ. Ageing and Parkinson's Disease: substantia nigra regional selectivity. Brain. 1991; 114:2283–2301. [PubMed: 1933245]

15. Vaillancourt DE, Spraker MB, Prodoehl J, Zhou XJ, Little DM. Effects of aging on the ventral and dorsal substantia nigra using diffusion tensor imaging. Neurobiology of Aging. 2012; 33:35–42. [PubMed: 20359780]

16. Samanez-Larkin GR, Kuhnen CM, Yoo DJ, Knutson B. Variability in Nucleus Accumbens Activity Mediates Age-Related Suboptimal Financial Risk Taking. J. Neurosci. 2010; 30:1426–1434. [PubMed: 20107069]

17. Schott BH, et al. Ageing and early-stage Parkinson's disease affect separable neural mechanisms of mesolimbic reward processing. Brain. 2007; 130:2412–2424. [PubMed: 17626038]

18. Cox KM, Aizenstein HJ, Fiez JA. Striatal outcome processing in healthy aging. Cogn Affect Behav Neurosci. 2008; 8:304–317. [PubMed: 18814467]

19. Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS. Associative learning of social value. Nature. 2008; 456:245–249. [PubMed: 19005555]

20. Cools R, D'Esposito M. Inverted-U Shaped Dopamine Actions on Human Working Memory and Cognitive Control. Biological Psychiatry. 2011; 69:e113–e125. [PubMed: 21531388]

21. Li J, Daw ND. Signals in Human Striatum Are Appropriate for Policy Update Rather than Value Prediction. The Journal of Neuroscience. 2011; 31:5504–5511. [PubMed: 21471387]

22. Eppinger B, Kray J, Mock B, Mecklinger A. Better or worse than expected? Aging, learning, and the ERN. Neuropsychologia. 2008; 46:521–539. [PubMed: 17936313]

23. Samanez-Larkin GR, Wagner AD, Knutson B. Expected value information improves financial risk taking across the adult life span. Social Cognitive and Affective Neuroscience. 2011; 6:207–217. [PubMed: 20501485]

24. Guitart-Masip M, et al. Go and no-go learning in reward and punishment: Interactions between affect and effect. Neuroimage. 2012; 62:154–166. [PubMed: 22548809]

25. Dickinson, A.; Balleine, B. The role of learning in the operation of motivational systems. John Wiley & Sons; New York: 2002.

26. Dickinson A, Smith J, Mirenowicz J. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav Neurosci. 2000; 114:468–483. [PubMed: 10883798]

27. Worthy DA, Gorlick MA, Pacheco JL, Schnyer DM, Maddox WT. With Age Comes Wisdom. Psychological Science. 2011; 22:1375–1380. [PubMed: 21960248]

28. de Wit S, et al. Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. Psychopharmacology (Berl). 2012; 219:621–631. [PubMed: 22134475]

29. Wunderlich K, Smittenaar P, Dolan Raymond J. Dopamine Enhances Model-Based over Model-Free Choice Behavior. Neuron. 2012; 75:418–424. [PubMed: 22884326]

30. Glascher J, Daw N, Dayan P, O'Doherty JP. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. Neuron. 2010; 66:585–595. [PubMed: 20510862]

31. Jocham G, Klein TA, Ullsperger M. Dopamine-Mediated Reinforcement Learning Signals in the Striatum and Ventromedial Prefrontal Cortex Underlie Value-Based Choices. The Journal of Neuroscience. 2011; 31:1606–1613. [PubMed: 21289169]

32. Shohamy D, Adcock RA. Dopamine and adaptive memory. Trends in Cognitive Sciences. 2010; 14:464–472. [PubMed: 20829095]

33. Nicola SM, Surmeier DJ, Malenka RC. Dopaminergic Modulation of Neuronal Excitability in the Striatum and Nucleus Accumbens. Annual Review of Neuroscience. 2000; 23:185–215.

34. Daw, Nathaniel D, Gershman, Samuel J, Seymour B, Dayan P, Dolan, Raymond J. Model-Based Influences on Humans' Choices and Striatal Prediction Errors. Neuron. 2011; 69:1204–1215. [PubMed: 21435563]

35. Simon DA, Daw ND. Neural Correlates of Forward Planning in a Spatial Decision Task in Humans. The Journal of Neuroscience. 2011; 31:5526–5539. [PubMed: 21471389]

36. Guitart-Masip M, et al. Action controls dopaminergic enhancement of reward representations. Proceedings of the National Academy of Sciences. 2012; 109:7511–7516.

37. Choi WY, Balsam PD, Horvitz JC. Extended Habit Training Reduces Dopamine Mediation of Appetitive Response Expression. The Journal of Neuroscience. 2005; 25:6729–6733. [PubMed: 16033882]

38. Klein-Flugge, Miriam C, Hunt, Laurence T, Bach, Dominik R, Dolan, Raymond J, Behrens, Timothy EJ. Dissociable Reward and Timing Signals in Human Midbrain and Ventral Striatum. Neuron. 2011; 72:654–664. [PubMed: 22099466]

39. Coxon JP, Van Impe A, Wenderoth N, Swinnen SP. Aging and Inhibitory Control of Action: Cortico-Subthalamic Connection Strength Predicts Stopping Performance. The Journal of Neuroscience. 2012; 32:8401–8412. [PubMed: 22699920]

40. Forstmann BU, et al. Cortico-subthalamic white matter tract strength predicts interindividual efficacy in stopping a motor response. Neuroimage. 2012; 60:370–375. [PubMed: 22227131]

41. Le Bihan D, Johansen-Berg H. Diffusion MRI at 25: Exploring brain tissue structure and function. Neuroimage. 2012; 61:324–341. [PubMed: 22120012]

42. Draganski B, et al. Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ). Neuroimage. 2011; 55:1423–1434. [PubMed: 21277375]

43. Li S-C, Sikström S. Integrative neurocomputational perspectives on cognitive aging, neuromodulation, and representation. Neuroscience & Biobehavioral Reviews. 2002; 26:795–808. [PubMed: 12470691]

44. Schonberg T, Daw ND, Joel D, O'Doherty JP. Reinforcement Learning Signals in the Human Striatum Distinguish Learners from Nonlearners during Reward-Based Decision Making. The Journal of Neuroscience. 2007; 27:12860–12867. [PubMed: 18032658]

45. Huys QJM, et al. Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. PLoS Comput Biol. 2011; 7:e1002028. [PubMed: 21556131]

46. Lutti A, Hutton C, Finsterbusch J, Helms G, Weiskopf N. Optimization and validation of methods for mapping of the radiofrequency transmit field at 3T. Magnetic Resonance in Medicine. 2010; 64:229–238. [PubMed: 20572153]

47. Reese TG, Heid O, Weisskoff RM, Wedeen VJ. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magnetic Resonance in Medicine. 2003; 49:177–182. [PubMed: 12509835]

48. Nagy Z, Weiskopf N, Alexander DC, Deichmann R. A method for improving the performance of gradient systems for diffusion-weighted MRI. Magnetic Resonance in Medicine. 2007; 58:763–768. [PubMed: 17899604]

49. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage. 2003; 20:870–888. [PubMed: 14568458]

50. Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage. 2007; 38:95–113. [PubMed: 17761438]

51. Knutson B, Cooper JC. Functional magnetic resonance imaging of reward prediction. Current Opinion in Neurology. 2005; 18:411–417. [PubMed: 16003117]

52. Brett, M.; Anton, J-L.; Valabregue, R.; Poline, J-B. Region of interest analysis using an SPM toolbox. 8th International Conference on Functional Mapping of the Human Brain; Sendai, Japan. 2002.

53. Forstmann BU, et al. The Speed-Accuracy Tradeoff in the Elderly Brain: A Structural Model-Based Approach. The Journal of Neuroscience. 2011; 31:17242–17249. [PubMed: 22114290]

54. Schwarzkopf DS, de Haas B, Rees G. Better ways to improve standards in brain-behavior correlation analysis. Frontiers in Human Neuroscience. 2012; 6
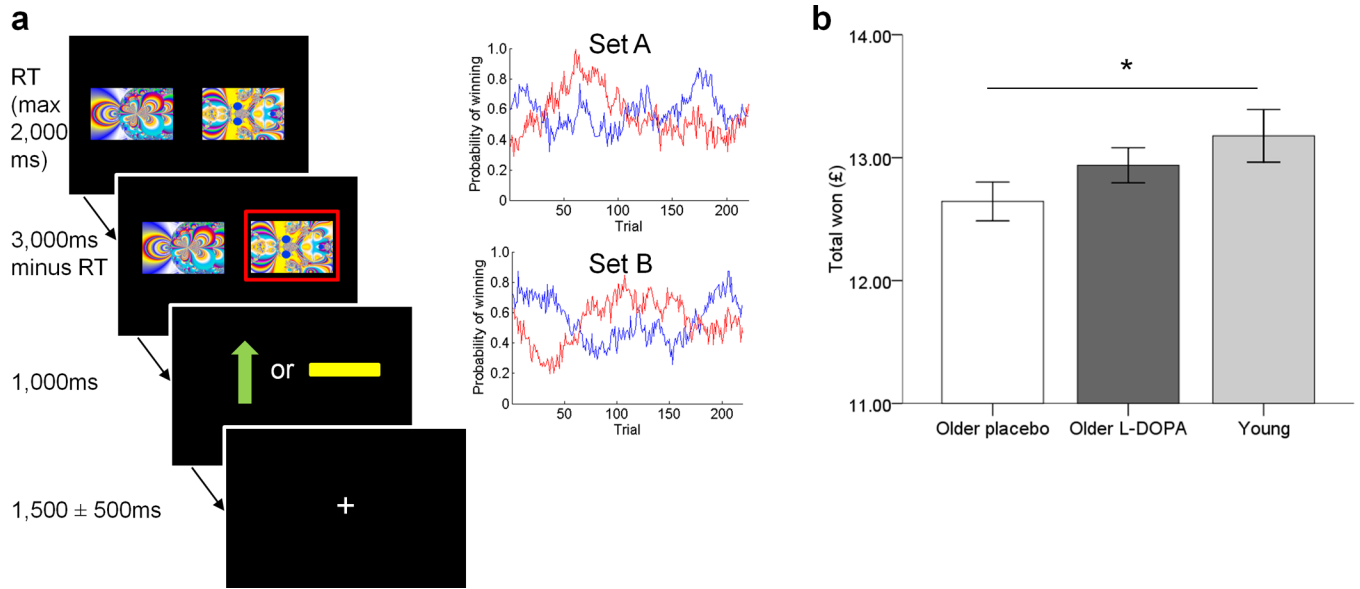
Figure 1. Two armed bandit task design and performance in young and older adults

a: On each trial, participants selected one of two fractal images which was then highlighted in a red frame. This was followed by an outcome where a green upward arrow indicated a win of 10 pence and a yellow horizontal bar indicated the absence of a win. If they did not choose a stimulus, the written message "you did not choose a picture" was displayed. The same pair of images was used throughout the task, although their position on the screen (left or right) varied. The task consisted of 220 trials separated into two sessions with a short break in between. Participants' earnings were displayed at the end of the task and given to them at the end of the test day. The probability of obtaining a reward associated with each image varied on a trial-by-trial basis according to a Gaussian random walk. Two different sets of probability distributions (Set A and Set B) were used on the two testing days, counterbalanced across the order of L-DOPA/placebo administration.

b: Older adults (n = 32) in the placebo condition won less money than young adults (n = 22). When the same older adults (n = 32) received L-DOPA, performance was similar to young adults. *p<0.05. Error bars indicate ±1SEM.
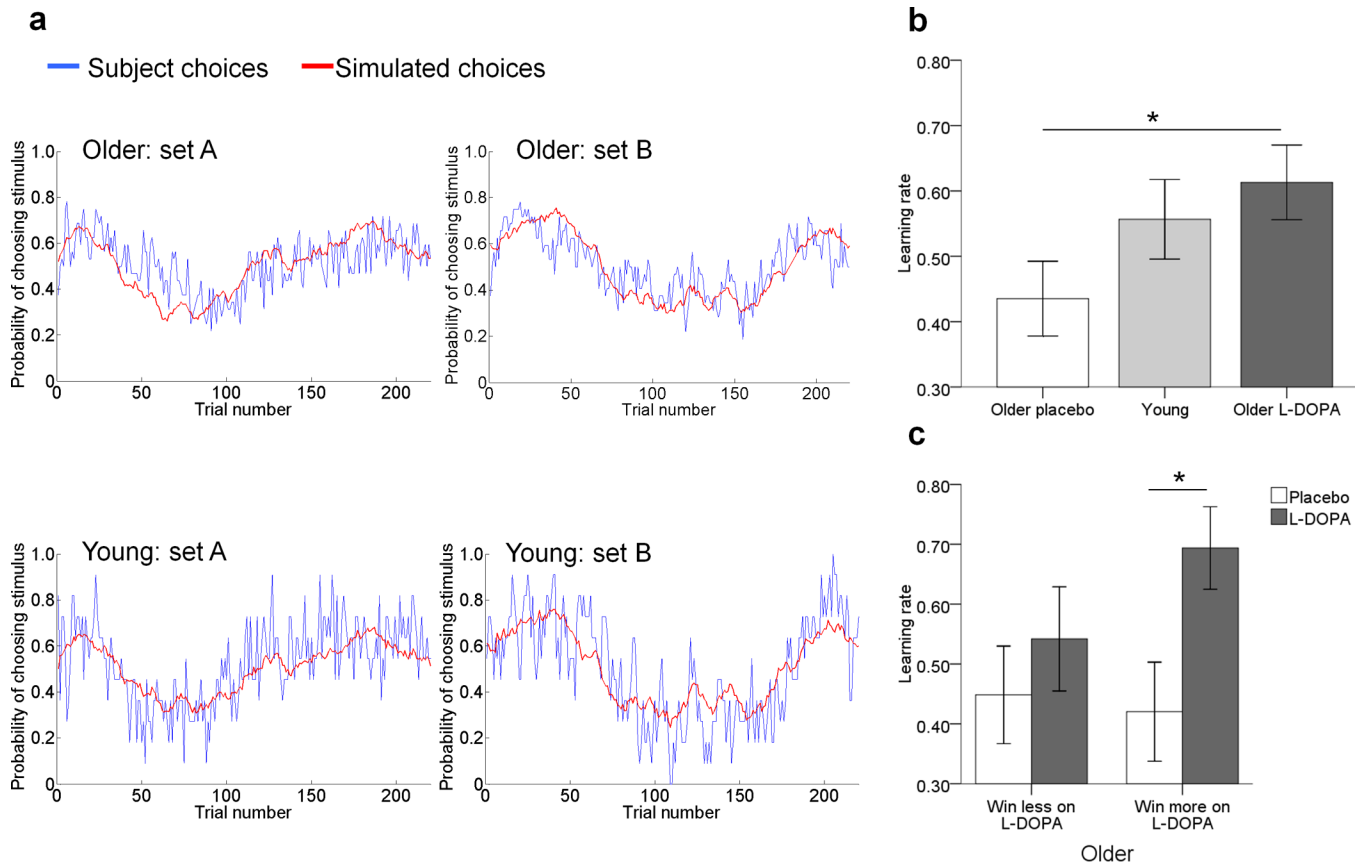
Figure 2. Reinforcement learning model and behaviour

a: For young and older adults, the predicted choices from the learning model (red) closely matched subjects' observed choices (blue). The red lines show the same time-varying probabilities, but evaluated on choices sampled from the model (see methods). Plots are shown for the two different sets of probability distributions used on the two test days.

b: Older adults (n = 32) had a higher learning rate under L-DOPA compared with placebo and did not differ from young adults (n = 22). *p<0.05 two-tailed. Error bars are ±1 SEM.

c: Older adults who won more on L-DOPA than placebo ('win more on L-DOPA', n = 15) had a significantly higher learning rate under L-DOPA than placebo, whereas learning rates did not differ between placebo and L-DOPA for older adults who won less on L-DOPA than placebo ('win less on L-DOPA', n = 17). *p<0.05 two-tailed. Error bars are ±1 SEM.
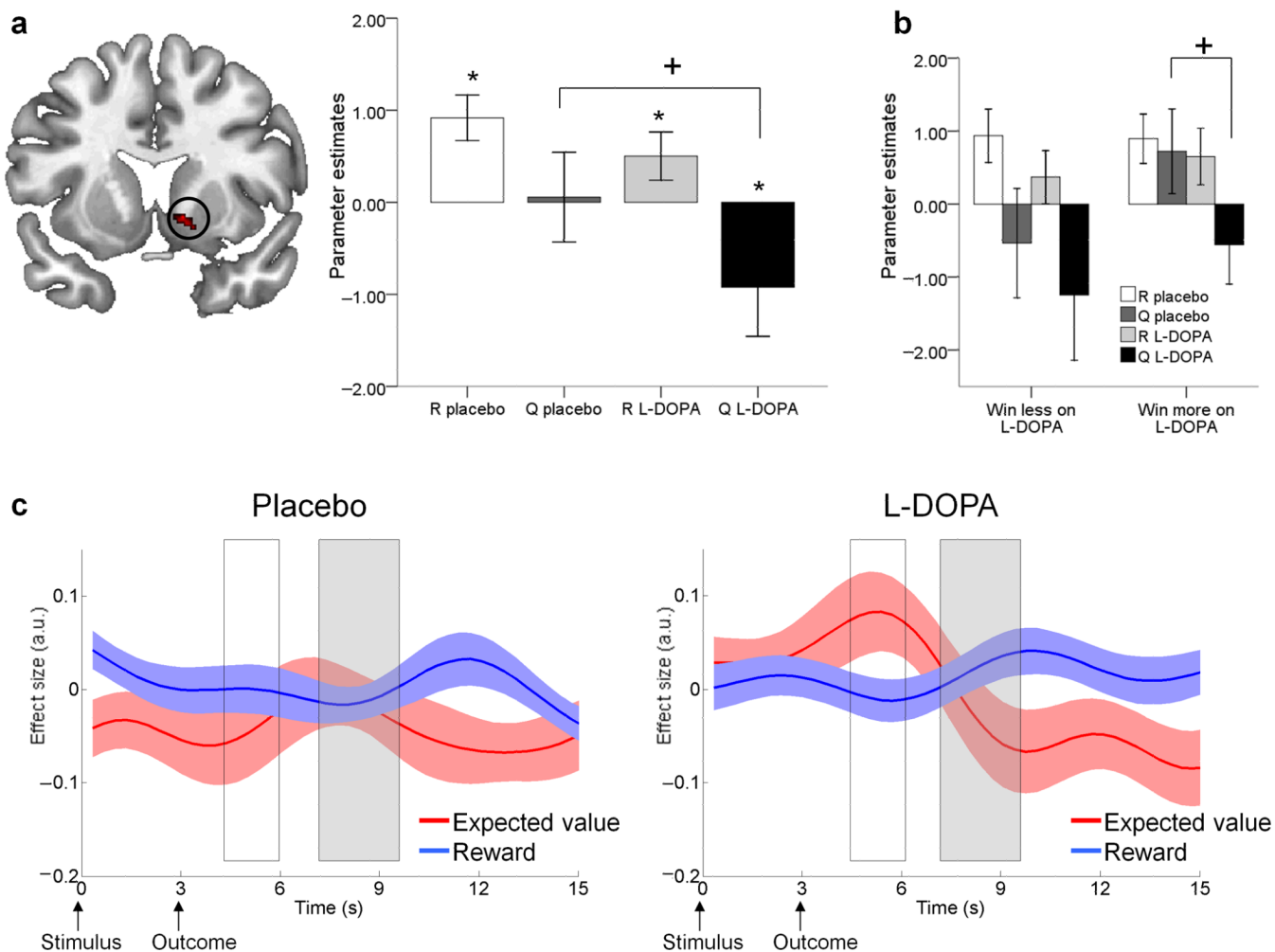
Figure 3. Reward prediction in the nucleus accumbens in 32 older adults

a: A region in the right nucleus accumbens showed greater BOLD activity for reward (R) than for expected value (Q) at the time of outcome ('putative' reward prediction error). However, the lack of a negative effect of Q under placebo meant this prediction error signal was incomplete (*one sample t-test p<0.05 one-tailed). L-DOPA increased the negative effect of Q (paired t-test, +p < 0.05 two-tailed) resulting in a 'canonical' prediction error signal (both a positive effect of R and negative effect of Q). Bars ±1 SEM.

b: Participants who 'win more on L-DOPA' (n = 15) only demonstrated a negative effect of Q under L-DOPA and not placebo (+paired t-test p < 0.05 two-tailed). R and Q parameter estimates did not differ between L-DOPA and placebo for participants who 'win less on L-DOPA', n = 17. Bars ±1 SEM.

c: Time course plots of the nucleus accumbens BOLD response to reward and expected value. White box corresponds with BOLD responses elicited at the time participants' made a choice; grey box corresponds with BOLD responses elicited when the outcomes were revealed. Under placebo the only reliable signal observed was a reward response. Under L-DOPA, a canonical reward prediction error was observed, involving a positive expectation of value at the time of the choice together with a positive reward response and a negative expectation of value at the time of the outcome. Reward anticipation (positive effect at the time of the choice) was only observed on L-DOPA. Solid lines are group means of the effect sizes, shaded areas represent ±1 SEM.
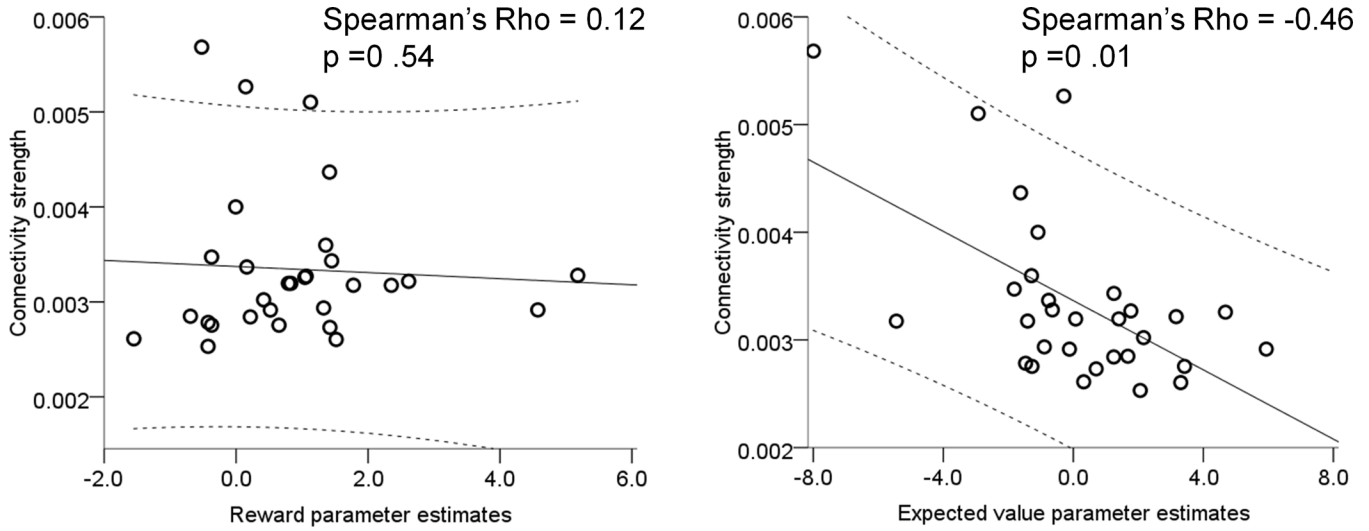
Figure 4. Nigro-striatal tract connectivity strength and functional prediction errors
Under placebo, individuals with higher white matter nigro-striatal tract connectivity strength (determined using DTI) had a more negative effect of expected value whereas there was no correlation with functional parameters estimates of reward. Each dot on the plots represents one subject (n = 30; note two participants are overlapping on the plot on the left), the solid line is the regression slope, dashed lines represent 95% confidence intervals.