Bidirectional modulation of reward-guided decision making by dopamine

Ana Antonia Dias Maile[1], Theo OJ Gründler[2], Monja I Froböse[1], Hannah Kurtenbach[1], Luca F Kaiser[1], Gerhard Jocham[1,2]

[1]Biological Psychology of Decision Making, Institute of Experimental Psychology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
[2]Center of Behavioral Brain Sciences, Otto von Guericke University Magdeburg, 39106 Magdeburg; Germany

Correspondence: Ana Antonia Dias Maile
Building 25.32.03.28
Heinrich Heine University Düsseldorf
Universitätsstraße 1
40225 Düsseldorf
+49 211 81-14123
ana.dias.maile@uni-duesseldorf.de

Abstract

The neuromodulator dopamine is known to play a key role in reward-guided decision making, where choice options are often characterized by multiple attributes. Different decision strategies can be used to merge these choice attributes with personal preferences (e.g. risk preferences) and integrate them into a single subjective value. While the influence of dopamine on risk preferences has been investigated, it is unknown whether dopamine is also involved in arbitrating between decision strategies. We investigated this using a reward-guided decision-making task which was performed by 31 healthy participants under the influence of the dopamine D_2/D_3-receptor antagonist amisulpride, the dopamine precursor L-DOPA, or placebo in a double-blind within-subject design. Notably, we observed that the dopaminergic interventions shifted the (overall) weighting of option attributes without changing how option attributes are integrated into a subjective value (decision strategy). These effects were bidirectional: Amisulpride reduced the degree to which choices were influenced by both reward magnitude and reward probability, whereas the opposite was observed under L-DOPA, where we found an increased effect of reward magnitude and reward probability on choice. These effects occurred in the absence of changes in statistically optimal behavior. Together, our data provide evidence for a role of dopamine in controlling the influence of value parameters on choice irrespective of decision strategies.

Introduction

Consider the daily commute to work and the choice of transportation: While biking boosts environmental and financial benefits, it is less comfortable and restricted by weather conditions. In contrast, driving provides flexibility and comfort, yet it has environmental and financial drawbacks. The choice of transportation thus follows from the direct comparison of various choice attributes, demanding the use of decision strategies and attribute weighting (e.g., risk preferences). Two lines of evidence suggest that decision strategies and preferences are controlled by dopamine. First, patients with the neurodegenerative disorder Parkinson's Disease (PD), which is characterized by a profound dorsal striatal dopamine depletion, exhibit motoric deficits (e.g., tremor and bradykinesia) often accompanied by an increase in risk aversive behavior [1, 2]. Conversely, treating the dopamine deficit with the dopamine-precursor levodopa (L-DOPA, [3]) improves the motor symptoms but it can also increase risk seeking behavior and even cause impulse control disorders such as pathological gambling [4-6].

Second, pharmacological studies in healthy volunteers have shown increased risk seeking after increasing dopamine concentrations by administering L-DOPA [7]. Conversely, similar effects were found when administering the D_2/D_3-receptor antagonist amisulpride [8]. Importantly, dopaminergic effects on risk preferences are valence-specific and were exclusively observed in the context of rewards, not losses [7, 8]. In previous work, choices were usually modelled using Prospect Theory, which posits that, following a non-linear distortion, reward probabilities and magnitudes are multiplicatively combined into an expected value (EV) which is then used to guide decisions [9, 10]. However, an alternative possibility would be that participants do not use such integrated EV at all, but instead rely on simpler heuristics, such as directly comparing attributes and using a weighted combination of attribute differences. Recent

evidence suggests that choices of both humans and non-human primates are best characterized by a mixture of such an additive and a multiplicative strategy [11-13]. These results further suggested that the degree to which either of the two strategies prevails varies as a function of uncertainty about option attributes [11]. Despite the known role of dopamine in both decision making and risk preferences, the effects of dopamine on arbitrating between decision strategies remain untested.

Therefore, in the present study we investigated the role of dopaminergic activity on decision strategies in reward-guided decision making, combining it with computational modeling approaches that take into account both multiplicative and additive strategies for value computation. Additionally, we sought to clarify the dopaminergic effect on risk preferences by including a parameter reflecting the relative balance between reliance on reward probability versus reward magnitude. Healthy participants performed a reward-guided decision-making task under the influence of either the D_2/D_3-receptor antagonist amisulpride (400 mg), the dopamine precursor L-DOPA (100 mg + 25 mg Carbidopa), or placebo in a double-blind within-subjects design. Notably, there was no significant dopaminergic effect on participants' decision strategies or risk preferences. Instead, the weighting of choice attributes was shifted. The degree to which participants' choices were governed by both reward magnitude and reward probability was diminished under amisulpride, but increased under L-DOPA.

Methods

Participants

A total of 33 participants took part in the study. Due to the pharmacological manipulation, participants were pre-screened on a separate day to exclude certain pre-existing medical conditions, including neurological or psychiatric disorders. Additionally, an electrocardiogram (ECG) was obtained from each potential participant

and assessed by a cardiologist. Only healthy participants without ECG abnormalities were allowed to participate in the study. Due to the menstrual cycle-dependent interaction between gonadal steroids and the dopaminergic system [14-16], we only included male participants, as in previous work [17, 18]. Further exclusion criteria were drug abuse, and use of psychoactive drugs or medication in the two weeks before the experiment. Participants were instructed to abstain from alcohol and any other drugs of abuse during the entire course of the study.

Out of the initial 33 participants, 31 fully completed the study. They were right-handed with normal or corrected-to-normal (*N* = 12) vision. On average, they were *M* = 25.71 years old (age range 21-32, *SD* = 3.20), body weight was *M* = 77.16 kg (weight range 62 – 90kg, *SD* = 7.46). Six participants reported occasional smoking. All were naive to the purpose of the study and gave written informed consent. The present study was approved by the local ethics committee of the Medical Faculty of the Otto-von-Guericke-University Magdeburg (internal reference: 129/13) and is in line with the Helsinki Declaration of 1975. Participants were compensated for their efforts at a fixed rate, plus an additional bonus that depended on their performance during the decision-making task.

Procedure

The study comprised three pharmacological MEG-sessions in a pseudorandomized order across participants. MEG recordings will be disregarded for the purpose of the present study. Identical procedures were followed on the three sessions, only the drug (or placebo) administered differed. A physician screened participants before each session. In a double-blind cross-over design, participants received a single oral dose of either the dopamine D_2/D_3-receptor antagonist amisulpride (400mg), the dopamine precursor levodopa (L-DOPA; 100 mg L-DOPA + 25mg cardidopa), or placebo. Because of the different pharmacokinetics of L-DOPA and amisulpride, a dummy

administration procedure was used. Participants ingested two pills separated by three hours, of which at least one always contained placebo. In the amisulpride condition, the active substance was contained in the first pill, whereas it was contained in the second pill in the L-DOPA condition. The decision-making task began approximately 1 h after ingestion of the second pill, corresponding to 1 h after L-DOPA administration and 3 h after amisulpride administration in line with the average time for the two drugs to reach peak plasma concentration [19, 20]. Immediately prior to the MEG recording, participants completed Bond & Lader visual analogue scales [BL-VAS, 21] and the trail-making task [TMT, 22] to assess drug effects on mood and visual attention. Heart rate and blood pressure were monitored prior to the first drug (or placebo) administration and after the study, before participants were released by the study physician. Sessions were separated by at least 8 days to ensure complete washout of the drug before the next session (elimination half-life of amisulpride 12 h, L-DOPA with cardidopa 1.5 h).

Reward-guided decision-making task

Participants completed 500 trials of a reward-guided decision-making task plus 12 additional trials to familiarize themselves with the task beforehand (Figure 1A). On each trial, choices were made between two options defined by two attributes each: a reward magnitude and a probability to obtain this reward. Reward magnitudes were presented as the width of a rectangular horizontal bar and reward probabilities as a percentage written underneath these bars. Therefore, both attributes were explicit and did not have to be learned. Reward outcomes were independent of each other, meaning either of the two options, both, or none of them could be rewarded. Option values were drawn from a reward schedule that was generated before the experiment, as in previous studies [23-26]. This reward schedule was designed such that correlations between factors of interest (in particular between chosen and unchosen

and between left and right option value) was minimized. Furthermore, we ensured that reward magnitude and probability were never identical for the two options. To make advantageous choices, participants needed to multiply magnitude and probabilities into an integrative value estimate referred to as EV. On some trials, however, both magnitude and probability of one option were higher than on the alternative option. We refer to these trials as ‘no brainer’ trials. These were limited to occur no more than in 17.8 % of the 500 trials (89 trials).

The two options were presented sequentially such that on every trial, the left option was always presented first, followed by a short delay (200 ms - 400 ms), which was followed by a presentation of the second option on the right side. Immediately after offset of the second option, a question mark appeared prompting participants to indicate their choice. Participants were instructed to refrain from responding while the second option was onscreen and only respond after the question mark had appeared. This design choice was made in order to separate sensorimotor cortical value representations from the execution of movement in the MEG signal. Note that this design precludes a meaningful analysis of reaction times.

Options were selected by button presses with the index finger of the left and right hand, respectively. If an option was rewarded on the current trial, the bar representing the reward magnitude turned green, if it was not rewarded, it turned red. When the chosen option was rewarded, an amount proportional to the reward magnitude of that option was added to the bonus and a blue progress bar displayed at the bottom of the screen indicated this to the participant. To motivate participants, the progress bar displayed earnings towards 2 € and reset itself after reaching this goal. All earned points counted towards the reward. Rewards were rounded to the next higher value in Euro. Even though the outcome of the unchosen option was irrelevant to the bonus, the outcome of both the chosen and the unchosen option was always

presented to illustrate independence of outcomes and stimulate risk seeking (i.e., occasional win of low reward probabilities). Outcome presentation was followed by an inter-trial interval (blank screen). The stimuli were presented on a grey (RGB: 60, 60, 60) background with a contrast optimized for the MEG recording chamber (white: 130, 130, 130; green: 46, 139, 60; red: 178, 70, 70; blue: 5,105,204).

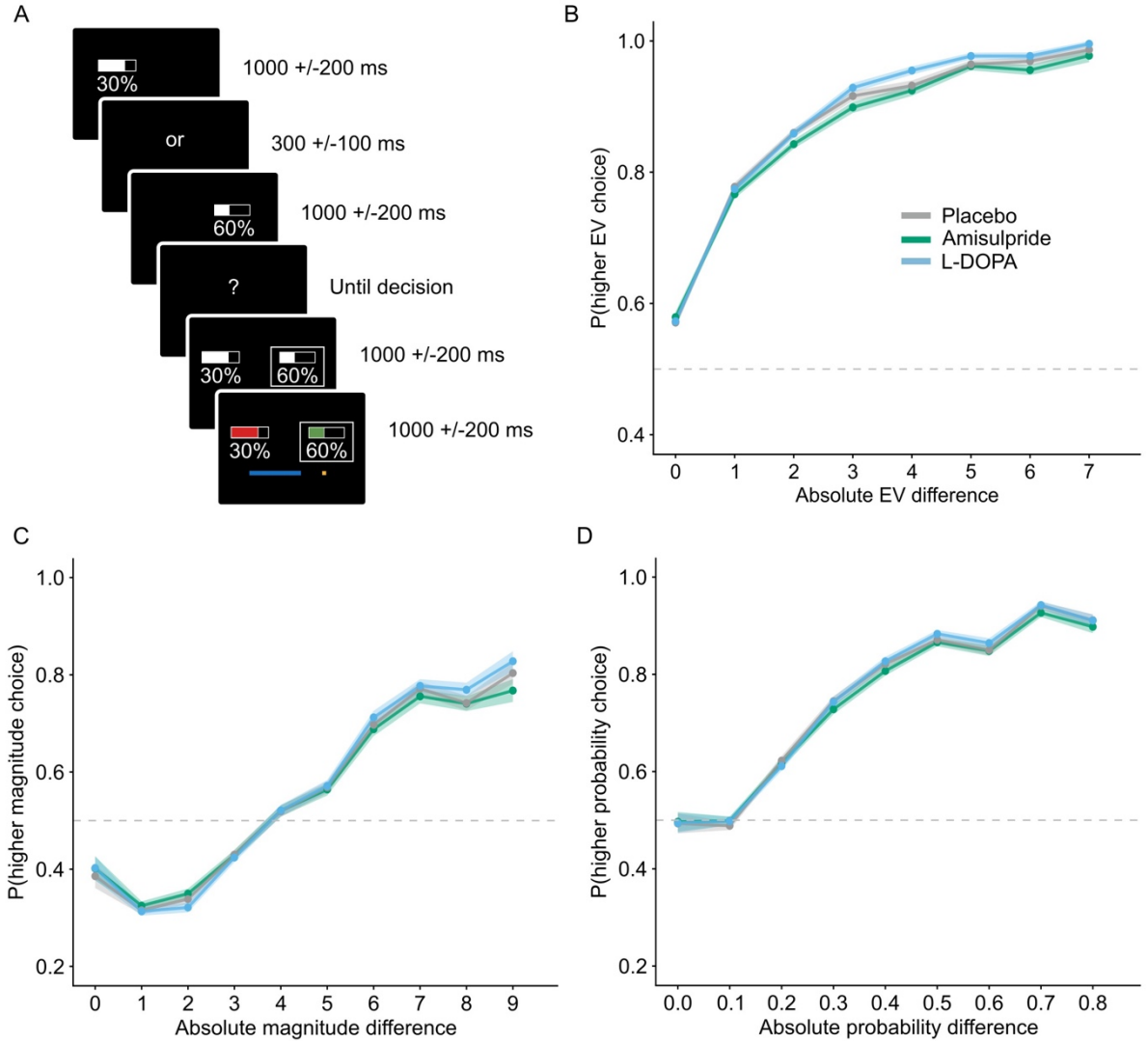

Figure 1. Schematic of experimental task and participants' behavior. (A) Task schematic and trial timeline. Each choice option was associated with a reward magnitude (width of the horizontal bar) and a reward probability (percentage). The left option was always presented first, followed by a short delay before the presentation of the right option. Participants were instructed to only respond once the question mark appeared. Lastly, they were given feedback on the outcome (reward/no reward) of both the chosen and the unchosen option and reward, if obtained, was added to the progress bar (blue bar at the bottom of the screen). (B) Probability of choosing the choice option with a higher expected value (EV) depending on the absolute EV difference between the two options. Note this is not the mean-centered EV and thus not independent of reward magnitude or probability. (C) Probability of choosing the option with higher reward magnitude depending on the absolute magnitude difference between the two options. (D) Probability of choosing the option with a higher reward probability depending on the absolute probability difference between the two options. In B-D, solid lines represent the average, shaded areas represent *SEM* across participants.

Statistical analyses

Mixed-effects modeling

Data analyses were conducted using R (R version 4.2.2; [27]). All data and codes are made available on OSF (https://osf.io/rnc94/). To evaluate drug effects on decision making, behavioral data was first analyzed using linear mixed modeling, implemented in the lme4 package (version 3.1-3; [28]). In contrast to analyses of variance (ANOVAs) and linear regressions, this approach allows taking into account both within- and between subject variability [29-31]. Since our main interest was in comparing each drug to placebo, linear mixed models were run separately for the two drugs. As the dependent variable was choice (left vs. right), we used general linear mixed effects models with a binomial distribution. Fixed effects included drug (amisulpride vs. placebo or L-DOPA vs. placebo), the option values (magnitude, probability, mean-centered EV) and a repetition bias (i.e., a tendency to choose the same choice side as in the trial before irrespective of choice attributes). Additionally, the model encompassed the interaction of drug with the option values and repetition bias. To analyze the effects of drug on option values, these were converted into the z-scaled difference scores between the two options. Hence, a drug-induced shift in risk preferences would be evident from a differential (or asymmetric) change of the interaction regression weights for reward magnitude and probability, respectively. Due to convergence issues, a full random-effects structure was not possible [32]. Nevertheless, linear mixed models included both a random intercept and a random slope for the within-subject factor drug. To investigate the task effects including data from all sessions irrespective of drug, a separate linear mixed model was set up that included the option values and a repetition bias but disregarded the factor drug. Furthermore, to investigate drug effects depending on the presentation order of the two sequentially presented options, an additional linear mixed model included option

values of each choice option individually instead of the difference score (tables S3 and S6). If we observed significant drug effects on control measures (BL-VAS, TMT, heart rate and blood pressure), these effects were added to the task models as fixed effects and we established that the effects of interest remained unchanged. A further control analysis encompassed adding the session number as fixed effects to control for putative order effects. All *p*-values are based on asymptotic Wald tests.

Computational modeling

To analyze the drug effects on the underlying cognitive processes determining participants' choice behavior, we used a hierarchical Bayesian model (similar to [33]). This method effectively incorporates the within-subject design enabling estimation of the group-level and subject-level parameters. However, due to practicability, we first identified the best fitting model using non-hierarchical models. The first model, the multiplicative model, assumes that reward magnitudes and probabilities are multiplicatively combined [11]:

$$sv_{i,t} = P_{i,t} \cdot M_{i,t} \quad (1)$$

Where $sv_{i,t}$ is the subjective value, and $M_{i,t}$ and $P_{i,t}$ are the reward magnitude and probability, respectively, of option *i* presented on trial *t* [11].

In contrast, in the additive model, subjective values are computed from a weighted additive combination of reward magnitudes and probabilities:

$$sv_{i,t} = \omega_p \cdot P_{i,t} + (1 - \omega_p) \cdot M_{i,t} \quad (2)$$

Where ω_p is the relative weight given to probabilities relative to magnitudes. Values of $\omega_p = 0.5$ thus indicate equal influence of probability and magnitude. Values of $\omega_p > 0.5$ indicate that choices are more strongly influenced by probabilities than magnitudes, indicating risk aversion. In contrast, values of $\omega_p < 0.5$ indicate risk seeking [13]. For

this purpose, magnitudes were scaled to values between 0.10-1.00 aligning them to the value range of reward probabilities.

The third model, the hybrid model assumes that participants use a weighted combination of both the multiplicative and additive strategy:

$$sv_{i,t} = \omega_{mult} \cdot P_{i,t} \cdot M_{i,t} + (1 - \omega_{mult}) \cdot (\omega_p \cdot P_{i,t} + (1 - \omega_p) \cdot M_{i,t}) \quad (3)$$

Where the relative allocation between the multiplicative and additive response strategy is governed by the parameter ω_{mult} [11, 13], which is bound between 0.00 - 1.00, with higher values of ω_{mult} indicating a dominance of the multiplicative response strategy. In all models, choices were modelled using a softmax choice rule including an inverse temperature τ to capture choice stochasticity, which then generates a probability to choose action *A* (in this case, left or right):

$$p_{A,t} = \frac{1}{1+e^{-(\Delta V_t \cdot \tau)}} \quad (4)$$

Where ΔV is the value difference (left minus right for *A* = left choice).

The linear mixed effect models revealed a (negative) choice repetition bias (see results). Therefore, we extended the best-fitting model by adding a choice bias. Furthermore, to relate to previous work and compare model fits, we used Prospect Theory. This included models where either only reward magnitude (equation 5), only reward probability (equation 6), or both (equation 7) were non-linearly distorted.

$$sv_{i,t} = P_{i,t} \cdot {M_{i,t}}^{\alpha} \quad (5)$$

$$sv_{i,t} = \frac{{P_{i,t}}^{\gamma}}{({P_{i,t}}^{\gamma} + (1 - P_{i,t})^{\gamma})^{\frac{1}{\gamma}}} \cdot M_{i,t} \quad (6)$$

$$sv_{i,t} = \frac{{P_{i,t}}^{\gamma}}{({P_{i,t}}^{\gamma} + (1 - P_{i,t})^{\gamma})^{\frac{1}{\gamma}}} \cdot {M_{i,t}}^{\alpha} \quad (7)$$

Where α and γ are the parameters that describe the non-linear warping of objective into subjective attributes. For this purpose, reward magnitudes were not rescaled and thus ranged between 1-10. Models were implemented using the NLoptr package (version 2.0.3; [34]) with bound optimization by quadratic approximation (bobyca, [35]). The maximum number of function evaluations (*maxeval*) was 10000 and the stopping criterion for relative change (*xtol_rel*) was set to 1.0e-08. Models were run with 100 iterations. The selection of the best-fitting model was based on the Bayes Information Criterion (BIC) for all testing sessions.

Next, the best-fitting model was implemented as a hierarchical Bayesian model to analyze the effects of the dopaminergic drugs. In this approach, group-level parameters (X) provide priors for the individual-level parameters (x), with $x \sim \mathcal{N}(X, \sigma)$. A half-Cauchy with a scale of 2 defined the hyperpriors for σ. With the exception for X_β, the hyperpriors for X were centered around 0: $X_{mult,p} \sim \mathcal{N}(0, 2)$, $X_\beta \sim \mathcal{N}(2, 3)$. An inverse logit transform constrained ω_{mult} and ω_p between 0.00-1.00. The inverse temperature τ was positively bounded through an exponential transform. Parameter initials were chosen by training the model on an independent dataset using a similar task [36]. We further extended the model by incorporating six parameter shifts (three per drug), enabling analysis of drug-dependent effects on the model parameters (ω_{mult}, ω_p and τ). These were unconstrained and their hyperpriors were specified by $\mathrm{N}(0, 3)$. Initial values of the parameter shifts were set to 0.00. Additionally, we used a similar approach for the best-fitting of the three variants of the Prospect Theory model, enabling a comparison of the effects of the dopaminergic drugs within a flexible attribute weighting model with the traditionally used model (Figure S2). Hierarchical Bayesian models were implemented using the RStan package (version 2.21.8; [37]), which enables full Bayesian inference with Markov chain Monte Carlo (MCMC)

sampling methods. There were four Markov chains, comprising 500 warm-up iterations and 2500 post burn-in iterations per chain (10000 total). The target average acceptance probability (adapt_delta) was set to 0.97. Model convergence was verified using the convergence and diagnostics criteria provided by RStan which include $\hat{R}$ < 1.05.

To evaluate the best-fitting model's ability to capture our data accurately, we generated 500 simulated datasets based on the posterior distributions of subject-level parameters obtained from the hierarchical Bayesian model. These simulated datasets were compared to key behavioral features of the real data e.g., likelihood of choosing the option with a higher EV, reward magnitude or reward probability.

Results

Bidirectional modulation of the reliance of choices on reward information by amisulpride and L-DOPA

In the reward-guided decision-making task, participants were more likely to choose the option with a higher EV, reward magnitude or probability, as evidenced by the descriptive statistics (Figure 1B-D). This is also confirmed by the main effects of task manipulation in the linear mixed models: EV, magnitude and probability, operationalized as a difference score between the two options, have a significant main effect on choice behavior (EV: $\beta \approx 0.90$, $SE \approx 0.12$, $z(39671) \approx 7.42$, $p < 0.001$; magnitude: $\beta \approx 2.51$, $SE \approx 0.24$, $z(39671) \approx 10.31$, $p < 0.001$; probability: $\beta \approx 3.79$, $SE \approx 0.26$, $z(39671) \approx 14.37$, $p < 0.001$). Accordingly, participants seem to weight magnitudes and probabilities more strongly than the EV as indicated by the numerically higher regression weights. As for the dopaminergic interventions, amisulpride decreased the effect of both reward magnitude and probability on choice ($\beta \approx -0.08$, $SE \approx 0.03$, $z(26143) \approx -2.85$, $p \approx 0.004$; and $\beta \approx -0.12$, $SE \approx 0.03$, $z(26143) \approx -3.62$, $p < 0.001$; for the interaction of amisulpride with magnitude and probability, respectively,

Figure 2A-B). In contrast, L-DOPA increased the weighting of reward magnitude and probability on choice ($\beta \approx 0.07$, $SE \approx 0.03$, $z(26591) \approx 2.48$, $p \approx 0.013$; and $\beta \approx 0.11$, $SE \approx 0.04$, $z(26591) \approx 3.04$, $p \approx 0.002$; for the interaction of L-DOPA with magnitude and probability, respectively, Figure 2C-D). Notably, neither of the dopaminergic drugs had an effect on the influence of EV on choice ($\beta \approx -0.02$, $SE \approx 0.02$, $z(26143) \approx -1.18$, $p = 0.237$; and $\beta \approx 0.01$, $SE \approx 0.02$, $z(26591) \approx 0.43$, $p = 0.666$; for the interaction of amisulpride and L-DOPA with EV). Participants displayed a tendency to alternate between right and left choices from trial to trial, evident from a negative regression weight of the repetition bias ($\beta \approx -0.10$, $SE \approx 0.04$, $z(39671) \approx -2.53$, $p = 0.011$). This alternation bias however was not affected by the dopaminergic interventions ($\beta \approx -0.05$, $SE \approx 0.03$, $z(26143) \approx -1.53$, $p = 0.127$; and $\beta \approx 0.02$, $SE \approx 0.03$, $z(26591) \approx 0.62$, $p = 0.535$; for the interaction of amisulpride and L-DOPA with the repetition bias). Irrespectively, due to the significant impact of the repetition bias on choice behavior, it was taken into account in the following computational modeling. For a complete overview of the linear mixed model results see tables S1 and S3.

We further confirmed that the drug effects were robust to the following control measures: All dopaminergic effects were independent of the session order (tables S2 and S4), drug effects on mood, visual attention, heart rate or blood pressure (note that L-DOPA increased participants' calmness independently of the effects on choice behavior; tables S7-11). Given the sequential presentation of choice options, we further controlled for effect specific to the presentation order. The effects of L-DOPA and amisulpride on the weighting of reward probabilities and the effect of amisulpride on the weighting of magnitudes were independent of the presentation order (tables S3 and S6). The interaction between L-DOPA and reward magnitude was diminished for the second presented option ($\beta \approx 0.04$, $SE \approx 0.02$, $z(26604) \approx 1.89$, $p \approx 0.059$; tables S3 and S6).

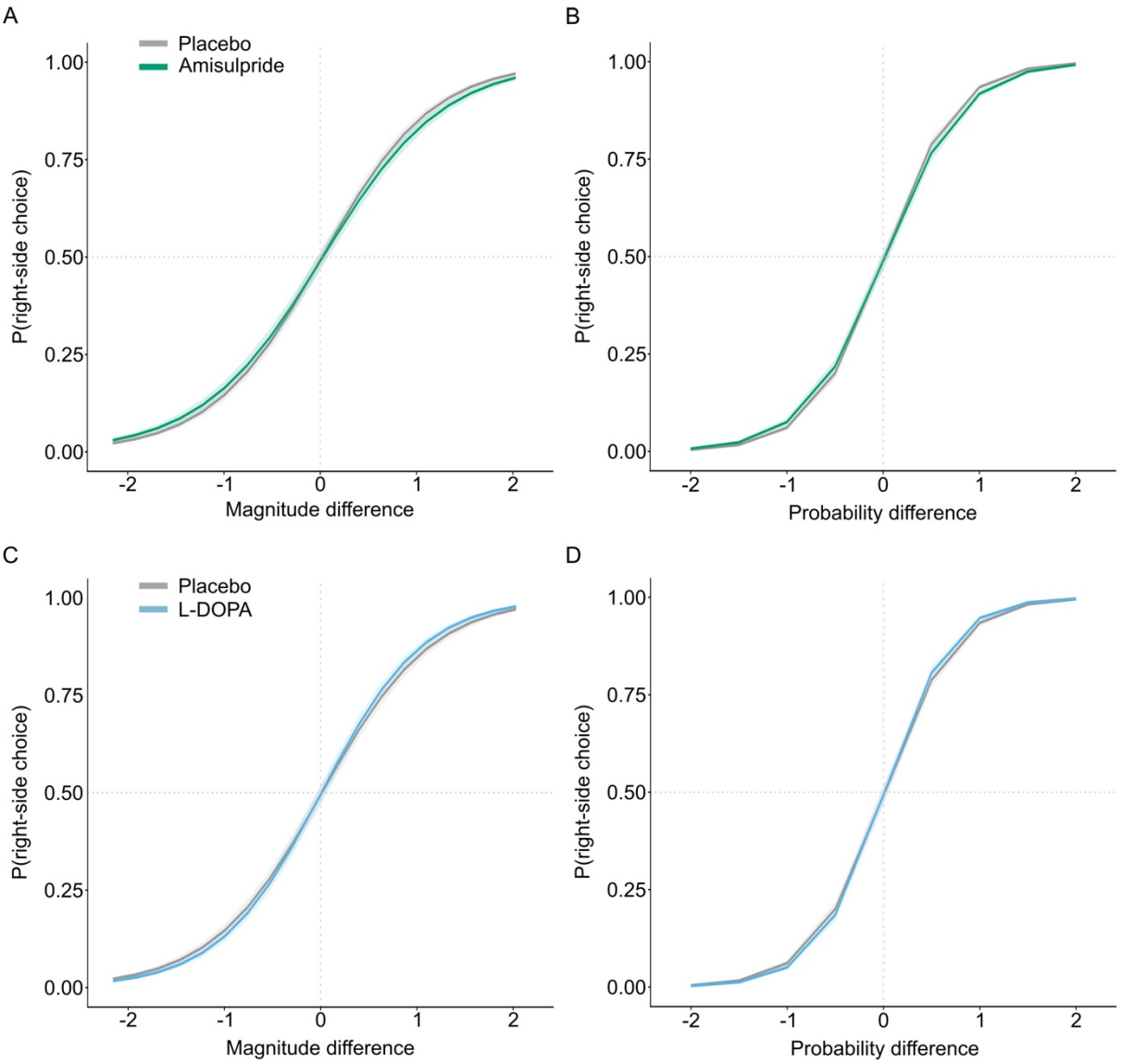

Figure 2. General linear mixed model results. The plots indicate the modelled probability to select the right option as a function of reward magnitude (A, C) or reward probability (B, D) difference between right minus left option (z-scored difference values). Amisulpride significantly decreased the influence of both reward magnitude (A) and reward probability (B) on choice. In contrast, L-DOPA significantly increased the influence of both reward magnitude (C) and reward probability (D) on choice. Solid lines represent the average, shaded areas represent SEM across participants.

Computational modeling shows no dopaminergic effect on decision strategies

The above results show that amisulpride and L-DOPA changed the degree to which choices were controlled by the individual choice attributes, magnitude and probability. Specifically, amisulpride decreased, and L-DOPA increased the effect of both reward magnitude and probability on choice behavior. Our computational models additionally allow us to further delineate whether these effects arise from an increase in choice stochasticity, as captured by the inverse softmax temperature (τ) and/or from a shift in

risk preference (ω_p), the relative importance of probability vs magnitude information. Additionally, the computational models allow insights into dopaminergic effects on the selection of decision strategies (ω_{mult}). The best-fitting model for all testing sessions was the hybrid model without choice bias (BIC *m* = 314.67; Table 1 for model fits). In this model the parameter ω_{mult} was estimated as *Mdn* = 0.49 (95%- HDI: 0.37 – 0.61, Figure 3A). This suggests participants used a hybrid response strategy comprising of both a multiplicative and an additive approach instead of a purely multiplicative strategy as traditionally assumed by Prospect Theory. However, this response strategy was not significantly shifted by amisulpride (*Mdn* = 0.02, 95%- HDI: -0.04 – 0.10, Figure 3B) or L-DOPA (*Mdn* = -0.01, 95%- HDI: -0.08 – 0.05, Figure 3B) as the posterior distributions overlap with zero. This indicates no dopaminergic effect on the selection of decision strategies. The weight parameter ω_p reflects the balance between reward probability and magnitude and thus gives insights into participants' risk preferences. It was estimated to be *Mdn* = 0.82 (95%-HDI: 0.71 – 0.93, Figure 3C). Hence, participants weighted probabilities more strongly than magnitudes causing risk aversive behavior. This parameter was not significantly shifted by amisulpride (*Mdn* = 0.04, 95%- HDI: -0.04 – 0.13, Figure 3D) or L-DOPA (*Mdn* = -0.01, 95%- HDI: -0.09 – 0.06, Figure 3D). Therefore, the shifted weighting of choice attributes as observed in the linear mixed models, cannot be attributed to a dopaminergic shift in risk preferences. The inverse softmax temperature τ is estimated *Mdn* = 14.56 (95%-HDI: 12.20 – 16.92, Figure 3E). It was not significantly shifted by amisulpride (*Mdn* = -0.46, 95%- HDI: -1.18– 1.02, Figure 3F) or L-DOPA (*Mdn* = 0.62, 95%- HDI: -0.68 – 1.94, Figure 3F). Hence, there is no dopaminergic effect on participants' choice stochasticity and thus this cannot clarify the dopaminergic shift of choice attributes found in the linear mixed models. Nevertheless, the direction of the drug shifts aligns with the direction of effects found in the linear mixed models: amisulpride decreases the reliance on choice attributes as

demonstrated here by the numerically decreasing inverse softmax temperature leading to an increased choice stochasticity. In contrast, L-DOPA increased the weighting of choice attributes aligned with the numerically increased inverse softmax temperature and hence a decreased choice stochasticity. Figure S1 shows simulated data from the used model reproducing the key features of our data. Further, the best-fitting Prospect Theory model included both a probability and magnitude distortion. Model results can be found in the Figure S2. Similar to the hybrid model, there were no significant drug shifts. Nevertheless, although it was not the best fitting model for our data, there was a trend effect for reduced probability distortion under amisulpride (*Mdn* = 0.11, 95%-HDI: -0.01 – 0.24), where 91,1% of the posterior distribution does not overlap with zero.

Table 1. Model fits for the different computational models. Bayes Information Criterion (BIC) was used to compare model fits. Lower BIC indicates better fit. The hybrid model without a bias is the best fitting model.

	BIC
Hybrid model	314.67
Hybrid model with repetition bias	319.56
Additive model	336.49
Multiplicative model	400.95
Prospect Theory with magnitude distortion	319.58
Prospect theory with probability distortion	364.93
Prospect Theory with magnitude and probability distortion	317.69

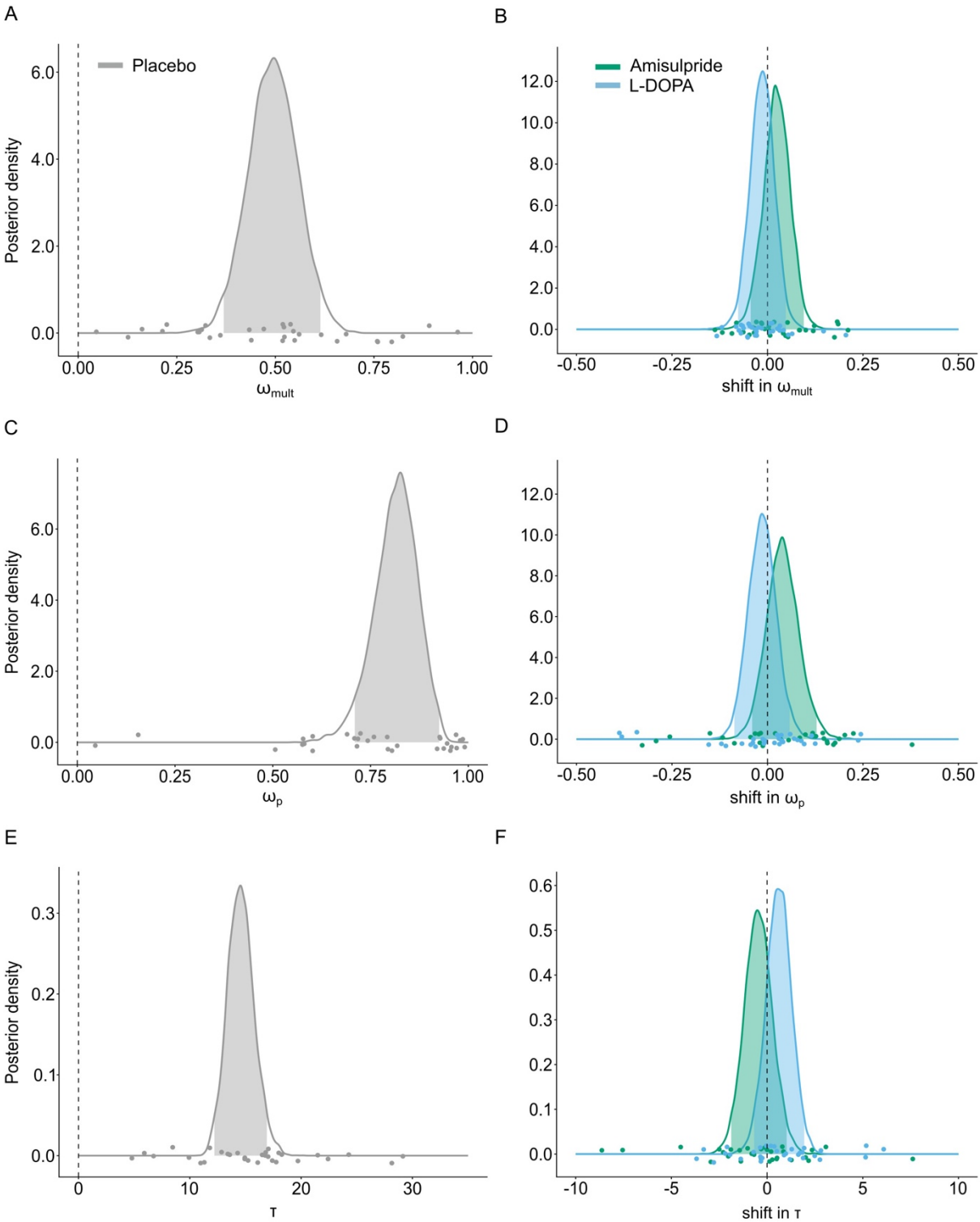

Figure 3. Posterior distributions of the group level estimate of the parameters from the hierarchical Bayesian Model. Panels on the left (A, C, E) are the posterior distributions of the three model parameters, the multiplicative weight ω_{mult}, the relative probability weight ω_P and the softmax inverse temperature τ. Panels on the right (B, D, F) are the posterior distributions of the shift parameters, color-coded for each drug. Shaded areas in the distributions are the 95%-CI. Dots represent single-subject estimates. (A) Participants use a hybrid response strategy comprising of both a multiplicative and an additive decision strategy, as evident from median ω_{mult} = 0.49. (B) This response strategy is not significantly affected by amisulpride or L-DOPA (95%-CI overlaps with zero). (C) Participants choices are guided more by reward probabilities than reward magnitudes, as evident from ω_P higher than 0.5 (median ω_P = 0.82). (D) This relative weighting is not significantly shifted by amisulpride or L-DOPA (95%-CI overlaps with zero). (E) The softmax inverse temperature τ is not significantly shifted by amisulpride or L-DOPA (F), since the 95%-CI overlaps with zero.

Discussion

We investigated the effects of dopaminergic manipulation on reward-guided decision making in a double-blind within-subject design in healthy participants. We found that increasing dopaminergic transmission with L-DOPA increased the degree to which participants' choices were governed by both option attributes, reward magnitude and probability. Blockade of D_2-like receptors with amisulpride had the opposite effect, decreasing the effect of both magnitude and probability on choice.

We investigated drug effects using two main approaches. First, linear mixed modeling allowed testing the task parameters that influenced participants choices, and how these were affected by drug. We found that the regression weights for reward magnitude and probability were decreased under amisulpride, and increased under L-DOPA. In contrast, the weight of the integrated EV was affected by neither of the drugs. Because choosing on the basis of EV represents statistically optimal behavior, this implies that drugs did not affect optimal behavior. We also observed that regression weights for the individual attributes, reward probability and reward magnitude, were higher than those for EV, an issue that we will return to in the next paragraph. Further, in agreement with previous work [38, 39], we found an alternation bias. That is, beyond the effects of value-related parameters, participants displayed a tendency to alternate between right and left choices from trial to trial. Again, this was not modulated by drug.

Our second approach is based on computational modeling. Traditionally, multi-attribute decision making is captured with models like Prospect Theory [9, 10]. These approaches typically involve applying non-linear weighting functions to reward magnitude and probability prior to multiplying them into an EV. This has recently been challenged by studies showing that behavior of both humans and non-human primates is best characterized by a more flexible mixture of decision strategies [11-13], comprised of both EV (multiplicative strategy) and a direct comparison of option

attributes (additive strategy). The dominance of either the multiplicative or additive strategy appears to depend on uncertainty where participants have been shown to shift to preferred use of the additive strategy during risky decision making or under heightened uncertainty [11, 12]. Our modeling results agree with these previous studies. We found that participants' behavior was best described by a hybrid model incorporating a multiplicative and an additive strategy, which outperformed classic multiplicative models like Prospect Theory. Investigation of the parameter ω_{mult}, which captures the relative contribution of the two strategies, showed that participants' subjective values were guided to a similar degree by the multiplicative and additive component. This is in line with the linear mixed modeling regression results, where we also observed that magnitudes and probabilities influenced choice behavior numerically stronger than EV. Further, we reasoned that the dopaminergic change in weighting of the individual attributes, without a concomitant change in the role of EV, should be reflected in choice stochasticity. Therefore, we expected softmax inverse temperature to be decreased under amisulpride, and increased under L-DOPA, corresponding to more and less stochastic choices, respectively. We found no significant effects of amisulpride or L-DOPA on any of the three parameters of the hierarchical Bayesian model. Nevertheless, numerically, the direction of the drug shifts on the softmax temperature does align with our prediction: Participants' choice behavior tended to get more stochastic under amisulpride, and less stochastic under L-DOPA. The lack of change in the relative probability weighting parameter ω_P and the multiplicative term ω_{mult} is again consistent with our regression-based results. A change of ω_P would only be coherent if a drug affected the impact of probability and magnitude in opposite direction - or at the very least if a change in the weight of probability was more/less pronounced than the change in the weight of magnitude. Likewise, a change in ω_{mult} should have readily been observed in the regression weight for EV.

Dopaminergic modulation of reward-guided decision making

Dopamine plays a key role in decision making, yet its involvement in arbitrating between multiplicative and additive decision strategies is not known. There is however evidence that dopamine influences risk preferences ([7, 8, 40-45]. For instance, increasing dopamine transmission with L-DOPA has been reported to increase risk seeking behavior [7, 41]. This effect appears to be mediated by participants' impulsivity and by their sensitivity to reward [43, 45]. In contrast, D_2-receptor antagonism has been associated with a decreased risk aversion, as evident from a reduced probability distortion [8, 40]. This is in line with our exploratory analyses on the Prospect Theory model (which was not the best fitting model for our data), where we also found a trend for amisulpride to reduce probability distortion (Figure S2). Our linear mixed modeling effects and the non-significant shift of ω_P in our hierarchical Bayesian model diverge from these previous results. This might be caused by our task design: First, our reward-guided decision-making task consisted exclusively of a gain frame, without a loss frame. Risk preferences, and especially loss aversion, might depend on losing instead of just not receiving a reward. However, previous studies that have also exclusively included a gain frame found dopaminergic effects that were equivalent to studies including both a gain and loss frame [7, 40]. Thus, this factor is unlikely to be the reason for our diverging effects. Second, the short offset in our task design between the presentation of the first and second choice option might have might have modulated the results by imposing a working memory component. According to the two-state model suggested by Seamans and Yang [46, 47], dopamine can switch networks in prefrontal cortex (PFC) between two different states. In state 1, dominated by D_2-receptor activation, multiple working memory representations can be held in PFC networks nearly simultaneously, but are vulnerable to disruption. Conversely, during state 2, dominated by activation of D_1-receptors, only particularly strong inputs can access PFC networks, which are then relatively robust to disruption. Therefore, state

1 allows exploration of the input space before state 2 focuses on a limited sets of representations leading to enhanced robustness of working memory representations. In our task, a dopaminergic effect on working memory should be mirrored especially in the first choice option, which needs to be maintained in working memory until the presentation of the second option. But testing the drug effects separating the value parameters for the two options did not reveal option-specific drug effects (only the interaction effect between L-DOPA and reward magnitude was no longer significant but only a trend, which likely is due to the reduced statistical power of this test). Still, we cannot rule out the possibility that the working memory delay might have influenced participants' behavior. Third, similar pharmacological studies have found no effects of (indirect) dopamine agonists or D_2-receptor antagonists on risk preferences [48-53].

In conclusion, participants' behavior was best described by a hybrid model where additive and multiplicative decision strategies jointly contributed to subjective value. Further, our data suggest that dopamine does not affect the relative dominance of either of the two decision strategies. However, the degree to which choices were controlled by individual attributes, was increased by L-DOPA and decreased by amisulpride. Since we did not find evidence for a stronger shift of one option attribute compared to the other, we would argue that this cannot be interpreted as change in risk preferences. Our data thus provide evidence for a role of dopamine in controlling the influence of value parameters on choice, irrespective of decision strategies and risk preferences.

Acknowledgements

We thank the Department of Neurology at the University Hospital of the Otto-von-Guericke University and the chair of the Department of Neurology, Prof. Dr. Hans-Jochen Heinze for invaluable support. Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf.

Author contributions

- Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work (AADM, TOJG, MIF, HK, LFK, GJ)
- Drafting the work or revising it critically for important intellectual content (AADM, MIF, GJ)
- Final approval of the version to be published (AADM, TOJG, MIF, HK, LFK, GJ)
- Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved (AADM, TOJG, MIF, HK, LFK, GJ)

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG JO 787/5-1), a grant from the Federal State of Saxony Anhalt and the European Regional Development Fund (ERDF 2014-2020) to GJ, project: Center for Behavioral Brain Sciences (FKZ: ZS/2016/04/78113), and a travel grant from the G.-A.-Lienert Foundation to TOJG.

Competing Interests

The authors have nothing to disclose.

References

1. Cherkasova, M.V., et al., *Dopamine replacement remediates risk aversion in Parkinson's disease in a value-independent manner.* Parkinsonism Relat Disord, 2019. **66**: p. 189-194.
2. Kobayashi, S., et al., *Dopaminergic influences on risk preferences of Parkinson's disease patients.* Cogn Affect Behav Neurosci, 2019. **19**(1): p. 88-97.
3. Connolly, B.S. and A.E. Lang, *Pharmacological Treatment of Parkinson Disease A Review.* Jama-Journal of the American Medical Association, 2014. **311**(16): p. 1670-1683.
4. Molina, J.A., et al., *Pathologic gambling in Parkinson's disease: a behavioral manifestation of pharmacologic treatment?* Mov Disord, 2000. **15**(5): p. 869-72.
5. Evans, A.H., et al., *Impulsive and compulsive behaviors in Parkinson's disease.* Mov Disord, 2009. **24**(11): p. 1561-70.
6. Weintraub, D., et al., *Association of dopamine agonist use with impulse control disorders in Parkinson disease.* Archives of Neurology, 2006. **63**(7): p. 969-973.
7. Rutledge, R.B., et al., *Dopaminergic Modulation of Decision Making and Subjective Well-Being.* J Neurosci, 2015. **35**(27): p. 9811-22.
8. Burke, C.J., et al., *Dopamine Receptor-Specific Contributions to the Computation of Value.* Neuropsychopharmacology, 2018. **43**(6): p. 1415-1424.
9. Tversky, A. and D. Kahneman, *The Framing of Decisions and the Evaluation of Prospects*, in *Studies in Logic and the Foundations of Mathematics*, R. Barcan Marcus, G.J.W. Dorn, and P. Weingartner, Editors. 1986, Elsevier. p. 503-520.
10. Kahneman, D. and A. Tversky, *Prospect theory: An analysis of decision under risk*, in *Handbook of the fundamentals of financial decision making: Part I*. 2013, World Scientific. p. 99-127.
11. Farashahi, S., et al., *Flexible combination of reward information across primates.* Nat Hum Behav, 2019. **3**(11): p. 1215-1224.
12. Figner, B. and N. Voelki, *Risky Decision Making in a Computer Card Game: An Information Integration Experiment.* Pol. Psychol. Bull., 2004. **35**.
13. Scholl, J., et al., *A role beyond learning for NMDA receptors in reward-based decision-making-a pharmacological study using d-cycloserine.* Neuropsychopharmacology, 2014. **39**(12): p. 2900-9.
14. Becker, J.B., T.E. Robinson, and K.A. Lorenz, *Sex differences and estrous cycle variations in amphetamine-elicited rotational behavior.* Eur J Pharmacol, 1982. **80**(1): p. 65-72.
15. Dreher, J.C., et al., *Menstrual cycle phase modulates reward-related neural function in women.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2465-70.
16. Creutz, L.M. and M.F. Kritzer, *Mesostriatal and mesolimbic projections of midbrain neurons immunoreactive for estrogen receptor beta or androgen receptors in rats.* J Comp Neurol, 2004. **476**(4): p. 348-62.
17. Jocham, G., T.A. Klein, and M. Ullsperger, *Differential modulation of reinforcement learning by D2 dopamine and NMDA glutamate receptor antagonism.* J Neurosci, 2014. **34**(39): p. 13151-62.
18. Jocham, G., T.A. Klein, and M. Ullsperger, *Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices.* J Neurosci, 2011. **31**(5): p. 1606-13.
19. Le Bricon, C., et al., *Pharmacokinetics of amisulpride and its enantiomers after single and repeated doses in healthy volunteers.* Eur J Drug Metab Pharmacokinet (special issue), 1996: p. 135-136.

20. Nutt, J.G., *Pharmacokinetics and pharmacodynamics of levodopa.* Mov Disord, 2008. **23 Suppl 3**: p. S580-4.
21. Bond, A. and M. Lader, *The use of analogue scales in rating subjective feelings.* British Journal of Medical Psychology, 1974.
22. Reitan, R.M., *Validity of the Trail Making Test as an indicator of organic brain damage.* Perceptual and motor skills, 1958. **8**(3): p. 271-276.
23. Jocham, G., et al., *A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex.* Nat Neurosci, 2012. **15**(7): p. 960-1.
24. Jocham, G., et al., *Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice.* Neuroimage, 2014. **100**: p. 498-506.
25. Hunt, L.T., R.J. Dolan, and T.E. Behrens, *Hierarchical competitions subserving multi-attribute choice.* Nat Neurosci, 2014. **17**(11): p. 1613-22.
26. Hunt, L.T., et al., *Trial-type dependent frames of reference for value comparison.* PLoS Comput Biol, 2013. **9**(9): p. e1003225.
27. R Development Core Team, *R: A Language and Environment for Statistical Computing*. 2022, R Foundation for Statistical Computing: Vienna, Austria.
28. Bates, D., et al., *Fitting Linear Mixed-Effects Models Using lme4.* Journal of Statistical Software, 2015. **67**(1): p. 1-48.
29. Brown, V.A., *An Introduction to Linear Mixed-Effects Modeling in R.* Advances in Methods and Practices in Psychological Science, 2021. **4**(1).
30. Harrison, X.A., et al., *A brief introduction to mixed effects modelling and multi-model inference in ecology.* PeerJ, 2018. **6**: p. e4794.
31. Jaeger, T.F., *Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models.* J Mem Lang, 2008. **59**(4): p. 434-446.
32. Barr, D.J., et al., *Random effects structure for confirmatory hypothesis testing: Keep it maximal.* J Mem Lang, 2013. **68**(3).
33. Swart, J.C., et al., *Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action.* Elife, 2017. **6**.
34. Johnson, S.G., *The NLopt nonlinear-optimization package*.
35. Powell, M.J., *The BOBYQA algorithm for bound constrained optimization without derivatives.* Cambridge NA Report NA2009/06, University of Cambridge, Cambridge, 2009. **26**.
36. Kurtenbach, H., et al., *The Role of Acetylcholine in Reward-Guided Decision Making Under Different Degrees of Uncertainty*. 2024.
37. Stan Development Team, *RStan: the R interface to Stan.* 2023.
38. Pape, A.A. and M. Siegel, *Motor cortex activity predicts response alternation during sensorimotor decisions.* Nat Commun, 2016. **7**: p. 13098.
39. Rogge, J., G. Jocham, and M. Ullsperger, *Motor cortical signals reflecting decision making and action preparation.* Neuroimage, 2022. **263**: p. 119667.
40. Ojala, K.E., et al., *Dopaminergic Drug Effects on Probability Weighting during Risky Decision Making.* Eneuro, 2018. **5**(2).
41. Rigoli, F., et al., *Dopamine Increases a Value-Independent Gambling Propensity.* Neuropsychopharmacology, 2016. **41**(11): p. 2658-67.
42. Gross, J., et al., *When Helping Is Risky: The Behavioral and Neurobiological Trade-off of Social and Risk Preferences.* Psychol Sci, 2021. **32**(11): p. 1842-1855.
43. White, T.L., C.W. Lejuez, and H. de Wit, *Personality and gender differences in effects of d-amphetamine on risk taking.* Exp Clin Psychopharmacol, 2007. **15**(6): p. 599-609.
44. Riba, J., et al., *Dopamine agonist increases risk taking but blunts reward-related brain activity.* PLoS One, 2008. **3**(6): p. e2479.

45. Petzold, J., et al., *Baseline impulsivity may moderate L-DOPA effects on value-based decision-making.* Scientific Reports, 2019. **9**.

46. Seamans, J.K., et al., *Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons.* J Neurosci, 2001. **21**(10): p. 3628-38.

47. Seamans, J.K. and C.R. Yang, *The principal features and mechanisms of dopamine modulation in the prefrontal cortex (vol 74, pg 1, 2004).* Progress in Neurobiology, 2004. **74**(5): p. 321-321.

48. Arrondo, G., et al., *Dopaminergic modulation of the trade-off between probability and time in economic decision-making.* Eur Neuropsychopharmacol, 2015. **25**(6): p. 817-27.

49. Zack, M. and C.X. Poulos, *A D2 antagonist enhances the rewarding and priming effects of a gambling episode in pathological gamblers.* Neuropsychopharmacology, 2007. **32**(8): p. 1678-86.

50. Symmonds, M., et al., *Assaying the effect of levodopa on the evaluation of risk in healthy humans.* PLoS One, 2013. **8**(7): p. e68177.

51. Evers, E.A., P. Stiers, and J.G. Ramaekers, *High reward expectancy during methylphenidate depresses the dopaminergic response to gain and loss.* Soc Cogn Affect Neurosci, 2017. **12**(2): p. 311-318.

52. Acheson, A. and H. de Wit, *Bupropion improves attention but does not affect impulsive behavior in healthy young adults.* Exp Clin Psychopharmacol, 2008. **16**(2): p. 113-23.

53. Hamidovic, A., U.J. Kang, and H. de Wit, *Effects of low to moderate acute doses of pramipexole on impulsivity and cognition in healthy volunteers.* J Clin Psychopharmacol, 2008. **28**(1): p. 45-51.

Figure legends

Figure 1. Schematic of the decision-making task and participants' behavior. (A) Task schematic and trial timeline. Each choice option was associated with a reward magnitude (width of the horizontal bar) and a reward probability (percentage). The left option was always presented first, followed by a short delay before the presentation of the right option. Participants were instructed to only respond once the question mark appeared. Lastly, they were given feedback on the outcome (reward/no reward) of both the chosen and the unchosen option and reward, if obtained, was added to the progress bar (blue bar at the bottom of the screen). (B) Probability of choosing the choice option with a higher expected value (EV) depending on the absolute EV difference between the two options. (C) Probability of choosing the option with higher reward magnitude depending on the absolute magnitude difference between the two options. (D) Probability of choosing the option with a higher reward probability depending on the absolute probability difference between the two options. In B-D, solid lines represent the average, shaded areas represent SEM across participants**.**

Figure 2. General linear mixed model results. The plots indicate the modelled probability to select the right option as a function of reward magnitude (A, C) or reward probability (B, D) difference between right minus left option (z-scored difference values). Amisulpride significantly decreased the influence of both reward magnitude (A) and reward probability (B) on choice. In contrast, L-DOPA significantly increased the influence of both reward magnitude (C) and reward probability (D) on choice. Solid lines represent the average, shaded areas represent SEM across participants.

Figure 3. Posterior distributions of the group level estimate of the parameters from the hierarchical Bayesian Model. Panels on the left (A, C, E) are the posterior distributions of the three model parameters, the multiplicative weight w_{mult}, the relative probability

weight w_P and the softmax inverse temperature τ. Panels on the right (B, D, F) are the posterior distributions of the shift parameters, color-coded for each drug. Shaded areas in the distributions are the 95%-CI. Dots represent single-subject estimates. (A) Participants use a hybrid response strategy comprising of both a multiplicative and an additive decision strategy, as evident from median ω_{mult} = 0.49. (B) This response strategy is not significantly affected by amisulpride or L-DOPA (95%-CI overlaps with zero). (C) Participants choices are guided more by reward probabilities than reward magnitudes, as evident from w_P higher than 0.5 (median w_P = 0.82). (D) This relative weighting is not significantly shifted by amisulpride or L-DOPA (95%-CI overlaps with zero). (E) The softmax inverse temperature τ is not significantly shifted by amisulpride or L-DOPA (F), since the 95%-CI overlaps with zero.

Table 1. Model fits for the different computational models. Bayes Information Criterion (BIC) was used to compare model fits. Lower BIC indicates better fit. The hybrid model without a bias is the best fitting model.