Impaired context-sensitive adjustment of behaviour in Parkinson's disease patients tested on and off medication: An fMRI study

Ima Trempler [a,b,*], Paul-Christian Bürkner [a], Nadiya El-Sourani [a], Ellen Binder [c,d], Paul Reker [c], Gereon R. Fink [c,d], Ricarda I. Schubotz [a,b,c]

[a] Department of Psychology, University of Muenster, 48149, Münster, Germany
[b] Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Muenster, 48149, Münster, Germany
[c] Faculty of Medicine and University Hospital Cologne, Department of Neurology, 50937, Cologne, Germany
[d] Institute of Neuroscience and Medicine (INM3), Cognitive Neuroscience, Research Centre Jülich, 52425, Jülich, Germany

ABSTRACT

The brain's sensitivity to and accentuation of unpredicted over predicted sensory signals plays a fundamental role in learning. According to recent theoretical models of the predictive coding framework, dopamine is responsible for balancing the interplay between bottom-up input and top-down predictions by controlling the precision of surprise signals that guide learning.

Using functional MRI, we investigated whether patients with Parkinson's disease (PD) show impaired learning from prediction errors requiring either adaptation or stabilisation of current predictions. Moreover, we were interested in whether deficits in learning over a specific time scale would be accompanied by altered surprise responses in dopamine-related brain structures. To this end, twenty-one PD patients tested on and off dopaminergic medication and twenty-one healthy controls performed a digit prediction paradigm. During the task, violations of sequence-based predictions either signalled the need to update or to stabilise the current prediction and, thus, to react to them or ignore them, respectively. To investigate contextual adaptation to prediction errors, the probability (or its inverse, *surprise*) of the violations fluctuated across the experiment.

When the probability of prediction errors over a specific time scale increased, healthy controls but not PD patients off medication became more flexible, i.e., error rates at violations requiring a motor response decreased in controls but increased in patients. On the neural level, this learning deficit in patients was accompanied by reduced signalling in the substantia nigra and the caudate nucleus. In contrast, differences between the groups regarding the probabilistic modulation of behaviour and neural responses were much less pronounced at prediction errors requiring only stabilisation but no adaptation. Interestingly, dopaminergic medication could neither improve learning from prediction errors nor restore the physiological, neurotypical pattern.

Our findings point to a pivotal role of dysfunctions of the substantia nigra and caudate nucleus in deficits in learning from flexibility-demanding prediction errors in PD. Moreover, the data witness poor effects of dopaminergic medication on learning in PD.

1. Introduction

To behave adaptively, we need to adjust our expectations to persistent environmental changes while sustaining the pursuit of our action goals despite temporary distractions. Environmental changes that do not match our expectations, i.e., prediction errors, are known to cause phasic dopamine signalling in the midbrain. Thereby, they trigger bottom-up processing guiding the adjustment of predictions and the initiation of behaviour (Schultz and Dickinson, 2000; Redgrave and Gurney, 2006;

Murty et al., 2011; D'Ardenne et al., 2012). This adjustment pertains to learning since future surprise can be minimised and behavioural implications of prediction errors become more predictable (Fiser et al., 2010; Friston et al., 2014).

The influence of predictions on behaviour is suggested to be regulated by tonic dopamine action, determining the relative weight or *precision* of bottom-up prediction errors for top-down predictions (Friston et al., 2012). By regulating phasic dopamine release, tonic dopamine has been found to modulate the surprise-driven learning rate, biases action

---

selection and sets transition thresholds between flexible and stable states, favouring either bottom-up sensory input or top-down predictions (Beeler et al., 2010; Beeler et al., 2012; Humphries et al., 2012; Yu et al., 2013; Barter et al., 2015).

During learning, predictions are adapted to a particular context, based on the probabilistic structure of the past (Behrens et al., 2007). Here, the cholinergic and the noradrenergic system is ascribed a role in computations of uncertainty (Yu and Dayan, 2005; Marshall et al., 2016). However, depending on whether a current context consists of prediction violations that either signal the need for adapting to lasting changes of the environment (flexibility-demanding prediction errors, hereafter) or are caused by temporary chance occurrences of uncommon events (stability-demanding prediction errors), the adoption of either a more flexible or stable state is required, respectively. The interplay of flexibility and stability in response selection is suggested to be balanced by the levels of dopamine in frontostriatal circuits (Williams and Castner, 2006; Cools and D'Esposito, 2011; Doll et al., 2011). Specifically, the striatum assumes the role of a gate for relevant versus irrelevant input into working memory (Badre, 2012; Keeler et al., 2014). High or low energy barriers of attractor networks within the prefrontal cortex are suggested to facilitate either shielding or updating of working memory representations (Durstewitz et al., 2000; Durstewitz and Seamans, 2008). These barriers are probably adjusted contingent on a prioritisation of perceptual input utilising relevance (Summerfield and Egner, 2009; Rauss et al., 2011).

Accordingly, previous studies found that patients with Parkinson's disease (PD), a disorder associated with loss of dopamine neurons in the substantia nigra, show difficulties in the selection and inhibition of motor responses (Wylie et al., 2009; Wylie et al., 2010) as well as in cognitive set-shifting, i.e., shifting or switching between particular stimulus-response links (Cools et al., 2001; Monchi et al., 2004). Moreover, Galea et al. (2012) showed that during action reprogramming requiring a switch from an expected to an unexpected response, PD patients show increased reaction times to unexpected events in contexts of predictable compared to unpredictable environments.

In the present study, we tested the idea that a dopamine deficit impairs response selection by impeding probabilistic inference over either flexible or stable states. To this end, we examined learning from different levels of probabilities of flexibility-demanding versus stability-demanding prediction errors in healthy controls and patients with akinetic-rigid PD on and off dopaminergic therapy during a digit prediction task. We used fMRI to assess whether activity in key dopaminergic regions varies as a function of learning from prediction errors, with altered signalling in PD patients. During the task, participants were required to indicate the occurrence of digit rule *switches*, as behaviourally *relevant* violations leading to an update of the predictive rule (revealing flexibility), and to ignore short interruptions, referred to as *drifts* hereafter, as behaviourally *irrelevant* prediction errors provoking a shielding of the predictive rule (revealing stability). Importantly, both the absolute frequency of switch and drift occurrences (prediction errors: predicted digits) as well as the relative proportion of switch and drift occurrences (switches: drifts) changed over time. Varying probability and predictability of these events were quantified as decay-dependent information-theoretic quantities, i.e., *surprise* and *entropy*, respectively (Harrison et al., 2011; see Methods for further details). Our hypotheses were focused on the effects of decay-dependent surprise, with the latter modelled as a regressor for analysing both behavioural performance as well as BOLD time series to assess learning from prediction errors on a trial-by-trial basis. Decay-dependent entropy was additionally modelled as a regressor of nuisance.

At first, behavioural data enabled us to test whether probabilistic inference to derive responses to different types of prediction errors differs between the groups. We expected that PD patients off medication would have problems to learn from prediction errors, i.e., to adopt flexible and stable states depending on the switch and drift probability, respectively (Friston et al., 2012) (Hypothesis 1, H1). On the neural level, we were

particularly interested in the contribution of brain regions, which are known to be rich in dopamine, i.e., the substantia nigra and the caudate nucleus. More specifically, we hypothesised that substantia nigra activity would be positively modulated by both switch and drift surprise in controls but not in PD patients off medication, reflecting phasic dopamine release independent of the events' identity (Redgrave and Gurney, 2006; Hölig and Berti, 2010) (Hypothesis 2, H2). In contrast, we hypothesised an increased activity of the caudate nucleus as a function of switch but not drift probability in controls compared with PD patients due to this regions' role in adaptive motor responses to prediction errors modulated by tonic dopamine (Bestmann et al., 2008; Galea et al., 2012; Marshall et al., 2016) (Hypothesis 3, H3). Finally, provided that dopamine is indeed substantially involved in these effects, learning from prediction errors should improve with dopaminergic medication in PD patients, also reflected in a restored surprise-dependent modulation of neural activation within the respective regions in patients on compared with patients off medication (Hypothesis 4, H4).

2. Materials and methods

2.1. Participants

Our sample was the same as reported in Trempler et al. (2018). 21 patients (6 females, mean age = 58.81 years, SD = 9.89, range = [40, 72]) meeting the United Kingdom Parkinson's Disease (UKPD) Society Brain Bank Criteria for idiopathic Parkinson's disease (Hughes et al., 1992) were recruited via the neurologic outpatient clinic of the University Hospital of Cologne, Germany. Hoehn and Yahr ratings ranged between I and III under regular medication (Hoehn and Yahr, 1967). During the screening session, the severity of symptoms was further defined according to the motor score of the Unified Parkinson's Disease Rating Scale (UPDRS) (Fahn and Elton, 1987). Based on the judgment of an experienced movement disorder specialist, only patients of the akinetic-rigid subtype were selected. This way, a clinically homogenous group could be ensured and potential movement artefacts could be minimised. Moreover, all participants scored between 19 and 30 points in the Parkinson Neuropsychometric Dementia Assessment (PANDA; 18–30 points = "age adequate cognitive performance") (Kalbe et al., 2008) and lower than 19 points in the Beck depression inventory-II (BDI-II; cut-off for depression: $\geq$ 20 points) (Hautzinger et al., 2006). The screening included a training session to ensure that patients would be able to perform the task under their regular dopaminergic medication.

Patients were tested twice, i.e., once with their regular medication ("ON"-state) and once without medication ("OFF"-state; after overnight withdrawal of dopaminergic medication, corresponding to at least 10 h after the last dose). Session order (OFF-ON and ON-OFF) was counterbalanced across the participants. Withdrawal affected motor performance as seen in a significant difference in UPDRS-score between patients ON (M = 19.62, SD = 7.48) and OFF (M = 27.14, SD = 9.46), $t(20)$ = 10.61, p < 0.001. A group of 21 healthy participants (6 females, mean age = 60.05 years, SD = 10.05, range = [36, 74]) matched to the patients regarding age and gender served as control subjects. Healthy controls did not receive any medication. They performed the training, the experiment, and all additional assessments on one day. No participant had undergone neurosurgical treatment for the disease or had a history of other neurological or psychiatric diseases.

The study was performed following the Declaration of Helsinki and had been approved by the ethics committee of the Medical Faculty of the University Hospital Cologne, Germany. Each participant submitted a signed informed consent notification and received reimbursement for participation plus travel expenses afterwards.

2.2. Task

During the task, a digit sequence was visually presented at the centre of a computer screen, in either ascending (1–2 – 3–4) or descending (4–3
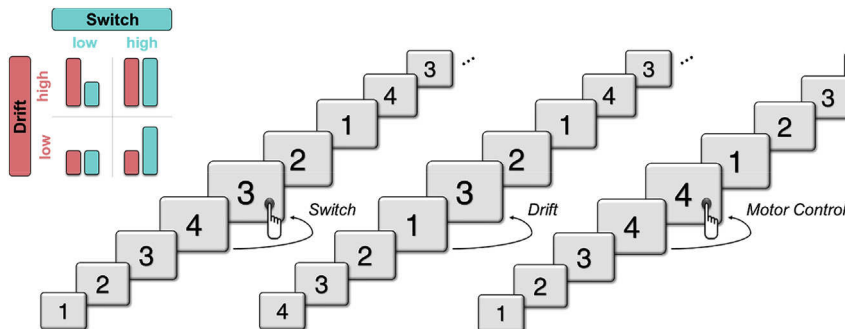
– 2–1) order (Fig. 1). To enable participants to predict forthcoming input, the sequence was repeated constantly, and digits succeeded one another for 1 s, separated by an inter-stimulus interval of 100 ms. Directional changes from ascending to descending digit sequences or vice versa (switches, hereafter) occurred at pseudorandom ordinal positions within the initial sequence. Subjects were asked to signal these events via button press (switch detection). Besides, single digits were omitted occasionally at variable positions without a temporal gap (drifts, hereafter), and participants were instructed to ignore these omissions (drift rejection). Switches and drifts never appeared at the same time, i.e., their occurrence was always unambiguous. During a motor control task, which was implemented to assess the individual mean reaction time, one digit of the sequence repeated continuously but maximally eight times until the participant pressed the response button. A total of 25 of these motor control trials were randomly interleaved across the experiment. A 6 s presentation of a fixation cross distributed across the experiment in 1.33% of the trials (n = 20) served as the baseline (i.e., rest trial).

To assess adaptation to different environments, the task was binned into 12 blocks that either had a high or low probability of switches, either paired with a high or low probability of drifts. Each block consisted of an average number of 125 trials in a full-factorial 2 (probability: high vs. low) x 2 (event: switch vs. drift) design. Transitions between block types resulting from this factor combination were balanced across the session. Probabilities were based on a pilot study, which assessed the performance of 12 PD patients during a staircase procedure of the task with different switch and drift frequencies. As a result, the maximum event frequency in unmixed blocks, in which switches and drifts occurred with the same frequency, was set to 16% (i.e., 8% per event type) and minimum event frequency was set to 8% (i.e., 4% per event type). In mixed (i.e., high-switch and low-drift or vice versa) blocks, the maximum frequency was set to 12%, whereas the minimum frequency was left at 4%. In this way, the difficulty level regarding the overall probability of events was kept constant across the experiment (except for unmixed low-frequency blocks). Stimulus presentation was pseudorandomised using the stochastic universal sampling method (Baker, 1987), which ensured a balanced distribution of switches and drifts across the blocks. Mean separation of the events was 6.24 (SD = 5.40).

The training consisted of ten blocks of 80 trials each and a probability of 16% for switch or drift occurrence. To enable participants to get accustomed to the task, presentation speed started at 1400 ms per digit and adapted block-wise with a decrease of 50 ms provided that the participant correctly reacted to 75% of the events. Besides, patients performed a short training before the scanner session with three blocks of 80 trials at the main experiment's digit presentation speed of 1 s. The randomisation was programmed using MATLAB R2012b (The MathWorks Inc., Natick, MA, USA) and stimuli were presented using Presentation 13.1 (Neurobehavioral Systems, San Francisco, CA, USA).

2.3. Probability model

In a Bayesian cognitive model, an observer's predictions of forthcoming sensory input are represented as probability distributions based on previous sensory input and prior knowledge. Due to the present task design, in which critical events were per se rather improbable and, thus, unpredicted as opposed to frequent standard digits, it is not plausible to assume an ideal Bayesian observer, which would base estimation of probabilities of these events on *all* previous events. Instead, we supposed that the observer would not be able to remember distant events. Consequently, we made use of a time-dependent decay model derived from Harrison et al. (2011) in which distant events are weighted less than recent ones. According to their formula, the expectations of an observer's model to observe different types of events at trial N are based on the weighted counts α of each type of event k (switch, drift, or standard digit) in the preceding trials $T = \{1, 2, ..., N - 1\}$, which are exponentially weighted by half-life τ:

$$\alpha_k(N) = \sum_{t \in T} \exp\left(-\frac{N - t - 1}{\tau}\right)\delta(x_t = k)$$

In this formula, $\delta(x_t = k)$ is equal to 1 if the event at time t corresponds to event type k and 0 otherwise. Further, it is assumed that trial N-1 has the weight 1 (i.e., is fully weighted), while the weights of more distant trials decrease at an exponential rate. Rest trials were neglected during the counting of the events. The weighted counts were computed based on the assumption that $\tau = 125$, according to the mean block length.

The probability of a particular event k occurring at trial N can then be calculated as (Bernardo and Smith, 2009):

$$\rho_k(N) = \frac{\alpha_k(N) + \alpha_k(0)}{\sum_{j=1}^{K}\alpha_j(N) + \alpha_j(0)}$$

In words, this probability is characterized by the weighted counts of event k relative to the sum of the counts of all possible events (switch, drift, and standard digit). The prior counts $\alpha_k(0)$ before observing the first trial were set to 1/3 for all events representing an uninformative prior (Jeffreys, 1946). Moreover, like previous studies, we used information theoretic indices, i.e., *surprise* and *entropy*, to quantify the amount of information provided by the current stimulus that could predict response accuracy and neural responses (e.g., Strange et al., 2005; Bestmann et al., 2008; Mars et al., 2008). The surprise $I_k(N)$ of an event, i.e., its improbability, is given by the negative logarithm of the probability:

$$I_k(N) = -\log p_k(N)$$

Conversely, entropy measures the average surprise of all possible events and quantifies the expected information of events regarding their predictability:

$$H(N) = -\sum_{k=1}^{K} \rho_k(N)\log\rho_k(N)$$

The varying extent to which each stimulus was locally unexpected,

Fig. 1. Schematic diagram of the task. Stimuli of a simple 4-digit sequence continuously followed each other with a duration of 1 s and an inter-stimulus interval of 100 ms. Subjects had to indicate changes from ascending to descending sequences (and vice versa) (switch), as displayed in the left row, via a button press. Moreover, they had to ignore the omission of a single digit (drift), as displayed in the middle row. During a motor control task, depicted on the right, one digit repeated continuously until the participant pressed the response button. As depicted in the top left diagram, the probabilities of switches and drifts varied block-wise across the experiment in a 2x2 design.
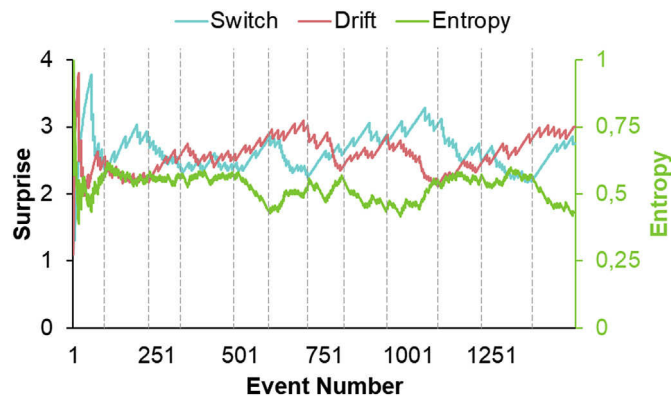
Fig. 2. Illustration of the decay-dependent information-theoretic measures *surprise,* for switches and drifts, and *entropy* varying throughout the experiment of one example participant. Surprise and entropy depended on half-live according to the mean block length of 125 trials (based on a formula derived from Harrison et al. (2011)). Surprise values were used to predict the participant's performance and BOLD activity at switches and drifts. Dashed vertical lines reflect boundaries between the different blocks.

i.e., its respective surprise value (Fig. 2), was used to explain error rates and fMRI BOLD response amplitudes at switches and drifts. In doing so, we aimed to assess whether learning differs between the two event types, and between healthy control participants and PD patients, regarding behavioural adaptation and its corresponding brain activity.

Finally, we also explored the predictive ability of models based on counts $\alpha_k(N)$ weighted by other half-lives (see supplementary material at https://osf.io/n5ugp/). We used approximate leave-one-out cross-validation (Vehtari et al., 2017) to compare the predictive ability of the different models. As detailed in the supplementary material, models based on a shorter τ were generally better at predicting response accuracy of both healthy controls and patients by switch and drift surprise; however, results based on a more extended specification of $\tau \geq 125$ were more sensitive to differences between controls and PD patients. Data and code of these analyses are available in OSF at https://osf.io/n5ugp/.

2.4. fMRI data acquisition

Whole-brain imaging data were collected on a 3 T S Magnetom Prisma MR tomograph using a TRTX-head coil. To minimise head motion, the head was tightly fixated with cushions. Functional images were acquired using a gradient T2*-weighted single-shot echo-planar imaging (EPI) sequence sensitive to blood oxygenation level dependent (BOLD) contrast (64 x 64 data acquisition matrix, 192 mm field of view, 90° flip angle, TR = 2000 ms, TE = 30 ms). Each volume consisted of thirty adjacent axial slices with a slice thickness of 4 mm and a gap of 1 mm. Images were acquired in ascending order along the AC-PC plane to provide a whole-brain coverage. Structural data were acquired for each participant using a standard Siemens 3D T1-weighted MPRAGE sequence for a detailed reconstruction of anatomy with isotropic voxels (1 x 1 × 1 mm) in a 256 mm field of view (256 x 256 matrix, 192 slices, TR = 2130, TE = 2.28). Stimuli were projected on a screen positioned behind the subject's head and were presented in the centre of the field of vision by a video-projector. Subjects viewed the screen by a 45° mirror, which was fixated on the top of the head coil and adjusted for each subject to provide a good view of the entire screen.

2.5. Behavioural data analysis

We assessed task performance by accurate detection of switches (hits), and correct non-responses to drifts (correct rejections), or, correspondingly, switch misses and false alarms at drifts. The motor control task was used to determine the 90%-quantile of each participant's

reaction times. This quantile served as an individual time window, in which button presses in response to switches and drifts were acknowledged as hits and false alarms, respectively. Using Bayesian logistic multilevel models in R (R Core Team, 2018) via the *brms* package using Stan (Bürkner, 2017; Carpenter et al., 2017), dichotomous erroneous responses, i.e., switch misses and false alarms at drifts, were predicted by decay-dependent information-theoretic indices (i.e., switch surprise and drift surprise as well as entropy dependent on half-live τ) in interaction with event type (i.e., switch and drift) and group. As regards the latter, two separate models were estimated according to our hypotheses, with differences between controls and PD patients OFF interpreted as effects of the disease (H1) and differences between PD patients ON and OFF interpreted as effects of dopamine medication (H4). No comparisons between controls and patients ON were carried out to avoid a confounding of disease and medication effects that might, for instance, relate to medication side effects. Finally, session was added as a factor to make sure that differences between controls and PD patients were not driven by retest-effects in PD patients on their second visit. For a summary of model parameters, including interaction terms, we report regression coefficients and 95% credible intervals (CIs; i.e., Bayesian confidence intervals). This means that there is a 95% probability that the respective parameter falls within this interval, given the evidence provided by the data (note that it would indicate statistical significance on a 5% level if the interval does not contain zero). For the factors group, session and event type, we used effect coding with −1 for healthy controls and 1 for patients OFF, −1 for patients ON and 1 for patients OFF, −1 for the first and 1 for the second session, and −1 for drifts and 1 for switches. Weak or non-informative default priors of the *brms* package were used (Bürkner, 2017).

2.6. fMRI data preprocessing

Brain image preprocessing and basic statistical analyses were conducted using SPM12 (Wellcome Department of Imaging Neuroscience, London, UK; see: http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). Functional images were slice-timed to the middle slice to correct for differences in slice acquisition time. To correct for three-dimensional motion, individual functional MR (EPI) images were realigned to the mean image. Motion correction estimates were inspected visually as those participants who exceeded a maximum of 3 mm head movements between two scans in the x, y, and z dimensions would have been excluded from further analyses. The anatomical scan was co-registered (rigid body transformation) to the mean functional image. Each subject's co-registered anatomical scan was segmented into native space tissue components. A group-specific template was created using DARTEL with default settings in SPM12. Functional images were then normalised to the MNI space by affine transformations using invertible and smooth deformations (flow fields) for each participant's native space to the template derived from the previous step through the DARTEL tool. Smoothing was also applied during DARTEL warping with a Gaussian kernel of 8 mm^3 full width at half-maximum.

To reduce effects of physiological noise (e.g., due to potential increased disease-specific motion or pulsatile artefacts in the midbrain), we performed a denoising procedure on the EPI data using the default settings of the CONN toolbox in MATLAB (Whitfield-Gabrieli and Nieto-Castanon, 2012), which implements the anatomical component-based noise correction method (aCompCor). Denoising included regressing out the first five principal components associated with white matter and cerebrospinal fluid as well as the motion parameters and their temporal derivatives from the BOLD signal. Finally, a 128 s temporal high-pass filter was applied to the data to remove low-frequency noise.

2.7. fMRI design specification

The statistical analysis was based on a least-squares estimation using

the general linear model (GLM) for serially autocorrelated observations (Friston et al., 1995; Worsley and Friston, 1995). The GLM included four regressors coding for onsets and durations of the specific event types, i.e., standard digits (std), switches, drifts, and motor control trials, which were then convolved with the canonical hemodynamic response function (HRF) and regressed against the observed fMRI data. Moreover, to model variability in the BOLD amplitude as a function of decay-dependent surprise and entropy, as outlined in detail above (cf. Probability model), two parametric modulators each were added to the switch and the drift regressor, i.e., one for decay-dependent surprise and one for decay-dependent entropy. We mean-centred each of these modulators before entering the GLM. Whenever two trials were separated by less than 2 s (i.e., less than one TR), only the first one was included in the GLM, whereas the second was not modelled and treated as part of the implicit baseline. Likewise, resting periods were not modelled and served as an implicit baseline (Pernet, 2014). The subject-specific six rigid-body transformations obtained from residual motion correction were included as covariates of no interest.

To ensure that neural activity during switches and drifts generally exceeded activation during standard digits, we conjoined one-sample *t*-tests of the contrasts switch > std and drift > std of healthy controls and PD patients OFF. To this end, the corresponding beta images per participant were submitted to a second-level two-way analysis of variance to perform a conjunction analysis testing against the conjunction null hypothesis (p(peak-level FWE) < 0.05, Nichols et al., 2005). Moreover, individual statistical maps for variations of BOLD amplitudes with surprise at switches and drifts were generated for each participant. We performed region of interest (ROI) analyses to test for BOLD activation differences between the groups in the substantia nigra and the caudate nucleus. The substantia nigra ROIs were derived from the probabilistic atlas of the basal ganglia (ATAG; Keuken et al., 2014). Anatomically defined ROIs of the left and right caudate nucleus were derived from the automated anatomical labelling (AAL) atlas and created using the SPM Wake Forest University (WFU) Pickatlas toolbox (http://www.fmri.wfubmc.edu/cms/software, version 2.3) (Maldjian et al., 2003). For the ROI analyses, we extracted the beta scores of switch and drift surprise and corresponding standard errors per voxel and subject. Using Bayesian linear multilevel models in R (R Core Team, 2018) via *brms* and Stan with default priors (Bürkner, 2017; Carpenter et al., 2017), we predicted these beta scores by event type, group, session, and ROI (left and right caudate nucleus and substantia nigra) while accounting for dependencies between responses belonging to the same voxel or subject. Furthermore, varying degrees of uncertainty in the beta scores (i.e., varying standard errors) were accounted for by using an inverse-variance weighting scheme with more precisely estimated scores receiving higher weight (Cooper et al., 2009; https://osf.io/n5ugp/). We tested hypothesised differences in the beta values in the selected ROIs between healthy controls and PD patients OFF, and between PD patients OFF and ON. The α-level was set to 5%, with a 90% CI for directional hypotheses. Further, in addition to regression coefficients and CIs, we report the posterior probability (*pp*) that the differences in neural activation between the groups show in the expected directions.

3. Results

3.1. Behavioural results

Bayesian logistic multilevel models on surprise and entropy estimates that depended on decay half-life τ (cf. Probability model) were used to predict behavioural errors, that is, switch misses and false alarms at drifts of healthy controls versus PD patients off medication, and PD patients on versus off medication. We found that the modulation of error rate by decay-dependent surprise and entropy depended on event type and group. Although there was no main effect of switch and drift surprise across the groups, the results are in accordance with H1 by revealing an interaction effect of SWITCH SURPRISE X GROUP X EVENT TYPE, as well as an

interaction effect of DRIFT SURPRISE X GROUP X EVENT TYPE: In controls, higher probability, i.e., lower surprise of switches led to more switch hits, whereas in PD patients OFF the rate of switch hits increased as a function of switch surprise. In contrast, while drift surprise did not modulate the error rate of healthy controls at drifts, the false alarm rate of PD patients OFF increased as a function of drift surprise (Fig. 3). Moreover, we observed that increasing entropy affected the error rate at switches more than at drifts in healthy controls only. Independently of their medication, PD patients revealed an interaction effect of DRIFT SURPRISE X EVENT TYPE, i.e., increasing drift surprise was accompanied by a higher rate of false alarms at drifts but also a lower rate of switch misses in both PD patients on and off medication. However, there were no differences between the medication states. Regression coefficients and corresponding 95% CIs for each predictor variable of the two models are given in Table 1 for controls and PD patients off medication and in Table 2 for PD patients on and off medication.

3.2. fMRI results

Main effects of switches and drifts. To first identify the network associated with prediction error processing in general, we run a conjunction analysis on the statistical maps of the two groups and the two tasks, i.e., of healthy controls and PD patients OFF during switch and drift processing [controls (switch > std) ∩ controls (drift > std) ∩ patients OFF (switch > std) ∩ patients OFF (drift > std)]. This analysis revealed higher activations in a network comprising - amongst others - the inferior parietal cortex as well as the inferior frontal gyrus extending into the anterior insula during switches and drifts compared to standard digits (Table 3) (Fig. 4).

Parametric effects of decay-dependent surprise. Bayesian linear multilevel models were employed to test our hypotheses (H2–H4) that decay-dependent surprise at switches and drifts modulated the BOLD response in defined ROIs, i.e., the substantia nigra and the caudate nucleus, differently in controls and PD patients OFF and ON. Thus, corresponding to the reported error rate effects, we addressed the hypothesised differences between healthy controls and PD patients OFF, and between PD patients ON and OFF.

In line with H2, the right substantia nigra activity showed a positive correlation with switch surprise in controls but not in patients off medication, R: $b_{HC\text{-}OFF} = 55.1$, 90%-CI = [3.49, 106.07], $pp(b > 0) = 0.96$; L: $b_{HC\text{-}OFF} = 1.05$, 90%-CI = [-54.15, 57.26], $pp(b > 0) = 0.51$. Regarding a modulation by drift surprise, parameter estimates were close to zero in both groups, that is, no differences between healthy controls and patients OFF were observed, R: $b_{HC\text{-}OFF} = 20.68$, 90%-CI = [-21.65, 63.16], $pp(b > 0) = 0.79$; L: $b_{HC\text{-}OFF} = 5.80$, 90%-CI = [-28.86, 40.80], $pp(b > 0) = 0.61$ (Fig. 5, top-left panel).

In support of H3, caudate nucleus activity covaried positively with switch probability in controls but not in PD patients OFF, L (left): $b_{HC\text{-}OFF} = -56.85$, 90%-CI = [-123.21, 9.21], $pp(b < 0) = 0.92$; R (right): $b_{HC\text{-}OFF} = -43.00$, 90%-CI = [-96.45, 11.55], $pp(b < 0) = 0.90$. In contrast, we found a positive covariation of drift surprise with caudate activity and there were no clear differences between the groups, L: $b_{HC\text{-}OFF} = 24.74$, 90%-CI = [-29.85, 81.19]; R: $b_{HC\text{-}OFF} = 11.41$, 90%-CI = [-45.56, 69.44] (Fig. 5, top-right panel).

Contrary to hypothesis H4, we did not find differences between PD patients ON and OFF in the substantia nigra for switch surprise, R: $b_{ON\text{-}OFF} = 21.4$, 90%-CI = [-23.39, 69.65], $pp(b < 0) = 0.78$; L: $b_{ON\text{-}OFF} = 0.70$, 90%-CI = [-46.58, 49.47], $pp(b < 0) = 0.51$. Likewise, modulation by drift surprise did not differ between the groups, R: $b_{ON\text{-}OFF} = 4.68$, 90%-CI = [-37.91, 47.06], $pp(b < 0) = 0.57$; L: $b_{ON\text{-}OFF} = 2.35$, 90%-CI = [-36.03, 40.07], $pp(b < 0) = 0.54$ (Fig. 5, bottom-left panel).

Regarding the caudate nucleus, medication did not restore the neural activation within this region. Instead, right caudate nucleus covaried positively with switch surprise in patients on compared to patients off medication, L: $b_{ON\text{-}OFF} = 16.44$, 90%-CI = [-42.49, 75.72], $pp(b < 0) = 0.32$; R: $b_{ON\text{-}OFF} = 47.21$, 90%-CI = [0.97, 93.29], $pp(b < 0) = 0.05$. No
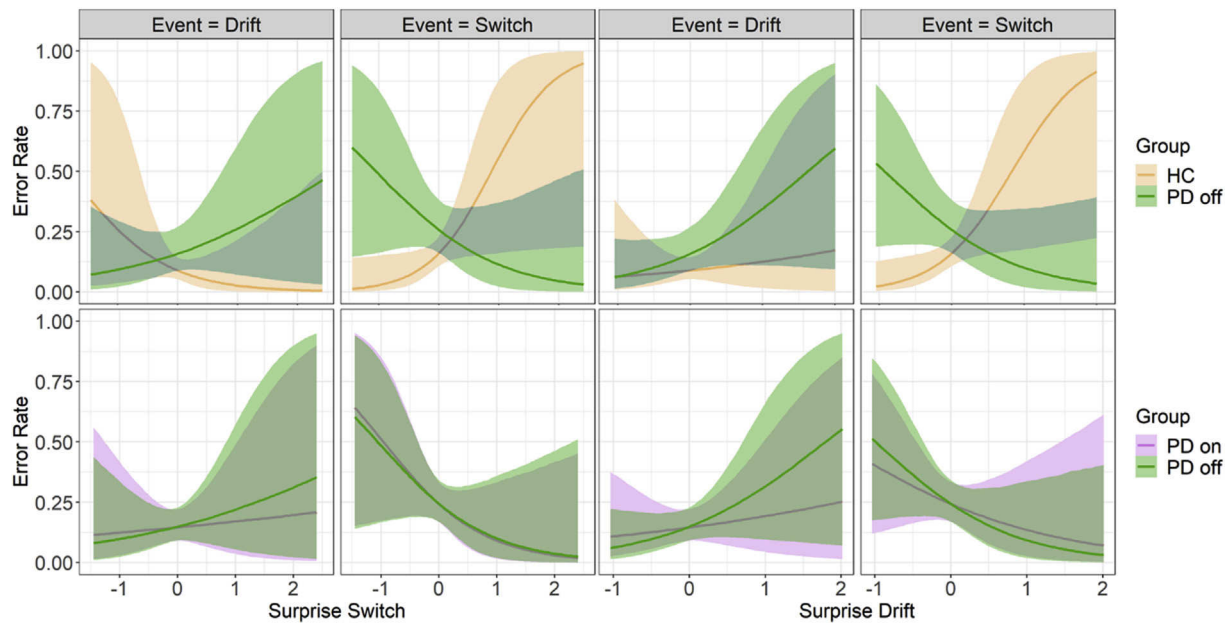
Fig. 3. Behavioural data. Effect of mean-centred decay-dependent drift surprise (left panel) and switch surprise (right panel) on error rate at drifts, i.e., false alarms, and at switches, i.e., misses, in healthy controls (HC) and PD patients off medication (top panel) and in PD patients off and on medication (lower panel). The solid lines depict the regression fit, and the shaded areas show the 95% credibility intervals.

Table 1
Bayesian logistic multilevel model predicting errors (i.e., switch misses and drift false alarms) of healthy controls and PD patients OFF.

Coefficient	b	l-95% CI	u-95% CI
Switch surprise	0.04	−0.85	0.93
Drift surprise	0.55	−0.29	1.38
Entropy	3.18	−1.99	8.28
Group: PD patients OFF	2.06	−0.53	4.61
Event type: Switch	−1.35	−3.86	1.20
Session: Second	−0.05	−0.43	0.31
Switch surprise x Group: PD patients OFF	−0.24	−1.10	0.66
Drift surprise x Group: PD patients OFF	−0.62	−1.41	0.20
Entropy x Group: PD patients OFF	−3.47	−8.42	1.63
Switch surprise x Event type: Switch	0.39	−0.47	1.23
Drift surprise x Event type: Switch	−0.13	−0.94	0.67
Entropy x Event type: Switch	3.32	−1.69	8.28
Group: PD patients OFF x Event type: Switch	3.09	0.56	5.62
Switch surprise x Group: PD patients OFF x Event type: Switch	−1.21	−2.06	−0.34
Drift surprise x Group: PD patients OFF x Event type: Switch	−0.96	−1.78	−0.15
Entropy x Group: PD patients OFF x Event type: Switch	−6.16	−11.16	−1.15

Table 2
Bayesian logistic multilevel model predicting errors (i.e., switch misses and drift false alarms) of PD patients ON and OFF.

Coefficient	b	l-95% CI	u-95% CI
Switch surprise	−0.41	−1.23	0.41
Drift surprise	−0.16	−0.92	0.58
Entropy	−1.30	−5.87	3.23
Group: PD patients ON	0.33	−1.94	2.68
Event type: Switch	1.16	−1.09	3.41
Session: Second	0.00	0.08	−0.08
Switch surprise x Group: PD patients ON	−0.10	−0.87	0.66
Drift surprise x Group: PD patients ON	−0.06	−0.83	0.69
Entropy x Group: PD patients ON	−0.67	−5.34	3.83
Switch surprise x Event type: Switch	−0.73	−1.49	0.05
Drift surprise x Event type: Switch	−0.79	−1.50	−0.08
Entropy x Event type: Switch	−1.67	−6.15	2.80
Group: PD patients ON x Event type: Switch	−0.53	−2.71	1.73
Switch surprise x Group: PD patients ON x Event type: Switch	0.05	−0.72	0.79
Drift surprise x Group: PD patients ON x Event type: Switch	0.26	−0.46	0.97
Entropy x Group: PD patients ON x Event type: Switch	1.06	−3.43	5.40

differences in the modulation by drift surprise were observed, L: b_{ON-OFF} = 22.1, 90%-CI = [-36.75, 82.40]; R: b_{ON-OFF} = 8.69, 90%-CI = [-33.37, 50.22] (Fig. 5, bottom-right panel).

4. Discussion

The present fMRI study aimed at gaining insight into the potential involvement of dopamine in learning from prediction errors requiring either flexible updating or stabilisation of current predictions. We addressed this issue by comparing the effects of prediction error probability based on a specific decay half-live on performance and activity in key dopaminergic regions of healthy control participants and PD patients on and off dopaminergic medication during a digit prediction task. We found that in healthy participants but not in PD patients increasing decay-dependent probability, that is, lower surprise of flexibility-demanding violations of a predictable digit sequence (i.e., *switches*) was

accompanied by a better switch detection (H1). Contrary to our hypothesis (H4), there were no performance differences between PD patients on and off medication. On the neural level, we observed a deficient modulation of decay-dependent surprise of these violations by the right substantia nigra (H2) and the caudate nucleus (H3) in PD patients off medication, but no changes of surprise responses under medication (H4). In the following, we will discuss our behavioural findings on impaired probability-dependent responding in PD patients, and will then go into more detail regarding our fMRI results.

Error rates of healthy controls and PD patients at switches were differentially affected by changes in the time-dependent probability of these event occurrences. Healthy participants detected more switches when switch probability within a certain time frame became particularly high. In contrast, PD patients, no matter whether on or off medication, reacted less flexibly to switches when these became more probable and thus did not adapt to high probability conditions. This finding indicates

Table 3

Maxima of activation from the conjunction analysis of the contrast images of switch > std and drift > std of healthy controls and PD patients off medication at $p < 0.05$ peak-level FWE-corrected. Labels are reported according to the AAL atlas. Entries in italics indicate sub-peak regions that are more than 8 mm apart within a cluster. MNI, Montreal Neurological Institute.

Region Label	Cluster Extent	t-value	MNI Coordinates		
			x	y	z
Right Inferior Parietal Lobule	785	13.02	45	−39	45
Superior Occipital Gyrus		9.29	33	−66	42
Precuneus		7.15	9	−66	51
Left Inferior Parietal Lobule	645	12.37	−39	−45	42
Superior Parietal Gyrus		9.84	−27	−66	45
Right Inferior Frontal Gyrus (pars opercularis)	852	11.38	48	9	21
Superior Frontal Gyrus		9.20	27	6	57
Left Precentral Gyrus	1425	11.01	−45	6	30
Supplementary Motor Area		10.98	−6	9	51
Right Insula	123	9.68	36	24	−3
Left Cerebellum	131	8.99	−6	−78	−27
Right Cerebellum	70	8.10	30	−63	−27
Left Precuneus	33	6.95	−6	−69	−48
Right Middle Temporal Gyrus	23	6.54	57	−51	−9
Right Lingual Gyrus	7	6.15	9	−24	−12
Right Thalamus	9	6.07	12	−9	0
Left Inferior Temporal Gyrus	1	5.48	−51	−57	−9
Left Inferior Frontal Gyrus (pars triangularis)	2	5.38	−54	18	−3
Left Thalamus	2	5.28	−15	−9	3

that PD patients are impaired in anticipating flexibility-demanding environmental changes due to difficulties in probabilistic learning. Thereby, this result supports recent studies reporting deficits in implicit contextual learning in PD patients (Perugini et al., 2016; Perugini and Basso, 2017) and extends previous results implicating dopamine in flexible responses to sensory prediction errors by suggesting a corresponding impairment in PD (Galea et al., 2012; Iglesias et al., 2013). However, since medication did not restore this deficit, further transmitter systems may be involved in learning to adjust one's behaviour to changing contextual demands (see below).

We found some evidence that PD patients compared with healthy controls are more susceptible to respond to rare, that is, highly surprising stability-demanding prediction errors requiring motor inhibition. Notably, previous research showed enhanced instead of lower distractor-resistance in PD patients (Cools et al., 2009). However, this alleged advantage in PD patients possibly results from inflexibility rather than active stabilisation (Uitvlugt et al., 2016). In the present study, the patients' inflexibility (and alleged stability) in reacting to sequential violations in general had a particular impact upon high-probability conditions, whereas PD patients became more prone to deliver responses when stimuli became more surprising – no matter whether these stimuli required a motor response or not. By that, the patients' somewhat arbitrary responses to rare unpredicted events reflected a deficient response selection (Humphries et al., 2006). Together, our behavioral findings thus indicate that PD patients have deficits in distinguishing between and learning from different types of unexpected events requiring either stabilisation or updating of prediction.

On the neural level, the right substantia nigra activity showed a positive correlation with switch surprise in controls but not in patients off medication. This reflects, in turn, a decrease of neural activity in the course of more frequent event occurrence, which has been regarded as a sign of learning (Turk-Browne et al., 2010; Schiffer et al., 2012). Paralleling our behavioural findings, learning from flexibility-demanding prediction errors in high-probability conditions thus appears to be accompanied by a relative activation decrease within the substantia nigra in healthy subjects, whereas there was no modulation of activity within this region in PD patients. Although this finding is in accordance with a broad literature reporting prediction error coding and learning in the
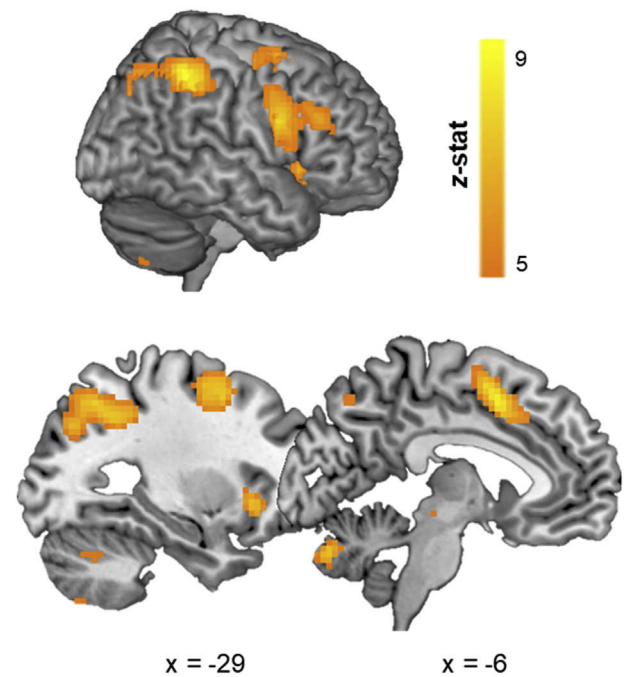
Fig. 4. fMRI activation at $p < 0.05$, peak-level FWE-corrected threshold for the conjunction analysis identifying the brain regions that were more active during switches and drifts relative to standard digits in both healthy controls and PD patients off medication.

midbrain dopaminergic nuclei (e.g. Schultz and Dickinson, 2000; Redgrave and Gurney, 2006; D'Ardenne et al., 2012), recent fMRI and PET studies report that substantia nigra specifically encode belief updates (i.e., Bayesian surprise) but not sensory (information-theoretic) surprise, that is, pure unexpectedness (Nour et al., 2018; Schwartenbeck et al., 2016). Accordingly, our finding can be explained by the behavioural gain that results from an adaptation to local environmental challenges by learning dynamically changing probabilities of unexpected events over a specific time scale. Moreover, contrary to our hypothesis, substantia nigra activity was not modulated by drift surprise in either group, suggesting that decay-dependent surprise signals within this region are specific to flexibility-demanding violations and thus indeed process information content rather than unexpectedness per se.

In healthy controls but not in PD patients off medication, BOLD amplitudes in the caudate nucleus increased with a higher probability of switch occurrence. The involvement of the caudate nucleus in predictive processing of flexibility-demanding events extends findings from previous fMRI studies and computational models that highlight striatal signalling in delivering gating input to frontal areas to allow flexible updating of cortical representations, possibly modulated by dopamine (e.g., O'Reilly and Frank, 2006; Stelzel et al., 2013). Notably, our results reveal evidence for a caudate activation increase as a function of drift surprise in healthy participants (instead of a decrease as was the case for switch surprise). This could indicate that environments in which stability-demanding events become more probable rather lead to a decrease in striatal signalling but, concurrently, increased striatal firing rates when events are highly unexpected. Although this is speculative, the direction of the reported correlations could be accounted for by the relationship between phasic and tonic dopamine release (Grace, 1991). It is suggested that dopaminergic neurons do not only respond to unpredicted events per se but also encode their precision by tonic dopamine release over longer time scales (Fiorillo et al., 2008; Friston et al., 2012). Therefore, the reciprocal relationship between anticipation and surprise with respective increased neural responses to predicted and unpredicted events might reflect tonic and phasic dopamine signals, respectively (Schmitz et al., 2003; O'Reilly and Frank, 2006; Yu et al., 2013).
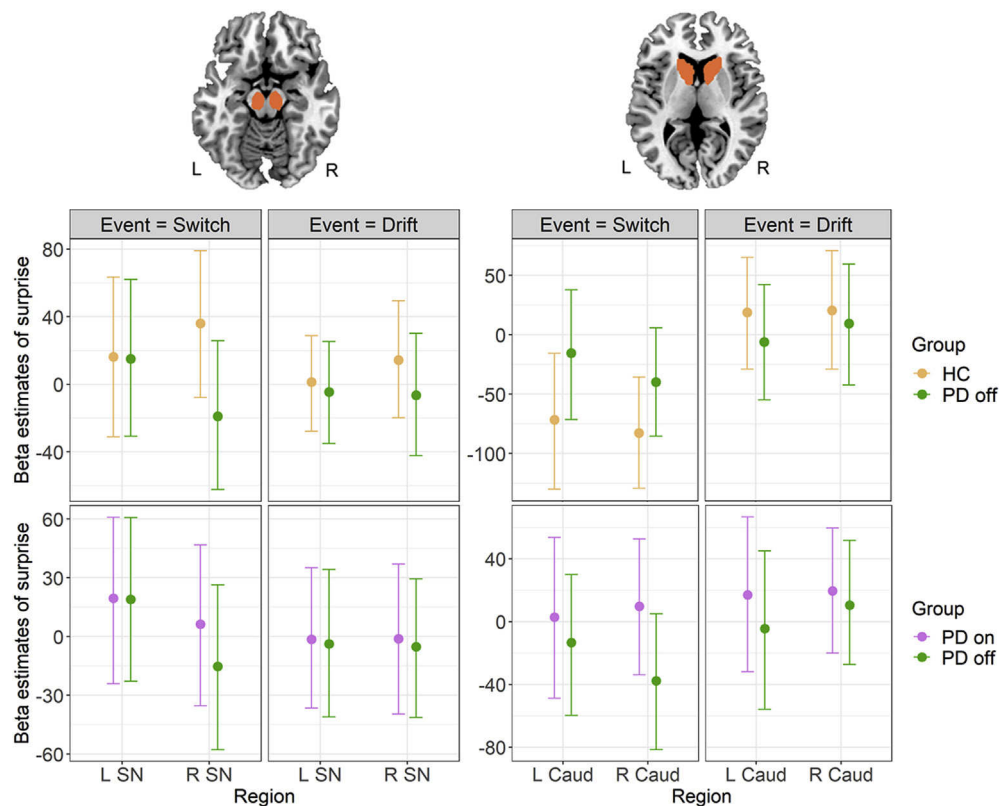
Fig. 5. Region of interest fMRI data. Beta estimates of the continuous modulation of activity in the left (L) and right (R) substantia nigra (SN), and caudate nucleus (Caud), by decay-dependent surprise at switches and drifts in healthy controls (HC) and PD patients off medication (top panel) and in PD patient off and on medication (lower panel).

Crucially, the dopaminergic medication did not restore learning from prediction errors, suggesting that not dopamine supply alone can explain the reported effects. Previous studies reported on heterogenous findings regarding the impact of dopaminergic therapy on learning. Studies demonstrate effects of medication on locomotor (Roemmich et al., 2014) and sensorimotor (Wolpe et al., 2018) adaptive learning as well as on reinforcement learning, but with mixed results on whether medication has beneficial and/or detrimental effects (e.g., Frank et al., 2004; Bódi et al., 2009; Argyelan et al., 2018; McCoy et al., 2019; see Meder et al., 2019, for a recent review). Similarly, with regard to uncertainty learning, i.e., when decisions are based on the integration of prior and current sensory information, some studies report an improvement due to dopaminergic medication (Wolpe et al., 2015; Vilares and Kording, 2017; Tomassini et al., 2019), while others did not find differences between PD patients on and off medication (Perugini and Basso, 2017; see Perugini et al., 2018).

It has been suggested that dopamine deficiency reduces the modulation of performance rather than learning itself (e.g., Beeler et al., 2010; Smittenaar et al., 2012). Accordingly, a recent fMRI study found that activation of the dorsal striatum associated with the decision event in a stimulus-response learning task increased on dopaminergic medication, whereas signals of the ventral striatum related to learning during a feedback event was depressed by medication (Hiebert et al., 2019). Thus, the particular role of dopamine in learning from prediction errors might consist in the modulation of correct response selection by exploiting already learned uncertainty representations (Beeler et al., 2010; Marshall et al., 2016). In contrast, signals of environmental uncertainty might rather be encoded by other neurotransmitters such as noradrenaline or acetylcholine (Yu and Dayan, 2002, 2005; Marshall et al., 2016). For example, the hippocampus has been associated with contextual learning by extracting statistical information to create a representation of the environmental volatility (Schapiro et al., 2014; Kluger and Schubotz,

2017). Hippocampal dysfunction in PD has been associated with cholinergic loss resulting in deficits in learning and memory, i.e., a progression towards dementia (Hall et al., 2014). Moreover, previous work highlights the role of noradrenaline depletion due to neuron loss within the locus coeruleus in PD, accompanied by cognitive inflexibility (Delaville et al., 2011; Vazey and Aston-Jones, 2012) and impaired inhibitory control (Borchert et al., 2016; Rae et al., 2016). Using fMRI, Ye et al. (2015) could show that PD patients' improvement in inhibitory control on atomoxetine, a noradrenaline reuptake inhibitor, was associated with increased functional and structural frontostriatal connectivity. Moreover, the authors could show that these beneficial effects on task performance could also be predicted by the patients' levodopa equivalent daily dose (Ye et al., 2016).

These findings suggest that dopamine release interacts with other neurotransmitter systems and possibly is affected by and influences uncertainty representations by means of the sensory inputs' goal relevance (Picciotto et al., 2012; Mizumori and Tryon, 2015; Aly and Turk-Browne, 2018). In the present study, we therefore assume that dopaminergic drugs do not enhance learning signals because these are rather provided by other neuromodulators. As a result, predictive strategies cannot be adapted to increasing demands on flexibility, resulting in a suboptimal performance. However, to exhibit the contribution of dopamine to impaired learning in PD, future studies with direct neural recordings should measure dopaminergic neuron responses to prediction errors of varying precision, for example in patients undergoing deep brain stimulation surgery.

Although not part of the hypotheses, it should be noted that increased time-dependent *drift* probability also improved flexible responding to switches; that is, both increased half-life weighted switch and drift occurrence led to better detection of switches in healthy subjects. Thus, contextual learning seems to rely on teaching signals provided by *all* types of violations but ultimately only impacts upon the motor response

to flexibility-demanding events. This finding supports previous accounts, according to which dopamine sensitises behaviour to higher-level precision estimates by highlighting surprising input in predictable contexts (Bestmann et al., 2014; Marshall et al., 2016). Because each violation elicited changes in both tracked measures, i.e., in switch as well as drift surprise, future studies could also investigate the effect of switch and drift surprise (and their interaction) on BOLD response at drifts and switches, respectively. Accordingly, we acknowledge that both switches and drifts are relevant for updating the expectation of long-run future event occurrence and are thus not qualitatively different in every respect. However, although the same process likely achieves learning of drift and switch probabilities, our findings suggest that the effects of this learning on appropriate response selection to the two event types differ. Further research is warranted to disentangle the specific processes related to uncertainty learning on the one hand and response selection on the other hand.

Finally, it is of importance to acknowledge the dependency of the present results on the time scale (i.e., half-life τ) we used to fit the probability model. Our exploratory results suggest that different temporal scales differentially impact on behavioural and neural responses (see supplementary material at https://osf.io/n5ugp/) so that the present results on differences between the groups in integrating past events must be considered against the background of the half-life we selected according to our experimental manipulation. It would have exceeded the scope of the present study to estimate individual half-lives, that is, to investigate individual differences in time scales over which sensory information is actually accumulated, but we consider this to be an important question of future research. In line with that, it has been suggested that the hierarchical organisation of the cortex is determined by specific time scales over which information is aggregated (Kiebel et al., 2008; Harrison et al., 2011). Moreover, subjects differ concerning the number of samples they use when coding probabilities (Trempler et al., 2017) and the time they spend within one representational state (Vidaurre et al., 2017). Further studies could investigate whether PD patients also represent probabilities of upcoming flexibility- and stability-demanding sensory input but on a different (probably shorter) time scale, as our additional analyses suggest. Moreover, it is reasonable to assume that the accumulated evidence for an event to occur probably consists of an interplay of its absolute probability and the time elapsed since its last occurrence. Previous studies provided evidence for the role of dopamine in gathering information from the passage of time (Pasquereau and Turner, 2014; Tomassini et al., 2016; Tomassini et al., 2019). Thus, future studies should elaborate on the dopaminergic modulation of learning from either flexibility- or stability-demanding prediction errors by taking temporal dynamics of prediction formation into account.

In sum, our study provides evidence that altered decay-dependent surprise-driven learning signals in the substantia nigra and the caudate nucleus, though unaffected by dopaminergic therapy, contribute to a deficient adaptation of behaviour in response to flexibility-demanding surprising events in PD. These findings provide novel insight into the specificity of dopamine in exploiting learning and corresponding deficits in PD.

Declaration of competing interest

The authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Ima Trempler: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing - original draft. **Paul-Christian Bürkner:** Methodology, Formal analysis, Data curation, Writing - review & editing. **Nadiya El-Sourani:** Conceptualization, Investigation, Writing - review & editing. **Ellen Binder:** Investigation, Writing - review & editing. **Paul Reker:** Investigation, Writing - review & editing. **Gereon R. Fink:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing - review & editing. **Ricarda I. Schubotz:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing - review & editing.

Acknowledgments

References

Aly, M., Turk-Browne, N.B., 2018. Flexible weighting of diverse inputs makes hippocampal function malleable. Neurosci. Lett. 680, 13–22.

Argyelan, M., Herzallah, M., Sako, W., DeLucia, I., Sarpal, D., Vo, A., et al., 2018. Dopamine modulates striatal response to reward and punishment in patients with Parkinson's disease: a pharmacological challenge fMRI study. Neuroreport 29 (7), 532.

Badre, D., 2012. Opening the gate to working memory. Proc. Natl. Acad. Sci. U. S. A 109, 19878–19879.

Baker, J.E., 1987. Reducing bias and inefficiency in the selection algorithm. In: Proceedings of the Second International Conference on Genetic Algorithms and their Application, pp. 14–21.

Barter, J.W., Li, S., Lu, D., Bartholomew, R.A., Rossi, M.A., Shoemaker, C.T., et al., 2015. Beyond reward prediction errors: the role of dopamine in movement kinematics. Front. Integr. Neurosci. 9, 39.

Beeler, J.A., Daw, N., Frazier, C.R., Zhuang, X., 2010. Tonic dopamine modulates exploitation of reward learning. Front. Behav. Neurosci. 4, 170.

Beeler, J.A., Frank, M.J., McDaid, J., Alexander, E., Turkson, S., Bernandez, M.S., et al., 2012. A role for dopamine-mediated learning in the pathophysiology and treatment of Parkinson's disease. Cell Rep. 2 (6), 1747–1761.

Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. Nat. Neurosci. 10, 1214–1221.

Bernardo, J.M., Smith, A.F., 2009. Bayesian Theory, vol.405. John Wiley & Sons.

Bestmann, S., Harrison, L.M., Blankenburg, F., Mars, R.B., Haggard, P., Friston, K.J., Rothwell, J.C., 2008. Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. Curr. Biol. 18 (10), 775–780.

Bestmann, S., Ruge, D., Rothwell, J., Galea, J.M., 2014. The role of dopamine in motor flexibility. J. Cognit. Neurosci. 27 (2), 365–376.

Bódi, N., Kéri, S., Nagy, H., Moustafa, A., Myers, C.E., Daw, N., et al., 2009. Reward-learning and the novelty-seeking personality: a between-and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain 132 (9), 2385–2395.

Borchert, R.J., Rittman, T., Passamonti, L., Ye, Z., Sami, S., Jones, S.P., et al., 2016. Atomoxetine enhances connectivity of prefrontal networks in Parkinson's disease. Neuropsychopharmacology 41 (8), 2171.

Bürkner, P.C., 2017. brms: an R package for Bayesian multilevel models using Stan. J. Stat. Software 80 (1), 1–28.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., et al., 2017. Stan: a probabilistic programming language. J. Stat. Software 76 (1), 1–32.

Cools, R., Barker, R.A., Sahakian, B.J., Robbins, T.W., 2001. Mechanisms of cognitive set flexibility in Parkinson's disease. Brain 124 (12), 2503–2512.

Cools, R., Miyakawa, A., Sheridan, M., D'Esposito, M., 2009. Enhanced frontal function in Parkinson's disease. Brain 133 (1), 225–233.

Cools, R., D'Esposito, M., 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol. Psychiatr. 69, e113–e125.

Cooper, H., Hedges, L.V., Valentine, J.C. (Eds.), 2009. The Handbook of Research Synthesis and Meta-Analysis. Russell Sage Foundation.

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L.E., Cohen, J.D., 2012. Role of prefrontal cortex and the midbrain dopamine system in working memory updating. Proc. Natl. Acad. Sci. U. S. A 109, 19900–19909.

Delaville, C., De Deurwaerdère, P., Benazzouz, A., 2011. Noradrenaline and Parkinson's disease. Front. Syst. Neurosci. 5, 31.

Doll, B.B., Hutchison, K.E., Frank, M.J., 2011. Dopaminergic genes predict individual differences in susceptibility to confirmation bias. J. Neurosci. 31 (16), 6188–6198.

Durstewitz, D., Seamans, J.K., Sejnowski, T.J., 2000. Neurocomputational models of working memory. Nat. Neurosci. 3, 1184–1191.

Durstewitz, D., Seamans, J.K., 2008. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol. Psychiatr. 64, 739–749.

Fahn, S., Elton, R.L., 1987. Unified rating scale for Parkinson's disease. Recent Dev. Parkinson's Dis. 153–163.

Fiorillo, C.D., Newsome, W.T., Schultz, W., 2008. The temporal precision of reward prediction in dopamine neurons. Nat. Neurosci. 11 (8), 966.

Fiser, J., Berkes, P., Orbán, G., Lengyel, M., 2010. Statistically optimal perception and learning: from behavior to neural representations. Trends Cognit. Sci. 14 (3), 119–130.

Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306 (5703), 1940–1943.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Shiner, T., FitzGerald, T., Galea, J.M., Adams, R., Brown, H., et al., 2012. Dopamine, affordance and active inference. PLoS Comput. Biol. 8 (1), e1002327.

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., Dolan, R.J., 2014. The anatomy of choice: dopamine and decision-making. Phil. Trans. R. Soc. B 369 (1655), 20130481.

Galea, J.M., Bestmann, S., Beigi, M., Jahanshahi, M., Rothwell, J.C., 2012. Action reprogramming in Parkinson's disease: response to prediction error is modulated by levels of dopamine. J. Neurosci. 32 (2), 542–550.

Grace, A.A., 1991. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41 (1), 1–24.

Hall, H., Reyes, S., Landeck, N., Bye, C., Leanza, G., Double, K., et al., 2014. Hippocampal Lewy pathology and cholinergic dysfunction are associated with dementia in Parkinson's disease. Brain 137 (9), 2493–2508.

Harrison, L., Bestmann, S., Rosa, M.J., Penny, W., Green, G.G., 2011. Time scales of representation in the human brain: weighing past information to predict future events. Front. Hum. Neurosci. 5, 37.

Hautzinger, M., Keller, F., Kühner, C., 2006. Beck Depressions-Inventar (BDI-II). Harcourt Test Services, Frankfurt.

Hiebert, N.M., Owen, A.M., Ganjavi, H., Mendonça, D., Jenkins, M.E., Seergobin, K.N., MacDonald, P.A., 2019. Dorsal striatum does not mediate feedback-based, stimulus-response learning: an event-related fMRI study in patients with Parkinson's disease tested on and off dopaminergic therapy. Neuroimage 185, 455–470.

Hoehn, M.M., Yahr, M.D., 1967. Parkinsonism: onset, progression, and mortality. Neurology 17 (5), 427-427.

Hölig, C., Berti, S., 2010. To switch or not to switch: brain potential indices of attentional control after task-relevant and task-irrelevant changes of stimulus features. Brain Res. 1345, 164–175.

Hughes, A.J., Daniel, S.E., Kilford, L., Lees, A.J., 1992. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J. Neurol. Neurosurg. Psychiatr. 55 (3), 181–184.

Humphries, M.D., Stewart, R.D., Gurney, K.N., 2006. A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. J. Neurosci. 26 (50), 12921–12942.

Humphries, M.D., Khamassi, M., Gurney, K., 2012. Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. Front. Neurosci. 6, 9.

Iglesias, S., Mathys, C., Brodersen, K.H., Kasper, L., Piccirelli, M., den Ouden, H.E., Stephan, K.E., 2013. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. Neuron 80 (2), 519–530.

Jeffreys, H., 1946. An invariant form for the prior probability in estimation problems. Proc. Roy. Soc. Lond. Math. Phys. Sci. 186 (1007), 453–461.

Kalbe, E., Calabrese, P., Kohn, N., Hilker, R., Riedel, O., Wittchen, H.U., Dodel, R., Otto, J., Ebersbach, G., Kessler, J., 2008. Screening for cognitive deficits in Parkinson's disease with the Parkinson neuropsychometric dementia assessment (PANDA) instrument. Park. Relat. Disord. 14 (2), 93–101.

Keeler, J.F., Pretsell, D.O., Robbins, T.W., 2014. Functional implications of dopamine D1 vs. D2 receptors: a 'prepare and select'model of the striatal direct vs. indirect pathways. Neuroscience 282, 156–175.

Keuken, M.C., Bazin, P.L., Crown, L., Hootsmans, J., Laufer, A., Müller-Axt, C., Sier, R., van der Putten, E.J., Schäfer, A., Turner, R., Forstmann, B.U., 2014. Quantifying inter-individual anatomical variability in the subcortex using 7 T structural MRI. Neuroimage 94, 40–46.

Kiebel, S.J., Daunizeau, J., Friston, K.J., 2008. A hierarchy of time-scales and the brain. PLoS Comput. Biol. 4 (11), e1000209.

Kluger, J., Schubotz, R.I., 2017. Strategic adaptation to non-reward prediction error qualities and irreducible uncertainty in fMRI. Cortex 97, 32–48.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19 (3), 1233–1239.

Mars, R.B., Debener, S., Gladwin, T.E., Harrison, L.M., Haggard, P., Rothwell, J.C., Bestmann, S., 2008. Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. J. Neurosci. 28 (47), 12539–12545.

Marshall, L., Mathys, C., Ruge, D., de Berker, A.O., Dayan, P., Stephan, K.E., Bestmann, S., 2016. Pharmacological fingerprints of contextual uncertainty. PLoS Biol. 14 (11), e1002575.

McCoy, B., Jahfari, S., Engels, G., Knapen, T., Theeuwes, J., 2019. Dopaminergic medication reduces striatal sensitivity to negative outcomes in Parkinson's disease. Brain 142 (11), 3605–3620.

Meder, D., Herz, D.M., Rowe, J.B., Lehéricy, S., Siebner, H.R., 2019. The role of dopamine in the brain-lessons learned from Parkinson's disease. Neuroimage 190, 79–93.

Mizumori, S.J., Tryon, V.L., 2015. Integrative hippocampal and decision-making neurocircuitry during goal-relevant predictions and encoding. Prog. Brain Res. 219, 217–242.

Monchi, O., Petrides, M., Doyon, J., Postuma, R.B., Worsley, K., Dagher, A., 2004. Neural bases of set-shifting deficits in Parkinson's disease. J. Neurosci. 24 (3), 702–710.

Murty, V.P., Sambataro, F., Radulescu, E., Altamura, M., Iudicello, J., Zoltick, B., et al., 2011. Selective updating of working memory content modulates meso-cortico-striatal activity. Neuroimage 57 (3), 1264–1272.

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25 (3), 653–660.

Nour, M.M., Dahoun, T., Schwartenbeck, P., Adams, R.A., FitzGerald, T.H., Coello, C., et al., 2018. Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. Proc. Natl. Acad. Sci. Unit. States Am. 115 (43), E10167–E10176.

O'Reilly, R.C., Frank, M.J., 2006. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. 18, 283–328.

Pasquereau, B., Turner, R.S., 2014. Dopamine neurons encode errors in predicting movement trigger occurrence. J. Neurophysiol. 113 (4), 1110–1123.

Pernet, C.R., 2014. Misconceptions in the use of the General Linear Model applied to functional MRI: a tutorial for junior neuro-imagers. Front. Neurosci. 8, 1.

Perugini, A., Ditterich, J., Basso, M.A., 2016. Patients with Parkinson's disease show impaired use of priors in conditions of sensory uncertainty. Curr. Biol. 26 (14), 1902–1910.

Perugini, A., Basso, M.A., 2017. Perceptual decisions based on previously learned information are independent of dopaminergic tone. J. Neurophysiol. 119 (3), 849–861.

Perugini, A., Ditterich, J., Shaikh, A.G., Knowlton, B.J., Basso, M.A., 2018. Paradoxical decision-making: a framework for understanding cognition in Parkinson's disease. Trends Neurosci. 41 (8), 512–525.

Picciotto, M.R., Higley, M.J., Mineur, Y.S., 2012. Acetylcholine as a neuromodulator: cholinergic signaling shapes nervous system function and behavior. Neuron 76 (1), 116–129.

Rae, C.L., Nombela, C., Rodríguez, P.V., Ye, Z., Hughes, L.E., Jones, P.S., et al., 2016. Atomoxetine restores the response inhibition network in Parkinson's disease. Brain 139 (8), 2235–2248.

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rauss, K., Schwartz, S., Pourtois, G., 2011. Top-down effects on early visual processing in humans: a predictive coding framework. Neurosci. Biobehav. Rev. 35 (5), 1237–1253.

Redgrave, P., Gurney, K., 2006. The short-latency dopamine signal: a role in discovering novel actions? Nat. Rev. Neurosci. 7, 967–975.

Roemmich, R.T., Hack, N., Akbar, U., Hass, C.J., 2014. Effects of dopaminergic therapy on locomotor adaptation and adaptive learning in persons with Parkinson's disease. Behav. Brain Res. 268, 31–39.

Schapiro, A.C., Gregory, E., Landau, B., McCloskey, M., Turk-Browne, N.B., 2014. The necessity of the medial temporal lobe for statistical learning. J. Cognit. Neurosci. 26 (8), 1736–1747.

Schiffer, A.M., Ahlheim, C., Wurm, M.F., Schubotz, R.I., 2012. Surprised at all the entropy: hippocampal, caudate and midbrain contributions to learning from prediction errors. PLoS One 7 (5), e36445.

Schmitz, Y., Benoit-Marand, M., Gonon, F., Sulzer, D., 2003. Presynaptic regulation of dopaminergic neurotransmission. J. Neurochem. 87, 273–289.

Schultz, W., Dickinson, A., 2000. Neuronal coding of prediction errors. Annu. Rev. Neurosci. 23, 473–500.

Schwartenbeck, P., FitzGerald, T.H., Dolan, R., 2016. Neural signals encoding shifts in beliefs. Neuroimage 125, 578–586.

Stelzel, C., Fiebach, C.J., Cools, R., Tafazoli, S., D'Esposito, M., 2013. Dissociable fronto-striatal effects of dopamine D2 receptor stimulation on cognitive versus motor flexibility. Cortex 49, 2799–2811.

Strange, B.A., Duggins, A., Penny, W., Dolan, R.J., Friston, K.J., 2005. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? Neural Network. 18 (3), 225–230.

Smittenaar, P., Chase, H.W., Aarts, E., Nusselein, B., Bloem, B.R., Cools, R., 2012. Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection–learning or performance? Eur. J. Neurosci. 35 (7), 1144–1151.

Summerfield, C., Egner, T., 2009. Expectation (and attention) in visual cognition. Trends Cognit. Sci. 13 (9), 403–409.

Trempler, I., Schiffer, A.M., El-Sourani, N., Ahlheim, C., Fink, G.R., Schubotz, R.I., 2017. Frontostriatal contribution to the interplay of flexibility and stability in serial prediction. J. Cognit. Neurosci. 29 (2), 298–309.

Trempler, I., Binder, E., El-Sourani, N., Schiffler, P., Tenberge, J.G., Schiffer, A.M., et al., 2018. Association of grey matter changes with stability and flexibility of prediction in akinetic-rigid Parkinson's disease. Brain Struct. Funct. 223 (5), 2097–2111.

Tomassini, A., Ruge, D., Galea, J.M., Penny, W., Bestmann, S., 2016. The role of dopamine in temporal uncertainty. J. Cognit. Neurosci. 28 (1), 96–110.

Tomassini, A., Pollak, T.A., Edwards, M.J., Bestmann, S., 2019. Learning from the past and expecting the future in Parkinsonism: dopaminergic influence on predictions about the timing of future events. Neuropsychologia 127, 9–18.

Turk-Browne, N.B., Scholl, B.J., Johnson, M.K., Chun, M.M., 2010. Implicit perceptual anticipation triggered by statistical learning. J. Neurosci. 30 (33), 11177–11187.

Uitvlugt, M.G., Pleskac, T.J., Ravizza, S.M., 2016. The nature of working memory gating in Parkinson's disease: a multi-domain signal detection examination. Cognit. Affect Behav. Neurosci. 16 (2), 289–301.

Vazey, E., Aston-Jones, G., 2012. The emerging role of norepinephrine in cognitive dysfunctions of Parkinson's disease. Front. Behav. Neurosci. 6, 48.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and waic. Stat. Comput. 27 (5), 1413–1432.

Vidaurre, D., Smith, S.M., Woolrich, M.W., 2017. Brain network dynamics are hierarchically organized in time. Proc. Natl. Acad. Sci. Unit. States Am. 114 (48), 12827–12832.

Vilares, I., Kording, K.P., 2017. Dopaminergic medication increases reliance on current information in Parkinson's disease. Nat. Hum. Behav. 1 (8), 0129.

Whitfield-Gabrieli, S., Nieto-Castanon, A., 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect. 2 (3), 125–141.

Williams, G.V., Castner, S.A., 2006. Under the curve: critical issues for elucidating D1 receptor function in working memory. Neuroscience 139 (1), 263–276.

Wolpe, N., Nombela, C., Rowe, J.B., 2015. Dopaminergic modulation of positive expectations for goal-directed action: evidence from Parkinson's disease. Front. Psychol. 6, 1514.

Wolpe, N., Zhang, J., Nombela, C., Ingram, J.N., Wolpert, D.M., Rowe, J.B., 2018. Sensory attenuation in Parkinson's disease is related to disease severity and dopamine dose. Sci. Rep. 8 (1), 15643.

Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited—again. Neuroimage 2 (3), 173–181.

Wylie, S.A., Van Den Wildenberg, W.P.M., Ridderinkhof, K.R., Bashore, T.R., Powell, V.D., Manning, C.A., Wooten, G.F., 2009. The effect of speed-accuracy strategy on response interference control in Parkinson's disease. Neuropsychologia 47 (8–9), 1844–1853.

Wylie, S.A., Ridderinkhof, K.R., Bashore, T.R., van den Wildenberg, W.P., 2010. The effect of Parkinson's disease on the dynamics of on-line and proactive cognitive control during action selection. J. Cognit. Neurosci. 22 (9), 2058–2073.

Ye, Z., Altena, E., Nombela, C., Housden, C.R., Maxwell, H., Rittman, T., et al., 2015. Improving response inhibition in Parkinson's disease with atomoxetine. Biol. Psychiatr. 77 (8), 740–748.

Ye, Z., Rae, C.L., Nombela, C., Ham, T., Rittman, T., Jones, P.S., et al., 2016. Predicting beneficial effects of atomoxetine and citalopram on response inhibition in Parkinson's disease with clinical and neuroimaging measures. Hum. Brain Mapp. 37 (3), 1026–1037.

Yu, A.J., Dayan, P., 2002. Acetylcholine in cortical inference. Neural Network. 15 (4–6), 719–730.

Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. Neuron 46 (4), 681–692.

Yu, Y., FitzGerald, T.H., Friston, K.J., 2013. Working memory and anticipatory set modulate midbrain and putamen activity. J. Neurosci. 33, 14040–14047.