



Motivational learning biases are differentially modulated by genetic determinants of striatal and prefrontal dopamine function

Anni Richter¹ · Lieke de Boer^{2,3} · Marc Guitart-Masip^{2,4} · Gusalija Behnisch¹ · Constanze I. Seidenbecher^{1,5} · Björn H. Schott^{1,5,6,7,8}

Received: 29 April 2021 / Accepted: 4 July 2021 / Published online: 24 July 2021
© The Author(s) 2021, corrected publication 2021

Abstract

Dopaminergic neurotransmission plays a pivotal role in appetitively motivated behavior in mammals, including humans. Notably, action and valence are not independent in motivated tasks, and it is particularly difficult for humans to learn the inhibition of an action to obtain a reward. We have previously observed that the carriers of the DRD2/ANKK1 TaqIA A1 allele, that has been associated with reduced striatal dopamine D2 receptor expression, showed a diminished learning performance when required to learn response inhibition to obtain rewards, a finding that was replicated in two independent cohorts. With our present study, we followed two aims: first, we aimed to replicate our finding on the DRD2/ANKK1 TaqIA polymorphism in a third independent cohort ($N=99$) and to investigate the nature of the genetic effects more closely using trial-by-trial behavioral analysis and computational modeling in the combined dataset ($N=281$). Second, we aimed to assess a potentially modulatory role of prefrontal dopamine availability, using the widely studied COMT Val108/158Met polymorphism as a proxy. We first report a replication of the above mentioned finding. Interestingly, after combining all three cohorts, exploratory analyses regarding the COMT Val108/158Met polymorphism suggest that homozygotes for the Met allele, which has been linked to higher prefrontal dopaminergic tone, show a lower learning bias. Our results corroborate the importance of genetic variability of the dopaminergic system in individual learning differences of action–valence interaction and, furthermore, suggest that motivational learning biases are differentially modulated by genetic determinants of striatal and prefrontal dopamine function.

Keywords Dopamine D2 receptor · TaqIA · COMT · Reward learning · Motivated learning · Action bias

Introduction

The impact of motivation on cognitive functions has been subject to intense investigation over the past 2 decades. While the influence of motivational salience on cognitive processes and goal-directed behavior is common knowledge nowadays, theories of instrumental learning have

until recently neglected the influence of outcome valence on action initiation. Two logically assumed independent axes of behavioral control, namely a valence axis running from reward to punishment, and an action axis running from vigor to inhibition, have been shown to interact (Guitart-Masip et al. 2012). To study this phenomenon, a *go/no-go* task was developed that independently dissociates, i.e.

✉ Anni Richter
anni.richter@lin-magdeburg.de

¹ Department of Behavioral Neurology, Leibniz Institute for Neurobiology, Brennekestr. 6, 39118 Magdeburg, Germany

² Ageing Research Centre, Karolinska Institute, Stockholm, Sweden

³ Present Address: Max Planck Institute for Human Development, Center for Lifespan Psychology, Berlin, Germany

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

⁵ Center for Behavioral Brain Sciences, Magdeburg, Germany

⁶ Department of Psychiatry and Psychotherapy, University Medicine Göttingen, Göttingen, Germany

⁷ Department of Neurology, University of Magdeburg, Magdeburg, Germany

⁸ German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany

orthogonalizes, action and valence, which includes the four conditions: *go to win*, *go to avoid losing*, *no-go to win*, and *no-go to avoid losing*. If the two axes of behavioral control, action and valence, would be independent, all conditions should be learned equally well. However, biased behavior, that is, an interaction of action and valence is observed, and the larger the bias the higher the coupling of action and valence, such that signals that predict reward are prepotently associated with behavioral activation, whereas signals that predict punishment are intrinsically coupled to behavioral inhibition. This finding has been robustly replicated in multiple studies (Guitart-Masip et al. 2012, 2014; Cavanagh et al. 2013; Chowdhury et al. 2013; Richter et al. 2014; de Berker et al. 2016; Swart et al. 2017, 2018; de Boer et al. 2019; Dorfman and Gershman 2019; Betts et al. 2020; Kuhnel et al. 2020; Perosa et al. 2020; van Nuland et al. 2020; Ereira et al. 2021). Understanding the neurocognitive mechanisms underlying this behavioral bias is thus important for developing more comprehensive theories of instrumental learning.

Numerous studies in a multitude of species, including humans, indicate the importance of dopamine (DA) in the neural manifestation of motivated behavior. According to a prevalent view in reinforcement learning and decision making, DA neurons signal reward prediction errors (Montague et al. 1996; Schultz et al. 1997; Bayer and Glimcher, 2005), in the form of phasic bursts for positive prediction errors and dips below baseline firing rate for negative prediction errors (Bayer et al. 2007), resulting in corresponding peaks and dips of DA availability in target structures, most prominently the striatum (McClure et al. 2003; O'Doherty et al. 2003, 2004; Pessiglione et al. 2006). In the striatum, increased DA release in response to an unexpected reward reinforces the direct pathway via activation of D1 receptors and thereby facilitates the future generation of *go* choices under similar circumstances, while dips in DA levels in response to an unexpected punishment reinforce the indirect pathway via reduced activation of D2 receptors, thereby facilitating the subsequent generation of *no-go* choices in comparable situations (Frank et al. 2004, 2007; Wickens et al. 2007; Hikida et al. 2010).

As the human dopaminergic system is subject to considerable genetic variability, several polymorphisms that have been associated with alterations in dopaminergic gene products (e.g., DRD2, COMT, DAT, and DARPP-32; see supplementary Figure S1) have been used to study naturally occurring differences in the dopaminergic system of healthy subjects. In line with the assumptions outlined above, we observed in a previous study (Richter et al. 2014) that the coupling of action and valence during learning was modulated by a genetic variant linked to striatal DA D2 receptor expression. We argued that A1 carriers with presumably less D2 receptors would be assumed to have less limitation of dopaminergic signaling

after negative prediction errors in the indirect pathway and a shift to a more action-oriented behavioral pattern mediated by the direct pathway (see Fig. 4). In line with that framework, in a recent study, de Boer et al. (2019) found a positive correlation between the strength of the action by valence interaction and dorsal striatal D1 receptor availability measured using positron emission tomography (PET). Therefore, striatal dopaminergic effects may be sufficient to explain biased motivational learning (Swart et al. 2017; de Boer et al. 2019). On the other hand, Guitart-Masip et al. (2014) observed that levodopa administration led to a reduced coupling of action and valence that cannot be explained by striatal action of DA. The authors attributed their observation to an effect on prefrontal cortex (PFC) functioning, where DA plays a role in facilitating working memory and attentional processes (Seamans and Yang, 2004; Hitchcott et al. 2007; Haber and Knutson, 2010) that may help to overcome the biased behavior. This effect of levodopa administration was recently replicated in patients with non-tremor Parkinson's disease (van Nuland et al. 2020), and studies investigating frontal network dynamics using electroencephalography further demonstrate that prefrontal control processes (as indexed by higher mid-frontal theta power) are important to overcome biased behavior (Cavanagh et al. 2013; Swart et al. 2018). Therefore, DA may influence these learning biases in a regionally specific manner.

Numerous previous studies have investigated the influence of candidate single-nucleotide polymorphisms (SNPs) of DA on instrumental learning (Frank et al. 2007; Klein et al. 2007; Frank & Hutchison, 2009; Jochem et al. 2009; Corral-Frias et al. 2016). As the expression of several key molecules of the dopaminergic system shows a characteristic regional distribution in the brain, genetically mediated differences may also provide some information about the contributions of different brain regions to DA-dependent learning and memory processes (Schott et al. 2006; Mier et al. 2010; Corral-Frias et al. 2016). In the current study, we aimed to examine differential contributions of two dopaminergic SNPs: the DRD2/ANKK1 TaqIA SNP (rs1800497) and the COMT Val108/158Met SNP (rs4680).

In PET studies, the DRD2/ANKK1 TaqIA polymorphism has repeatedly been linked to lower striatal D2 binding availability in carriers of the less common A1 allele (for review and meta-analysis, see Gluskin and Mickey 2016; Eisenstein et al. 2016). With respect to motivated behavior, Stice et al. (2012) found stronger midbrain activation in A1 carriers compared with A2 homozygotes on reward expectancy, and Stelzel et al. (2010) reported generally increased striatal BOLD signaling in A1 carriers. In addition, relative to A2 homozygotes, A1 carriers showed poorer performance in avoiding actions associated with punishment and lower activations of PFC and striatum during processing of negative

feedback (Klein et al. 2007; Frank and Hutchison, 2009; Jocham et al. 2009).

Furthermore, there is evidence of associations of the A1 allele with psychiatric disorders such as addictions—most notably alcohol dependence (for a meta-analysis, see Wang et al. 2013; for reviews, see Samochowiec et al. 2014 and Koeneke et al. 2020)—and ADHD (for a meta-analysis, see Pan et al. 2015). In addition, it was initially hypothesized that there was an advantage of the A1 allele in schizophrenia disorders in terms of lower risk (Dubertret et al. 2004) and better response to haloperidol (Schafer et al. 2001). However, while a meta-analysis (Yao et al. 2015) failed to confirm a significant association between schizophrenia and the TaqIA polymorphism, an association with another DRD2 SNP was reaffirmed, and findings from a genome-wide association study also support the relevance of DRD2 polymorphisms in schizophrenia disorders (Schizophrenia Working Group of the Psychiatric Genomics 2014).

Moreover, behavioral experiments and questionnaire studies have been able to show associations between the A1 allele and higher scores on the personality traits reward dependence, impulsivity, curiosity (novelty seeking), and extraversion (Noble et al. 1998; Eisenberg et al. 2007; Lee et al. 2007; Smillie et al. 2010).

Catechol-O-methyltransferase (COMT) plays a key role in the breakdown of DA in the PFC (Kaenmaki et al. 2010; Schott et al. 2010), whereas its role in striatal DA inactivation has been shown to be of lesser importance (Yavich et al. 2007; Korn et al. 2021). The frequent Val108/158Met SNP in the *COMT* gene (chromosome 22) leads to an amino acid exchange from valine (Val) to methionine (Met). In Met carriers, reduced enzymatic activity and increased prefrontal DA availability have been observed, presumably due to lower thermostability of the enzyme (Chen et al. 2004). This SNP has mainly been investigated with respect to PFC-dependent executive functions (for reviews, see Frank and Fossella, 2011; Klanker et al. 2013), and a meta-analysis of functional magnetic resonance imaging (fMRI) studies confirmed that Met carriers show more efficient performance in executive functions and higher neural activations during emotion processing (Mier et al. 2010). In the context of motivated behavior, the Met allele has been associated with more successful reward learning (for a meta-analysis, see Corral-Frias et al. 2016). Moreover, Met allele carriers adapt behavior more rapidly on a trial-to-trial basis during reinforcement learning (Frank et al. 2007; Frank and Hutchison 2009).

We have previously shown in two independent cohorts that carriers of the A1 allele of the DRD2/ANKK1 TaqIA polymorphism show a rather selective deficit in learning to inhibit an action to receive a reward (Richter et al. 2014). With our present study, we followed two aims: first, we

aimed to replicate our finding on the TaqIA polymorphism in a third independent cohort and to investigate the nature of the genetic effects more closely using trial-by-trial behavioral analysis and computational modeling in the combined dataset ($N=281$). Second, we aimed to assess a potentially modulatory role of prefrontal DA availability, using the widely studied COMT Val108/158Met polymorphism as a proxy. Regarding the DRD2/ANKK1 TaqIA SNP, we hypothesized that, in line with our previous observations (Richter et al. 2014), A1 carriers would show a higher coupling of action and valence. With respect to the COMT polymorphism, we hypothesized that, given the preferential role of COMT in PFC versus striatal DA availability, carriers of the low-activity Met allele would more readily overcome the learning bias and show less coupling of valence with action.

Materials and methods

Participants

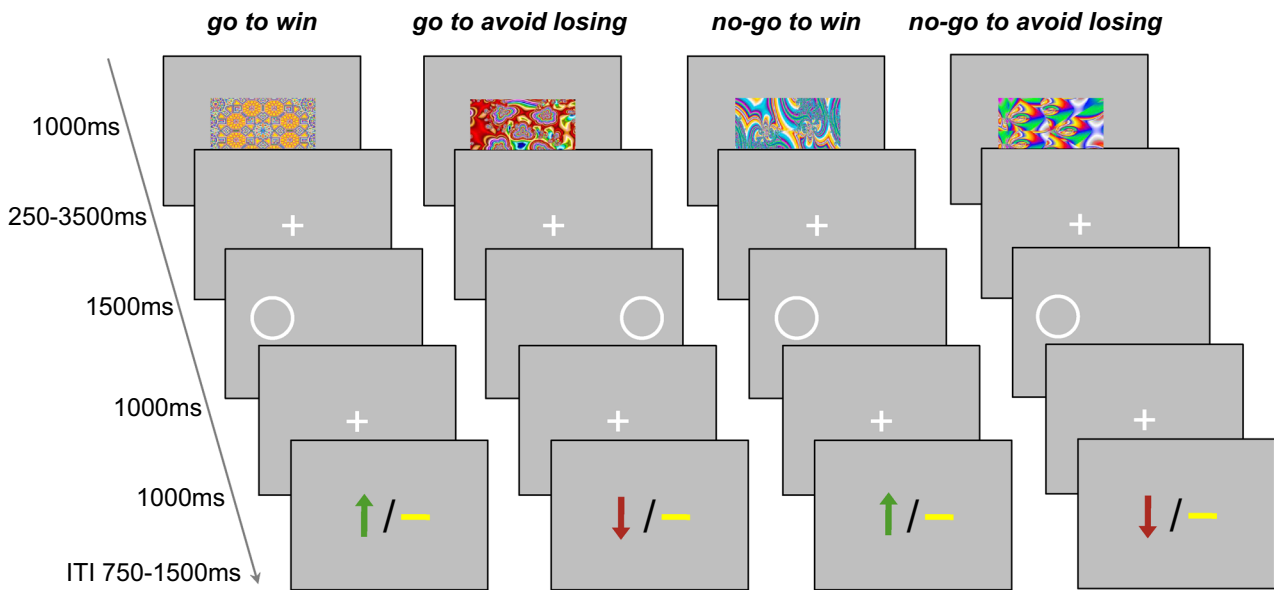
In addition to our previously described two cohorts of 87 and 95 participants (Richter et al. 2014), 99 newly recruited participants were tested (55 females and 44 males; age: range 20–34 years, mean 25.2 years, $SD=2.6$ years; demographic description of all three samples in Supplementary Table S1). According to self-report, all participants were of European ethnicity, right-handed, had obtained at least a university entrance diploma (Abitur) as educational certificate, had no present or past neurological or mental disorder, alcohol or drug abuse, did not use centrally acting medication, and had no history of psychosis or bipolar disorder in a first-degree relative. Additionally, given the design of the experiment, regularly gambling was defined as an exclusion criterion for participation.

All participants gave written informed consent in accordance with the Declaration of Helsinki and received financial compensation for participation. The study was approved by the Ethics Committee of the Faculty of Medicine at the Otto von Guericke University of Magdeburg.

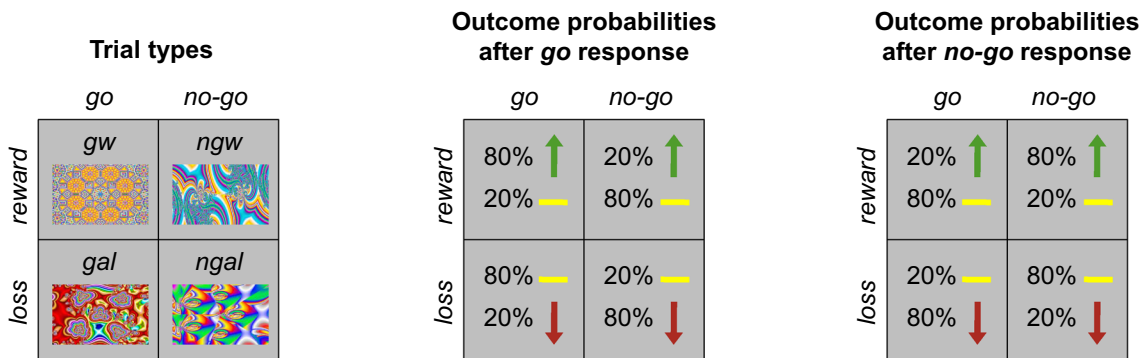
Genotyping

Genomic DNA was extracted from blood leukocytes using the KingFisher™ Duo Prime Purification System (Thermo Scientific™) according to the manufacturer's protocol. Genotyping of the SNPs DRD2/ANKK1 TaqIA (NCBI accession number: rs1800497) and COMT Val108/158Met (rs4680) was performed using PCR-based restriction fragment length analysis according to previously described protocols (Schott et al. 2006; Wimber et al. 2011; Richter et al. 2013, 2014,

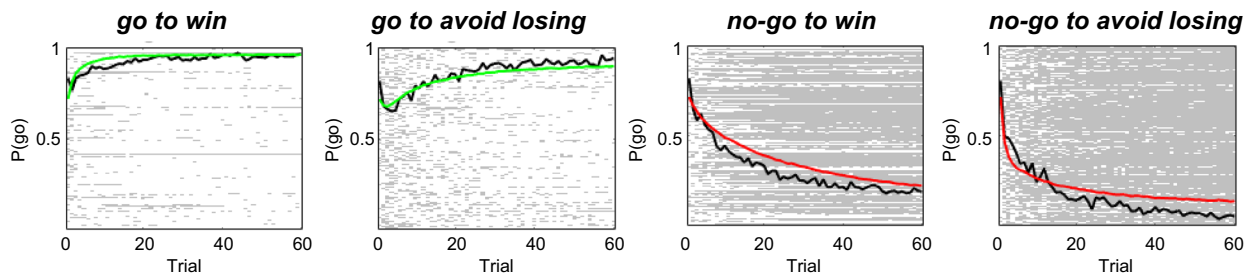
(A) Trial Timeline



(B) Stimulus characteristics



(C) Simulated choice data



2017). A1 carriers of the TaqIA SNP were grouped together (A1 + : A1/A1 and A1/A2; A1 - : A2/A2) as in the previous studies (Klein et al. 2007; Frank and Hutchison, 2009; Jocham et al. 2009; Stelzel et al. 2010; Stice et al. 2012; Richter et al. 2013, 2014, 2017).

Paradigm

We used a previously employed go/no-go learning task with orthogonalized action requirements and outcome valence (Guitart-Masip et al. 2012). Detailed descriptions of the task have been presented previously (Richter et al. 2014; Betts

Fig. 1 Experimental paradigm and participant performance. **A** Probabilistic monetary *go/no-go* task. Fractal cues indicate the condition—a combination of action (*go* or *no-go*) and valence (*reward* or *punishment*). On *go* trials, subjects press a button for the side of a circle. On *no-go* trials, they withhold a response. Arrows indicate *rewards* (upward) or *punishments* (downward). Horizontal bars symbolize the absence of a *reward* or *punishment*. ITI, intertrial interval. **B** The schematics represent for each condition the nomenclature (left), the possible outcomes and their probabilities after a *go* response (middle), and the possible outcomes and their probability after a *no-go* response (right). **C** Simulated choice data according to the model parameters of the winning model. Colored lines represent the simulated group mean probability of performing a *go* on each trial (green for *go* conditions, where *go* is the correct response; red for *no-go* conditions, where *no-go* is the correct response). Black lines indicate the group mean for participants' actual *go* responses on each trial. In the plot area, each row represents one participant's choice behavior for each trial (281 × 60 pixels). A white pixel reflects that a participant chose *go* on that trial; a gray pixel represents *no-go*. Participants made more *go* responses to *win* vs. *avoid losing* cues, reflecting the motivational bias. Overall, they successfully learned whether to make a *go* response or not (proportion of *go* responses increases for *go* cues and decreases for *no-go* cues). Figures (A) and (B) adapted from Richter et al. (2014)

et al. 2020). Figure 1A displays the trial timeline. Briefly, each trial consisted of the presentation of a fractal cue, a target detection task, and a probabilistic outcome. First, one out of four abstract fractal cues was displayed. Prior to the beginning of the task, participants were informed that a fractal indicated i) whether they would subsequently be required to perform a target detection task by pressing a button (*go*) or not (*no-go*) and ii) the possible valence of the outcome of the subjects' behavior (reward/no reward or punishment/no punishment). Importantly, subjects were not instructed with respect to the contingencies of each fractal image and had to learn them by trial and error. There were four trial types: press the correct button in the target detection task to gain a reward of 0.50 € [*“go to win”* (*gw*)]; press the correct button to avoid a punishment of −0.50 € [*“go to avoid losing”* (*gal*)]; do not press a button to gain a reward [*“no-go to win”* (*ngw*)]; do not press a button to avoid punishment [*“no-go to avoid losing”* (*ngal*)]. The outcome was probabilistic (see Fig. 1B). To avoid incidental effects of specific cue images, the association of the fractal images with the specific conditions (*go* vs. *no-go** reward vs. punishment) was randomized across participants. The task included 240 trials (60 trials per condition) and was divided into four sessions. Subjects were told that they would be paid their earnings of the task up to a total of 25 € and a minimum of 7 €. Before starting the actual learning task, subjects performed 10 trials of the target detection task to familiarize themselves with the speed requirements.

Statistical analysis

Accuracy was analyzed using IBM® SPSS® Statistics version 21. The percentage of correct choices in the target detection task (button press in *go* trials and omission of responses in *no-go* trials) was collapsed across time bins of 30 trials per condition. To assess the learning enhancement, the slope was calculated by subtracting the mean values in the first half of the experiment from the mean values of the second half of the experiment (slope = mean [2nd half] − mean [1st half]).

For the replication of our previous study (Richter et al. 2014) in the new cohort ($N=99$), we compared DRD2/ANKK1 TaqIA genotype groups with a *t* test for independent samples and investigated task effects with a mixed analysis of variance (ANOVA) with time (1st/2nd half), action (*go/no-go*), and valence (*win/avoid losing*) as within-subject factors.

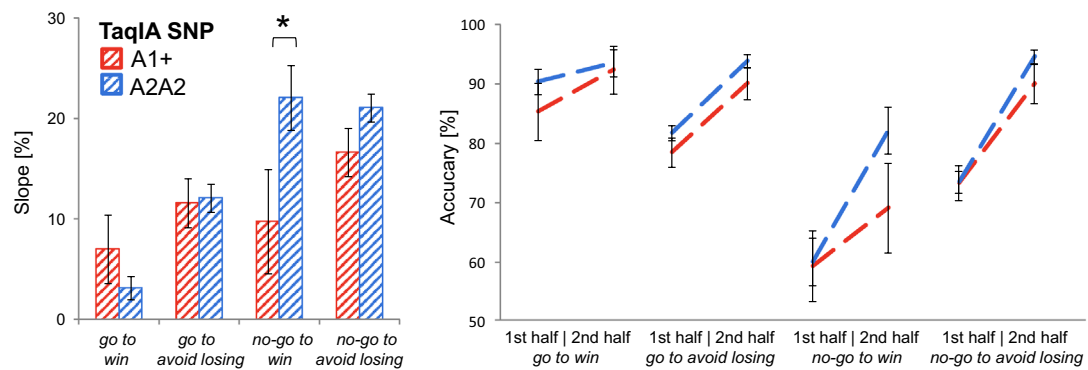
Then, by combining all three datasets ($N=281$), we included the two genotypes as between-subject factors in the analysis and added cohort (three cohorts represented in two dichotomous dummy coded variables for cohort 2 and 3), and age and gender as covariates (analysis of covariance, ANCOVA). The increased number of participants allowed us to run a logistic regression on the trial-by-trial *go* responses as in Swart et al. (2017) which more accurately analyzes the data, as it is closer to the actual behavior of each participant by including inter- and intraindividual variability (see supplementary methods for details).

Unless stated otherwise, independent samples *t* tests were used as post hoc tests, and the significance threshold was set to 0.05, two-tailed. Whenever Levene's test was significant, statistics were adjusted, but for better readability, uncorrected degrees of freedom are reported.

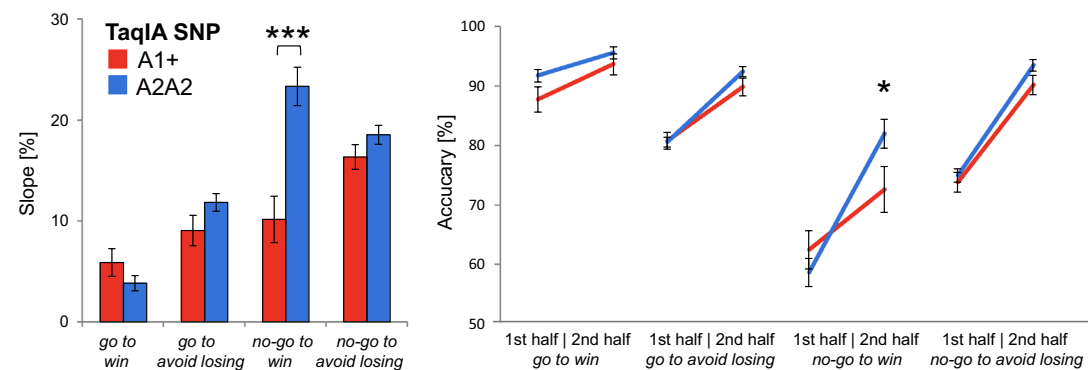
Computational modeling of task performance

Computational modeling of task performance was employed using MATLAB® R2016B (Mathworks®). We used a previously published modeling procedure (Huys et al. 2011; Guitart-Masip et al. 2012). Detailed descriptions of the reinforcement learning models as well as the model fitting procedure and comparison have been described in a recent study of age effects in the same task (Betts et al. 2020). Briefly, we constructed six nested reinforcement learning models to fit participants' behavior (Table 2). The base model was a Q-learning algorithm (Sutton and Barto 1998) that used a Rescorla–Wagner update rule to independently track the action value of each choice (*go*; *no go*), given each fractal image, with a learning rate (ϵ) as a free parameter. In this model, the probability of choosing one action on a trial was

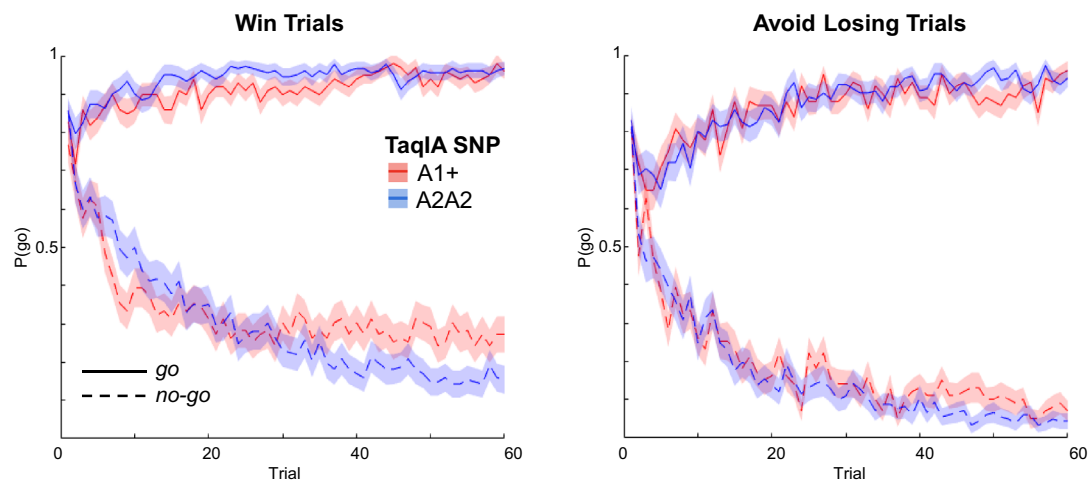
(A) TaqIA SNP effects on accuracy in the third cohort (N=99)



(B) TaqIA SNP effects on accuracy in the entire sample (N=281)



(C) TaqIA SNP effects on trial-by-trial behavior in the entire sample (N=281)



a sigmoid function of the difference between the action values scaled by a slope parameter that was parameterized as sensitivity to reward (ρ). This basic model was augmented with an irreducible noise parameter (ξ) and then further

expanded by adding a static bias parameter to the value of the *go* action (b). Furthermore, we allowed for separate sensitivities to rewards (ρ_{win}) and punishments (ρ_{lose}). As in our recent study of age effects (Betts et al. 2020), the model

Fig. 2 Effects of DRD2/ANKK1 TaqIA genotype on choice performance. **A** and **B** Effects of DRD2/ANKK1 TaqIA genotype on choice performance in the third cohort ($N=99$) and in the entire sample ($N=281$). Compared to the A2 homozygotes, A1 carriers showed a diminished learning to withhold an action to receive a reward. Left panels: bar plots show mean differences between correct response rates (\pm SEM) during second half versus the first half of trials for each condition. This score represents the observed fourfold interaction of *action* \times *valence* \times *time* \times *genotype*. Right panels: line charts show mean values of correct responses (\pm SEM) in the first and the second half of trials for all four conditions. Post hoc comparisons via t tests: $*p < 0.05$, $***p < 0.001$. **C** Trial-by-trial proportions of *go* responses (\pm SEM) to *go* cues (solid lines) and *no-go* cues (dashed lines) across cue types. *Win* and *avoid losing* condition separately and colors depict DRD2/ANKK1 TaqIA genotypes. TaqIA A1 carriers showed an enhanced effect of cue valence on *go* responding especially in the *no-go to win* condition with further progress of the experiment (lines are mostly separated). Adapted scripts of Swart et al. (2017) were used to generate figures

was then extended by adding a constant Pavlovian value of 1 or -1 to the value of the *go* action as soon as the first reward for *win* cues or the first punishment for *avoid losing* cues, respectively, was encountered. This fixed Pavlovian value was weighted by a further free parameter (Pavlovian parameter) into the value of the *go* action (π). Model comparisons demonstrated a better fit compared to a variable Pavlovian value used in the previous studies (Guitart-Masip et al. 2012; Cavanagh et al. 2013; de Boer et al. 2019) (see Table 2). As in the previous reports (Huys et al. 2011; Guitart-Masip et al. 2012), we employed a hierarchical Type II Bayesian procedure using maximum likelihood to fit simple parameterized distributions for higher level statistics of the parameters. All six computational models were fit to the data using a single distribution for all participants. This fitting procedure was, therefore, blind to the existence of different genotype groups with putatively different parameter values. Models were compared using the integrated Bayesian Information Criterion (iBIC) with small iBIC values indicating a model that fits the data better after penalizing for the number of data points associated with each parameter. Finally, we assessed genotype-related effects on all modeling parameters using IBM® SPSS® Statistics version 21. To test for differences regarding specific model parameters, we calculated t tests for independent samples. As one could not exclude that not one specific parameter but a combination of them differed between genotypes, we performed a multivariate test of differences—a linear discriminant analysis (LDA). The purpose of LDA was to find a linear combination of the six model parameters that gives the best possible separation between the genotype groups. This method simultaneously accounts for differences in combinations of variables between groups over and beyond differences across single multiple variables (Ramos and Liow 2012).

Results

Reduced learning performance in DRD2/ANKK1 TaqIA A1 carriers

In our previous study (Richter et al. 2014), we observed that in the *no-go to win* condition, DRD2/ANKK1 TaqIA A1 carriers showed a significantly diminished improvement from the first to the second half of the experiment compared to A2 homozygotes (cohort 1: $t_{85} = -2.78$, $p = 0.007$; cohort 2: $t_{93} = -2.16$, $p = 0.033$). As expected, we replicated this finding in our current sample (cohort 3: $t_{97} = 2.05$, $p = 0.043$; Fig. 2A). In all other conditions, A1 carriers and A2 homozygotes did not significantly differ (all $p > 0.100$), nor in gender ($p = 0.621$), age ($p = 0.749$), the number of smokers and nonsmokers ($p = 0.084$), or in the COMT Val108/158Met genotype distribution ($p = 0.901$).

Furthermore, we also analyzed task effects and replicated previous results showing an action by valence interaction on overall task performance (Guitart-Masip et al. 2012, 2014; Cavanagh et al. 2013; Chowdhury et al. 2013; Richter et al. 2014; de Berker et al. 2016; Swart et al. 2017, 2018; de Boer et al. 2019; Dorfman and Gershman, 2019; Betts et al. 2020; Kuhnel et al. 2020; Perosa et al. 2020; van Nuland et al. 2020; Ereira et al. 2021); see supplementary results and Table S2 for details).

Genotyping results in the entire sample

Our further analyses of genetically driven effects were performed in the entire sample comprising all three cohorts ($N = 281$ participants). Within this group, 99 carriers of the DRD2/ANKK1 A1 allele (35.2%; 10 A1/A1 and 89 A1/A2) and 182 A2 homozygotes were identified. For the COMT Val108/158Met polymorphism, 83 subjects were Met homozygous, 70 subjects were Val homozygous, and the remaining 128 subjects were heterozygous. These distributions are within the expected range for a European population (see Supplementary Table S3; NCBI ALFA project release version: 20201027095038; (Phan et al. 2020). Genotype frequencies were in Hardy–Weinberg equilibrium (all $p > 0.145$), and there was no linkage between the two polymorphisms ($p = 0.971$; for detailed demographics, see Table 1).

To further control for effects of population stratification, genotyping was also performed for a variety of additional polymorphisms with a known distribution in European populations (see Supplementary Table S3). The distributions were in line with previously reported frequencies and did not differ between genotype groups of the

DRD2/ANKK1 and COMT polymorphisms (all $p > 0.112$), thus making genetic inhomogeneity of the tested population unlikely.

DRD2/ANKK1 TaqIA and COMT genotypes differentially modulate motivational learning biases

In line with our previous work (Richter et al. 2014), we observed for the DRD2/ANKK1 TaqIA SNP a significant *genotype* \times *time* \times *action* \times *valence* interaction ($F_{1,271} = 11.18$, $p = 0.001$; see Fig. 2B), as well as significant interactions of *genotype* \times *time* ($F_{1,271} = 11.08$, $p = 0.001$) and *genotype* \times *time* \times *action* ($F_{1,271} = 11.94$, $p = 0.001$). *Post hoc* comparisons revealed that A1 carriers exhibited an overall significantly worse learning performance throughout the experiment compared to A2 homozygotes (overall slope: $t_{279} = -3.72$, $p < 0.001$, Cohen's $d = 0.47$). This effect was solely carried by the *no-go* conditions (*no-go* slope: $t_{279} = -4.56$, $p < 0.001$, Cohen's $d = 0.58$; *go* slope: $p = 0.748$), and specifically by the *no-go to win* condition (*ngw* slope: $t_{279} = -4.41$, $p < 0.001$, Cohen's $d = 0.54$; all other conditions: all $p > 0.087$). As displayed in Fig. 2B and C, the DRD2/ANKK1 TaqIA A1 carriers reached their learning asymptote earlier and to a lower level. They significantly differed in performance from the A2 homozygotes only during the second half of the experiment, pointing to different learning capacities (overall 2nd half: $t_{279} = -2.21$, $p = 0.028$, Cohen's $d = 0.35$; *no-go* 2nd half: $t_{279} = -2.28$, $p = 0.024$, Cohen's $d = 0.29$; *ngw* 2nd half: $t_{279} = -2.06$, $p = 0.041$, Cohen's $d = 0.26$; equivalent 1st half comparisons: all $p > 0.340$). A summary of the statistics is displayed in Supplementary Tables S4 and S5.

The combined datasets allowed for a logistic regression on the trial-by-trial *go* responses (see supplementary results and Figure S2 for details). This analysis confirmed the ANCOVA results with A1 carriers showing significantly diminished *no-go to win* performance in the course of the experiment (Fig. 2C).

For the COMT Val108/158Met polymorphism, we observed a trend toward a significant four-way interaction *genotype* \times *time* \times *action* \times *valence* ($F_{2,271} = 2.96$, $p = 0.053$). Met homozygotes showed significantly increased learning throughout the experiment in the *no-go to win* (*ngw* slope: $t_{209} = 2.02$, $p = 0.045$; Fig. 3) and the *go to avoid losing* conditions (*gl* slope: $t_{209} = 2.48$, $p = 0.014$) compared to heterozygotes (other conditions: all $p > 0.922$). The logistic regression did not show an effect of COMT genotype ($p = 0.381$; see supplementary results and Figure S3 for details).

In light of previous evidence that Met homozygotes have a higher response bias relative to Val carriers (Lancaster et al. 2012, 2015; Goetz et al. 2013; Corral-Frias et al. 2016), in an additional analysis, participants were

separated into Met homozygotes (Met/Met) and Val allele carriers (Val/Val and Val/Met). The ANCOVA revealed a significant *genotype* \times *time* \times *action* \times *valence* interaction ($F_{1,273} = 4.30$, $p = 0.039$) as well as a significant main effect of COMT genotype ($F_{1,273} = 4.55$, $p = 0.034$) and interestingly also a significant interaction of the COMT with the TaqIA genotype ($F_{1,273} = 3.88$, $p = 0.050$). The latter finding indicates a beneficial effect of Met homozygosity on overall performance in A1 carriers ($t_{97} = 2.31$, $p = 0.024$) but not in A2 homozygotes ($p = 0.971$).

We controlled for potential effects in reaction times (participants were explicitly instructed to respond accurately) and false responses in the target detection task (i.e., left when the target was on the right side of the display or vice versa) and found no significant differences between genotype groups ($p > 0.187$; see supplement for details).

Computational modeling of task performance

To identify components of the observed asymmetry during learning, we constructed six nested reinforcement learning models to fit participants' behavior (Table 2). Our computational modeling approach demonstrated that the marked asymmetry in learning could be best accounted for by the model including separate parameters for sensitivity to rewards and punishments as well as a learning rate, an irreducible noise parameter, a constant *go* bias parameter, and a constant Pavlovian bias parameter (see Table 2), which is consistent with our recently published lifetime study on motivational learning (Betts et al. 2020). The simulations of the winning model are presented in Fig. 1C. Neither one specific model parameter (independent samples t tests: all $p > 0.119$), nor a linear combination of the parameters (LDA: all $p > 0.636$) showed significant genotype-related differences.

Discussion

In the present study, we investigated how genetic determinants of striatal and prefrontal DA function modulate learning biases when action and valence are experimentally orthogonalized. Using the previously established valenced *go/no-go* task (Guitart-Masip et al. 2012), we provide independent confirmation for a selective deficit of DRD2/ANKK1 TaqIA A1 carriers in learning to inhibit an action to obtain a reward. Moreover, our exploratory analysis yielded preliminary evidence that COMT Met homozygotes show superior learning during trials with incongruent coupling of action and valence. Due to previous knowledge about their neurophysiological consequences, the genetic polymorphisms studied here allow conclusions about differential contributions of striatal and prefrontal DA function to

COMT SNP effects on accuracy in the entire sample (N=281)

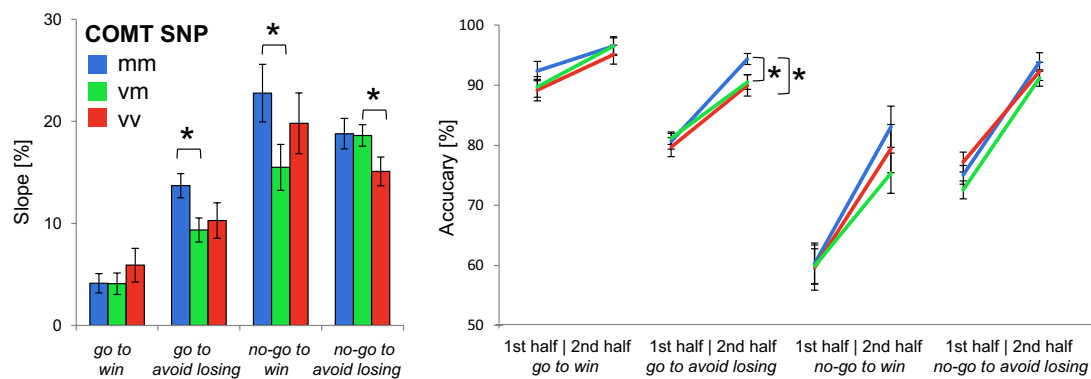


Fig. 3 Effects of COMT genotype on choice performance in the entire sample. Left panels: bar plots show mean differences between correct response rates (\pm SEM) during second half versus the first half of trials for each condition. This score represents the observed fourfold interaction of *action* \times *valence* \times *time* \times *genotype*. Right pan-

els: line charts show mean values of correct responses (\pm SEM) in the first and the second half of trials for all four conditions. Met homozygotes showed increased learning throughout the experiment in the *no-go to win* and *go avoid losing* condition relative to heterozygotes. Post hoc comparisons via *t* tests: $*p < 0.05$

instrumental control mechanisms (Schott et al. 2006; Mier et al. 2010; Corral-Frias et al. 2016).

Selective modulation of the *no-go to win* condition by DRD2/ANKK1 TaqIA genotype

For the DRD2/ANKK1 TaqIA polymorphism, we replicated our previous observation (Richter et al. 2014) that A1 carriers show a stronger coupling of action and valence in a third independent cohort. As in our previous study, A1 carriers exhibited a specific impairment in learning to withhold actions in reward contexts. When combining all three datasets ($N = 281$), we could more closely investigate the nature of this effect.

D2-type DA receptors are primarily expressed in the striatum (*post-mortem* autoradiography: Joyce et al. 1991; Kessler et al. 1993; Hall et al. 1996; in vivo PET: Okubo et al. 1999; MacDonald et al. 2009). They function as both postsynaptic inhibitory receptors and as presynaptic autoreceptors that regulate neurotransmission via negative feedback (Bello et al. 2011, for reviews, see Wolf and Roth, 1990; Schmitz et al. 2003). While DRD2 is, albeit sparsely, expressed in extrastriatal regions (2–8% of the expression level in the striatum, Suhara et al. 1999) and cortically mediated effects can thus not be excluded, differences for the DRD2/ANKK1 TaqIA genotypes have thus far only been observed for the striatum—with lower DRD2 expression or binding availability in A1 carriers (*post-mortem* autoradiography: Noble et al. 1991; Thompson et al. 1997; Ritchie and

Noble, 2003; in vivo PET: for review and meta-analysis, see Gluskin and Mickey 2016; Eisenstein et al. 2016).

Those techniques cannot differentiate between presynaptic and postsynaptic D2 receptors. Thus genetically mediated differences in dopamine-dependent learning processes may to some extent be attributable to reduced availability of presynaptic autoinhibitory D2 receptors, which in turn may underlie the previously reported increased DA synthesis capacity in A1 carriers (Laakso et al. 2005; Fig. 4). Two SNPs of the DRD2 gene, rs2283265 and rs1076560, have previously been associated with alternative splicing and a rather selective decrease of presynaptic D2 receptor expression (Zhang et al. 2007). Notably, in a motivational learning study, the haplotype linked to lower presynaptic D2 receptor availability was associated with relatively impaired avoidance learning, but intact approach learning (Frank and Hutchison 2009). However, it is not possible to separate in this study whether the effects were actually due to the aversive nature of the feedback or to poorer *no-go* learning, because there was no control of the coupling of action and valence. Nevertheless, that finding is compatible with the possibility that the rather selective deficit of A1 carriers in the *no-go to win* condition observed in the present study may, at least in part, be attributable to reduced presynaptic D2 receptor density.

Another factor that comes into play are the assumed different functions in reward learning of dorsal striatal regions that include the caudate nucleus and putamen specifically involved in learning about actions and their reward consequences, and ventral striatal regions, encompassing the

Table 1 Descriptive data of the entire sample regarding DRD2/ANKK1 TaqIA and COMT Val108/158Met genotypes

DRD2/ANKK1 TaqIA	A1 +		A1 –		A1 + > A1 –	
Gender (<i>N</i> women/men)	44/55		102/80		$\chi^2 = 3.46$ $p = .063$	
Age in years (M \pm SD)	25.1 \pm 3.1		24.6 \pm 2.6		$t_{279} = 1.30$ $p = .195$	
Non-smokers/smokers (<i>N</i>)	70/29		143/39		$\chi^2 = 2.16$ $p = .141$	
COMT (<i>N</i> MM/VM/VV)	30/45/24		53/83/46		$\chi^2 = 0.06$ $p = .971$	
COMT Val108/158Met	MM	VM	VV	MM > VV	VM > VV	MM > VM
Gender (<i>N</i> women/men)	43/40	64/64	39/31	$\chi^2 = 0.59, p = .743$		
Age in years (M \pm SD)	25.0 \pm 2.7	24.9 \pm 2.8	24.3 \pm 3.0	$t_{151} = 1.36$ $p = .176$	$t_{196} = 1.37$ $p = .174$	$t_{209} = 0.09$ $p = .931$
Non-smokers/smokers (<i>N</i>)	63/20	93/35	57/13	$\chi^2 = 1.90, p = .387$		
TaqIA (<i>N</i> A1 +/A1 –)	30/53	45/83	24/46	$\chi^2 = 0.06, p = .971$		

Demographic data are pooled across all three cohorts (cohort 1 and 2 from Richter et al. (2014), and the newly investigated cohort 3). *N* = number, *M* = mean, *SD* = standard deviation, *VM*: Val/Met heterozygotes, *VV*: Val homozygotes, *A1 +*: carriers of the *A1* allele, *A1 –*: *A2* homozygotes

nucleus accumbens classically linked to expected value representations (Wickens et al. 2003, 2007; O’Doherty et al. 2004).

While differences in DRD2 binding availability of DRD2/ANKK1 TaqIA *A1* allele carriers have been observed for all striatal subregions (putamen, caudate, and nucleus accumbens; Eisenstein et al. 2016), studies using the valenced *go/no-go* learning task investigating regionally specific striatal functions thus far only observed correlations with the dorsal striatum. De Boer et al. (2019) investigated cortical and striatal sources of variance in D1 receptor availability in humans using PET and could show that higher levels of endogenous D1 receptor availability in the dorsal striatum were related to biases during learning. Perosa et al. (2020) analyzed voxel-based morphometry using 7 Tesla MRI images and could show that individual differences in learning rate in older adults were related to the volume of the caudate nucleus. Relatedly, an fMRI study in young adults using a variation of the task that does not require learning (Guitart-Masip et al. 2011) demonstrated an association between the anticipation of action value and activity in the dorsal striatum suggesting its crucial role for evaluating the weight of an action. Thus, it is tempting to speculate that the observed effects of the DRD2/ANKK1 TaqIA genotype on motivational biases may be more related to dorsal striatal action learning as compared to ventral striatal functions in reward value representations, but clearly future studies are needed to answer this issue.

Effects of the COMT Val108/158Met polymorphism and a potential role for prefrontal dopamine

Beyond replicating and expanding our findings on the DRD2/ANKK1 TaqIA polymorphism, the larger sample size of our three combined samples made it possible to investigate the effects of and potential interactions with the COMT Val108/158Met polymorphism.

The role of COMT in DA clearance has been subject to extensive research since the first studies suggesting a role for the COMT Val108/158Met polymorphism in human PFC function (Egan et al. 2001; Weinberger et al. 2001). Despite some evidence for a role for membrane-bound COMT in striatal DA metabolism (Laatikainen et al. 2013), converging evidence from animal studies and human *post-mortem* investigations suggests that COMT is primarily important for DA inactivation in the PFC, whereas its role in the striatum appears to be quantitatively negligible in most cases (Huotari et al. 2002; Matsumoto et al. 2003; Yavich et al. 2007; Kaenmaki et al. 2010; Korn et al. 2021). This has been attributed to the sparse cortical expression of the DA transporter (DAT; Chen et al. 2004; Kaenmaki et al. 2010; Tunbridge, 2010). Therefore, the COMT polymorphism has mostly been studied in relation to PFC-dependent executive functions (for reviews, see Frank and Fossella 2011; Klanker et al. 2013; for a meta-analysis, see Mier et al. 2010). With respect to motivated behavior, homozygosity for the Met allele has been associated with relatively increased reward learning (for a meta-analysis, see Corral-Frias et al. 2016). In our study, Met homozygosity is associated with stronger learning enhancement during Pavlovian conflict (i.e., incongruent

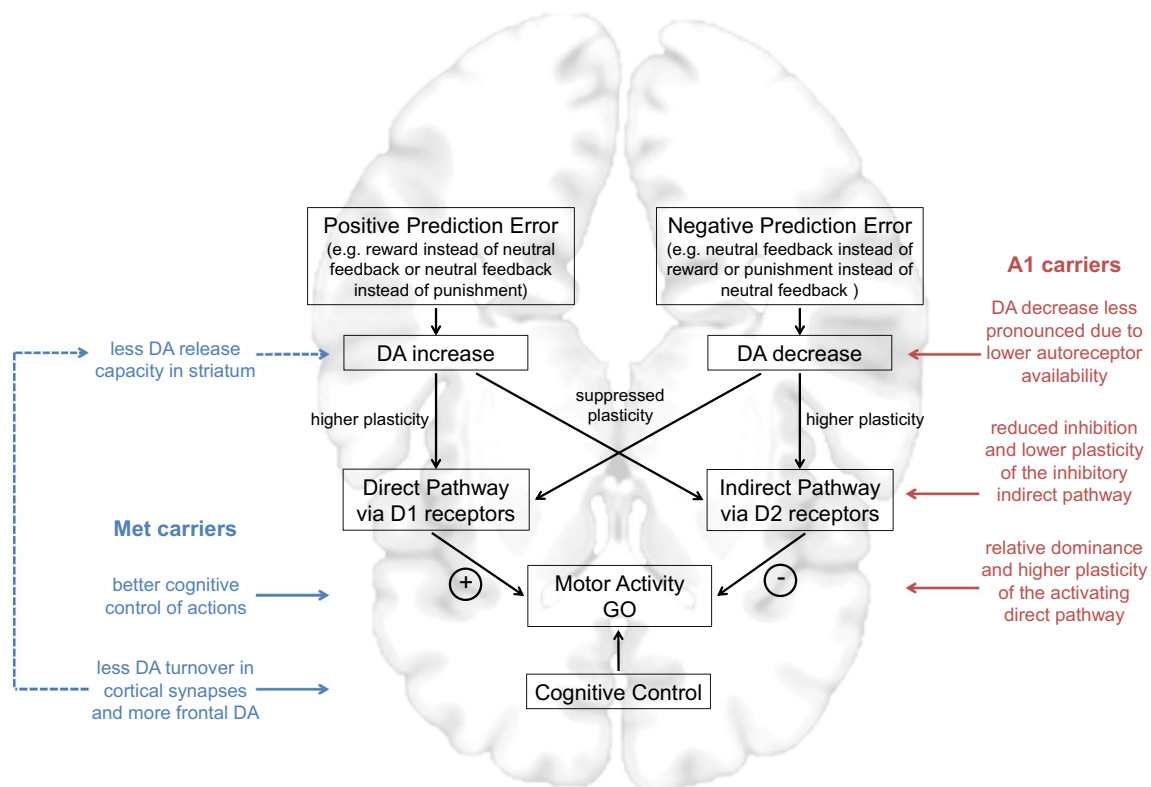


Fig. 4 A model of genetically driven contributions to the coupling of action and valence during learning. DA neurons signal positive reward prediction errors by phasic bursts and negative prediction errors by dips below baseline firing rate. While the first reinforces the direct pathway via activation of D1 receptors and thereby facilitates the future generation of *go* choices, the second reinforces the indirect pathway via reduced activation of D2 receptors and thus facilitates the future generation of *no-go* choices in comparable situations. A1 carriers would be assumed to have reduced D2 receptor-binding capacity decreasing autoinhibition of dopaminergic signaling after

negative prediction errors in the indirect pathway and a shift to a more action-oriented behavioral pattern mediated by the direct pathway. COMT Val108/158Met Met carriers would be assumed to have higher frontal DA availability facilitating working memory and attentional processes. Moreover, indirect downstream effects on striatal DA regulation may add on improving performance under Pavlovian conflict in Met compared to Val homozygotes. The MNI template brain from MRICroGL (“mni152”) was used in this illustration. Figure adapted from Richter et al. (2014)

coupling of action and valence) throughout the experiment—thus, improved performance when motivational biases are involved. This may be related to COMT’s impact on prefrontal DA levels and prefrontal function. It should be noted, though, that despite the majority of studies showing a minor role for COMT in striatal DA metabolism, there is evidence for a delicately balanced mutual regulation of prefrontal and striatal DA turnover (Akil et al. 2003). Animal studies suggest that transgenic mice with increased COMT activity, equivalent to the relative increase in activity observed with the human COMT Val allele, do not only show deficits in PFC-dependent tasks (e.g., stimulus–response learning and working memory), but also increased DA release capacity in the striatum (Simpson et al. 2014). This finding corroborates earlier human neuroimaging studies that reported higher midbrain DA synthesis capacity in Val compared to Met homozygotes (Akil et al. 2003; Meyer-Lindenberg et al.

2005). Therefore, to the extent that the COMT genotype affects prefrontal function, it may contribute to motivational learning not only because of its biological effects in the PFC but also because of indirect downstream effects on striatal DA regulation (Fig. 4). Thus, compared with the Val allele, the Met allele, which is likely associated with relatively increased prefrontal DA signaling, would result in relatively decreased disinhibition of mesencephalic DA activity, e.g., in neuronal populations projecting to the striatum (Akil et al. 2003; Fig. 4).

Limitations

A limitation in the interpretation of our data that is also common in other studies on this topic lies in the fact that the molecular mechanisms underlying the observed effects are still under debate. It is well known that the Taq1A

Table 2 Integrated Bayesian information criteria (iBIC) for tested models

Model no	Model parameters	No. of parameters	Likelihood	Pseudo- R^2	iBIC
1	ε, ρ	2	-23,463	0.498	46,970
2	ε, ρ, ξ	3	-23,314	0.501	46,695
3	$\varepsilon, \rho, \xi, b$	4	-21,798	0.534	43,685
4	$\varepsilon, \rho_{win}, \rho_{lose}, \xi, b$	5	-21,334	0.544	42,779
5	$\varepsilon, \rho_{win}, \rho_{lose}, \xi, b, \pi_{variable}$	6	-21,137	0.548	42,406
6	$\varepsilon, \rho_{win}, \rho_{lose}, \xi, b, \pi_{constant}$	6	-21,106	0.549	42,346

Boldface type: winning model statistics, ε : learning rate, ρ_{win} : weighting of reward on win trials, ρ_{lose} : weighting of punishments on lose trials. ξ : irreducible noise, b : go bias, π : Pavlovian bias, iBIC: integrated Bayesian information criterion (smaller iBIC values indicate a better model fit). Descriptives for the parameters in the winning model (M \pm SD): $\varepsilon = 0.26 \pm 0.15$, $\rho_{win} = 15.32 \pm 13.30$, $\rho_{lose} = 7.51 \pm 4.03$, $\xi = 0.96 \pm 0.06$, $b = 1.10 \pm 0.74$, $\pi_{constant} = 0.65 \pm 0.57$

polymorphism is not located within the DRD2 gene but 10 kb downstream of its termination codon on chromosome 11q23.1, within the coding region of the adjacent ankyrin repeat and kinase domain containing 1 (ANKK1) gene (Dubertret et al. 2004; Neville et al. 2004). The molecular mechanisms underlying the effects of ANKK1 TaqIA on striatal DRD2 availability have not been conclusively established. Multiple mechanisms have been proposed, including linkage disequilibrium (Duan et al. 2003; Ritchie and Noble, 2003; Fossella et al. 2006; Doehring et al. 2009; Richter et al. 2017) or a potential direct interaction of ANKK1 with the D2 receptor at protein level, potentially modulated by the TaqIA polymorphism (Hoenicka et al. 2010; Garrido et al. 2011; Ponce et al. 2016); for a review, see Ponce et al. 2009; see Supplementary Discussion for details). Similarly, for the COMT Val108/158Met polymorphism, it remains to be determined how COMT-dependent DA inactivation in brain regions with low DAT expression is realized. There is only limited evidence for extracellular activity of membrane-bound COMT (Chen et al. 2011), and the predominant evidence points to intracellular orientation and activity, requiring a DAT-independent uptake mechanism (Myohanen et al. 2010; Schott et al. 2010; see Supplementary Discussion).

Moreover, we only investigated two dopaminergic SNPs, and it must be noted that there are several additional genetic variants in the dopaminergic system that could be important for the generation and overcoming of motivational learning biases. In the Supplementary Discussion, we summarize the previous results on motivated behavior, focusing on the commonly investigated DAT1 VNTR rs28363170, the DARPP-32 rs907094, and the DRD2 C957T rs6277 polymorphism. Owing to the sample size, those polymorphisms were not investigated in the present study.

A further limitation lies in our modeling approach, which failed to reflect the very robust and replicated effect of the DRD2/ANKK1 TaqIA SNP on learning gain throughout the experiment in the *no-go to win* condition and on

the time-dependent valence effect on individual *go/no-go* responses. One explanation could be that the model space does not include the computational mechanism to differentiate, for example, instrumental from Pavlovian contributions. This should be addressed in future studies.

Conclusion

With our study, we demonstrate by assessing the contributions of two well-studied genetic polymorphisms that DRD2/ANKK1 TaqIA A1 carriers with presumably reduced striatal D2 receptor-binding capacity and less autoinhibition of striatal dopaminergic signaling after negative prediction errors in the indirect pathway showed a shift to a more action-oriented and biased behavioral pattern. COMT Val108/158Met Met homozygotes, who presumably exhibit higher prefrontal DA activity, showed less biased learning, possibly reflecting more efficient frontal control.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00702-021-02382-4>.

Acknowledgements We are grateful to Herta Flor for valuable comments on the manuscript. We thank Iris Mann, Catherine Libeau, and Timo Lemme for help with testing.

Author contributions AR, LdB, MG-M, CIS, and BHS wrote the manuscript. AR, MG-M, and BHS conceptualized the study design. AR and GB collected the data. AR, LdB, and GB analyzed and curated data.

Funding Open Access funding enabled and organized by Projekt DEAL. This project was supported by the Deutsche Forschungsgemeinschaft (SFB 779/A08 and SFB 1436/A05 to CIS and BHS as well as RI 2964-1 to AR). Work in the laboratory of BHS was supported by the EU/EFRE-funded "Autonomy in Old Age" Research Alliance of the State of Saxony-Anhalt. MG-M and LdB were supported by a research grant from the Swedish Research Council (VT521-2013-2589) awarded to MG-M. The funding agencies had no role in the design of the study or interpretation of the data.

Declarations

Conflict of interest The authors have no conflicts of interest, financial or otherwise, to report.

Ethics approval The study was approved by the Ethics Committee of the Faculty of Medicine at the Otto von Guericke University of Magdeburg.

Consent to participate and publish All individual participants included in the study gave written informed consent in accordance with the Declaration of Helsinki and received financial compensation for participation.

Data and code availability The datasets generated during and/or analyzed during the current study as well as the code are available from the corresponding author on reasonable request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akil M, Kolachana BS, Rothmond DA, Hyde TM, Weinberger DR, Kleinman JE (2003) Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. *J Neurosci: off J Soc Neurosci* 23:2008–2013
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141
- Bayer HM, Lau B, Glimcher PW (2007) Statistics of midbrain dopamine neuron spike trains in the awake primate. *J Neurophysiol* 98:1428–1439
- Bello EP, Mateo Y, Gelman DM, Noain D, Shin JH, Low MJ, Alvarez VA, Lovinger DM, Rubinstein M (2011) Cocaine supersensitivity and enhanced motivation for reward in mice lacking dopamine D2 autoreceptors. *Nat Neurosci* 14:1033–1038
- Betts MJ, Richter A, de Boer L, Tegelbeckers J, Perosa V, Baumann V, Chowdhury R, Dolan RJ, Seidenbecher C, Schott BH, Duzel E, Guitart-Masip M, Krauel K (2020) Learning in anticipation of reward and punishment: perspectives across the human lifespan. *Neurobiol Aging* 96:49–57
- Cavanagh JF, Eisenberg I, Guitart-Masip M, Huys Q, Frank MJ (2013) Frontal theta overrides pavlovian learning biases. *J Neurosci: off J Soc Neurosci* 33:8541–8548
- Chen J, Lipska BK, Halim N, Ma QD, Matsumoto M, Melhem S, Kolachana BS, Hyde TM, Herman MM, Apud J, Egan MF, Kleinman JE, Weinberger DR (2004) Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain. *Am J Hum Genet* 75:807–821
- Chen J, Song J, Yuan P, Tian Q, Ji Y, Ren-Patterson R, Liu G, Sei Y, Weinberger DR (2011) Orientation and cellular distribution of membrane-bound catechol-O-methyltransferase in cortical neurons: implications for drug development. *J Biol Chem* 286:34752–34760
- Chowdhury R, Guitart-Masip M, Lambert C, Dolan RJ, Duzel E (2013) Structural integrity of the substantia nigra and subthalamic nucleus predicts flexibility of instrumental learning in older-age individuals. *Neurobiol Aging* 34:2261–2270
- Corral-Frias NS, Pizzagalli DA, Carre JM, Michalski LJ, Nikolova YS, Perlis RH, Fagerness J, Lee MR, Conley ED, Lancaster TM, Haddad S, Wolf A, Smoller JW, Hariri AR, Bogdan R (2016) COMT Val(158) Met genotype is associated with reward learning: a replication study and meta-analysis. *Genes Brain Behav* 15:503–513
- de Berker AO, Tirole M, Rutledge RB, Cross GF, Dolan RJ, Bestmann S (2016) Acute stress selectively impairs learning to act. *Sci Rep* 6:29816
- de Boer L, Axelsson J, Chowdhury R, Riklund K, Dolan RJ, Nyberg L, Backman L, Guitart-Masip M (2019) Dorsal striatal dopamine D1 receptor availability predicts an instrumental bias in action learning. *Proc Natl Acad Sci USA* 116:261–270
- Doehring A, Hentig N, Graff J, Salamat S, Schmidt M, Geisslinger G, Harder S, Lotsch J (2009) Genetic variants altering dopamine D2 receptor expression or function modulate the risk of opiate addiction and the dosage requirements of methadone substitution. *Pharmacogenet Genomics* 19:407–414
- Dorfman HM, Gershman SJ (2019) Controllability governs the balance between Pavlovian and instrumental action selection. *Nat Commun* 10:5826
- Duan J, Wainwright MS, Comeron JM, Saitou N, Sanders AR, Gelenter J, Gejman PV (2003) Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Hum Mol Genet* 12:205–216
- Dubertret C, Gouya L, Hanoun N, Deybach JC, Ades J, Hamon M, Gorwood P (2004) The 3' region of the DRD2 gene is involved in genetic susceptibility to schizophrenia. *Schizophr Res* 67:75–85
- Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, Goldman D, Weinberger DR (2001) Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc Natl Acad Sci USA* 98:6917–6922
- Eisenberg DT, Mackillop J, Modi M, Beauchemin J, Dang D, Lisman SA, Lum JK, Wilson DS (2007) Examining impulsivity as an endophenotype using a behavioral approach: a DRD2 TaqI A and DRD4 48-bp VNTR association study. *Behav Brain Funct* 3:2
- Eisenstein SA, Bogdan R, Love-Gregory L, Corral-Frias NS, Koller JM, Black KJ, Moerlein SM, Perlmutter JS, Barch DM, Hershey T (2016) Prediction of striatal D2 receptor binding by DRD2/ANKK1 TaqIA allele status. *Synapse* 70:418–431
- Ereira S, Pujol M, Guitart-Masip M, Dolan RJ, Kurth-Nelson Z (2021) Overcoming Pavlovian bias in semantic space. *Sci Rep* 11:3416
- Fossella J, Green AE, Fan J (2006) Evaluation of a structural polymorphism in the ankyrin repeat and kinase domain containing 1 (ANKK1) gene and the activation of executive attention networks. *Cogn Affect Behav Neurosci* 6:71–78
- Frank MJ, Fossella JA (2011) Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacol: off Publ Am Coll Neuropsychopharmacol* 36:133–152
- Frank MJ, Hutchison K (2009) Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience* 164:131–140
- Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007) Genetic triple dissociation reveals multiple roles for

- dopamine in reinforcement learning. *Proc Natl Acad Sci USA* 104:16311–16316
- Garrido E, Palomo T, Ponce G, Garcia-Consuegra I, Jimenez-Arriero MA, Hoenicka J (2011) The ANKK1 protein associated with addictions has nuclear and cytoplasmic localization and shows a differential response of Ala239Thr to apomorphine. *Neurotox Res* 20:32–39
- Gluskin BS, Mickey BJ (2016) Genetic variation and dopamine D2 receptor availability: a systematic review and meta-analysis of human in vivo molecular imaging studies. *Transl Psychiatry* 6:e747
- Goetz EL, Hariri AR, Pizzagalli DA, Strauman TJ (2013) Genetic moderation of the association between regulatory focus and reward responsiveness: a proof-of-concept study. *Biol Mood Anxiety Disord* 3:3
- Guitart-Masip M, Fuentemilla L, Bach DR, Huys QJ, Dayan P, Dolan RJ, Duzel E (2011) Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J Neurosci: off J Soc Neurosci* 31:7867–7875
- Guitart-Masip M, Huys QJ, Fuentemilla L, Dayan P, Duzel E, Dolan RJ (2012) Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* 62:154–166
- Guitart-Masip M, Economides M, Huys QJ, Frank MJ, Chowdhury R, Duzel E, Dayan P, Dolan RJ (2014) Differential, but not opponent, effects of L-DOPA and citalopram on action learning with reward and punishment. *Psychopharmacology* 231:955–966
- Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacol: off Publ Am Coll Neuropsychopharmacol* 35:4–26
- Hall H, Farde L, Halldin C, Hurd YL, Pauli S, Sedvall G (1996) Autoradiographic localization of extrastriatal D2-dopamine receptors in the human brain using [¹²⁵I]epidepride. *Synapse* 23:115–123
- Hikida T, Kimura K, Wada N, Funabiki K, Nakanishi S (2010) Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* 66:896–907
- Hitchcott PK, Quinn JJ, Taylor JR (2007) Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cereb Cortex* 17:2820–2827
- Hoenicka J, Quinones-Lombrana A, Espana-Serrano L, Alvira-Botero X, Kremer L, Perez-Gonzalez R, Rodriguez-Jimenez R, Jimenez-Arriero MA, Ponce G, Palomo T (2010) The ANKK1 gene associated with addictions is expressed in astroglial cells and upregulated by apomorphine. *Biol Psychiat* 67:3–11
- Huotari M, Gogos JA, Karayiorgou M, Koponen O, Forsberg M, Raasmaja A, Hyttinen J, Mannisto PT (2002) Brain catecholamine metabolism in catechol-O-methyltransferase (COMT)-deficient mice. *Eur J Neurosci* 15:246–256
- Huys QJ, Cools R, Golzer M, Friedel E, Heinz A, Dolan RJ, Dayan P (2011) Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput Biol* 7:e1002028
- Jocham G, Klein TA, Neumann J, von Cramon DY, Reuter M, Ullsperger M (2009) Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. *J Neurosci: off J Soc Neurosci* 29:3695–3704
- Joyce JN, Janowsky A, Neve KA (1991) Characterization and distribution of [¹²⁵I]epidepride binding to dopamine D2 receptors in basal ganglia and cortex of human brain. *J Pharmacol Exp Ther* 257:1253–1263
- Kaenmaki M, Tammimaki A, Myohanen T, Pakarinen K, Amberg C, Karayiorgou M, Gogos JA, Mannisto PT (2010) Quantitative role of COMT in dopamine clearance in the prefrontal cortex of freely moving mice. *J Neurochem* 114:1745–1755
- Kessler RM, Whetsell WO, Ansari MS, Votaw JR, de Paulis T, Clanton JA, Schmidt DE, Mason NS, Manning RG (1993) Identification of extrastriatal dopamine D2 receptors in post mortem human brain with [¹²⁵I]epidepride. *Brain Res* 609:237–243
- Klanker M, Feenstra M, Denys D (2013) Dopaminergic control of cognitive flexibility in humans and animals. *Front Neurosci* 7:201
- Klein TA, Neumann J, Reuter M, Hennig J, von Cramon DY, Ullsperger M (2007) Genetically determined differences in learning from errors. *Science* 318:1642–1645
- Koeneke A, Ponce G, Troya-Balseca J, Palomo T, Hoenicka J (2020) Ankyrin repeat and kinase domain containing 1 gene, and addiction vulnerability. *Int J Mol Sci* 21:2516
- Korn C, Akam T, Jensen KHR, Vagnoni C, Huber A, Tunbridge EM, Walton ME (2021) Distinct roles for dopamine clearance mechanisms in regulating behavioral flexibility. *Mol Psychiatry*. <https://doi.org/10.1038/s41380-021-01194-y>
- Kuhnel A, Teckentrup V, Neuser MP, Huys QJM, Burrasch C, Walter M, Kroemer NB (2020) Stimulation of the vagus nerve reduces learning in a go/no-go reinforcement learning task. *Eur Neuropsychopharmacol* 35:17–29
- Laakso A, Pohjalainen T, Bergman J, Kajander J, Haaparanta M, Solin O, Syvalahti E, Hietala J (2005) The A1 allele of the human D2 dopamine receptor gene is associated with increased activity of striatal L-amino acid decarboxylase in healthy subjects. *Pharmacogenet Genomics* 15:387–391
- Latikainen LM, Sharp T, Harrison PJ, Tunbridge EM (2013) Sexually dimorphic effects of catechol-O-methyltransferase (COMT) inhibition on dopamine metabolism in multiple brain regions. *PLoS ONE* 8:e61839
- Lancaster TM, Linden DE, Heerey EA (2012) COMT val158met predicts reward responsiveness in humans. *Genes Brain Behav* 11:986–992
- Lancaster TM, Heerey EA, Mantripragada K, Linden DE (2015) Replication study implicates COMT val158met polymorphism as a modulator of probabilistic reward learning. *Genes Brain Behav* 14:486–492
- Lee SH, Ham BJ, Cho YH, Lee SM, Shim SH (2007) Association study of dopamine receptor D2 TaqI A polymorphism and reward-related personality traits in healthy Korean young females. *Neuropsychobiology* 56:146–151
- MacDonald SW, Cervenka S, Farde L, Nyberg L, Backman L (2009) Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. *Neuropsychologia* 47:2299–2304
- Matsumoto M, Weickert CS, Akil M, Lipska BK, Hyde TM, Herman MM, Kleinman JE, Weinberger DR (2003) Catechol O-methyltransferase mRNA expression in human and rat brain: evidence for a role in cortical neuronal function. *Neuroscience* 116:127–137
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346
- Meyer-Lindenberg A, Kohn PD, Kolachana B, Kippenhan S, McNerney-Leo A, Nussbaum R, Weinberger DR, Berman KF (2005) Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. *Nat Neurosci* 8:594–596
- Mier D, Kirsch P, Meyer-Lindenberg A (2010) Neural substrates of pleiotropic action of genetic variation in COMT: a meta-analysis. *Mol Psychiatry* 15:918–927
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci: off J Soc Neurosci* 16:1936–1947
- Myohanen TT, Schendzielorz N, Mannisto PT (2010) Distribution of catechol-O-methyltransferase (COMT) proteins and enzymatic activities in wild-type and soluble COMT deficient mice. *J Neurochem* 113:1632–1643

- Neville MJ, Johnstone EC, Walton RT (2004) Identification and characterization of ANKK1: a novel kinase gene closely linked to DRD2 on chromosome band 11q23.1. *Hum Mutat* 23:540–545
- Noble EP, Blum K, Ritchie T, Montgomery A, Sheridan PJ (1991) Allelic association of the D2 dopamine receptor gene with receptor-binding characteristics in alcoholism. *Arch Gen Psychiatry* 48:648–654
- Noble EP, Ozkaragoz TZ, Ritchie TL, Zhang X, Belin TR, Sparkes RS (1998) D2 and D4 dopamine receptor polymorphisms and personality. *Am J Med Genet* 81:257–267
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454
- Okubo Y, Olsson H, Ito H, Lofti M, Suhara T, Halldin C, Farde L (1999) PET mapping of extrastriatal D2-like dopamine receptors in the human brain using an anatomic standardization technique and [11C]FLB 457. *Neuroimage* 10:666–674
- Pan YQ, Qiao L, Xue XD, Fu JH (2015) Association between ANKK1 (rs1800497) polymorphism of DRD2 gene and attention deficit hyperactivity disorder: a meta-analysis. *Neurosci Lett* 590:101–105
- Perosa V, de Boer L, Ziegler G, Apostolova I, Buchert R, Metzger C, Amthauer H, Guitart-Masip M, Duzel E, Betts MJ (2020) The role of the striatum in learning to orthogonalize action and valence: a combined PET and 7 T MRI aging study. *Cereb Cortex* 30:3340–3351
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045
- Phan L, J Y, Zhang H, Qiang W, Shekhtman E, Shao D, Revoe D, Villamarin R et al (2020) ALFA: allele frequency aggregator. National Center for Biotechnology Information, U.S. National Library of Medicine, Bethesda
- Ponce G, Perez-Gonzalez R, Aragues M, Palomo T, Rodriguez-Jimenez R, Jimenez-Arriero MA, Hoenicka J (2009) The ANKK1 kinase gene and psychiatric disorders. *Neurotox Res* 16:50–59
- Ponce G, Quinones-Lombrana A, Martin-Palanco NG, Rubio-Solsona E, Jimenez-Arriero MA, Palomo T, Hoenicka J (2016) The addiction-related gene Ankk1 is oppositely regulated by D1R- and D2R-like dopamine receptors. *Neurotox Res* 29:345–350
- Ramos SDS, Liow SJR (2012) Discriminant Function Analysis. <https://doi.org/10.1002/9781405198431.wbeal0335>
- Richter A, Richter S, Barman A, Soch J, Klein M, Assmann A, Libeau C, Behnisch G, Wustenberg T, Seidenbecher CI, Schott BH (2013) Motivational salience and genetic variability of dopamine D2 receptor expression interact in the modulation of interference processing. *Front Hum Neurosci* 7:250
- Richter A, Guitart-Masip M, Barman A, Libeau C, Behnisch G, Czerny S, Schanze D, Assmann A, Klein M, Duzel E, Zenker M, Seidenbecher CI, Schott BH (2014) Valenced action/inhibition learning in humans is modulated by a genetic variant linked to dopamine D2 receptor expression. *Front Syst Neurosci* 8:140
- Richter A, Barman A, Wustenberg T, Soch J, Schanze D, Deibele A, Behnisch G, Assmann A, Klein M, Zenker M, Seidenbecher C, Schott BH (2017) Behavioral and neural manifestations of reward memory in carriers of low-expressing versus high-expressing genetic variants of the dopamine D2 receptor. *Front Psychol* 8:654
- Ritchie T, Noble EP (2003) Association of seven polymorphisms of the D2 dopamine receptor gene with brain receptor-binding characteristics. *Neurochem Res* 28:73–82
- Samochowiec J, Samochowiec A, Puls I, Bienkowski P, Schott BH (2014) Genetics of alcohol dependence: a review of clinical studies. *Neuropsychobiology* 70:77–94
- Schafer M, Rujescu D, Giegling I, Guntermann A, Erfurth A, Bondy B, Moller HJ (2001) Association of short-term response to haloperidol treatment with a polymorphism in the dopamine D(2) receptor gene. *Am J Psychiatry* 158:802–804
- Schizophrenia Working Group of the Psychiatric Genomics, C (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511:421–427
- Schmitz Y, Benoit-Marand M, Gonon F, Sulzer D (2003) Presynaptic regulation of dopaminergic neurotransmission. *J Neurochem* 87:273–289
- Schott BH, Seidenbecher CI, Fenker DB, Lauer CJ, Bunzeck N, Bernstein HG, Tischmeyer W, Gundelfinger ED, Heinze HJ, Duzel E (2006) The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging. *J Neurosci: off J Soc Neurosci* 26:1407–1417
- Schott BH, Frischknecht R, Debska-Vielhaber G, John N, Behnisch G, Duzel E, Gundelfinger ED, Seidenbecher CI (2010) Membrane-bound catechol-O-methyl transferase in cortical neurons and glial cells is intracellularly oriented. *Front Psychiatry* 1:142
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599
- Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74:1–58
- Simpson EH, Morud J, Winiger V, Biezonski D, Zhu JP, Bach ME, Malleret G, Polan HJ, Ng-Evans S, Phillips PE, Kellendonk C, Kandel ER (2014) Genetic variation in COMT activity impacts learning and dopamine release capacity in the striatum. *Learn Mem* 21:205–214
- Smillie LD, Cooper AJ, Proitsis P, Powell JF, Pickering AD (2010) Variation in DRD2 dopamine gene predicts Extraverted personality. *Neurosci Lett* 468:234–237
- Stelzel C, Basten U, Montag C, Reuter M, Fiebach CJ (2010) Frontostriatal involvement in task switching depends on genetic differences in d2 receptor density. *J Neurosci: off J Soc Neurosci* 30:14205–14212
- Stice E, Yokum S, Burger K, Epstein L, Smolen A (2012) Multilocus genetic composite reflecting dopamine signaling capacity predicts reward circuitry responsivity. *J Neurosci: off J Soc Neurosci* 32:10093–10100
- Suhara T, Sudo Y, Okauchi T, Maeda J, Kawabe K, Suzuki K, Okubo Y, Nakashima Y, Ito H, Tanada S, Halldin C, Farde L (1999) Extrastriatal dopamine D2 receptor density and affinity in the human brain measured by 3D PET. *Int J Neuropsychopharmacol* 2:73–82
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. The MIT Press, Cambridge, Massachusetts
- Swart JC, Frobose MI, Cook JL, Geurts DE, Frank MJ, Cools R, den Ouden HE (2017) Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife*, 6. <https://doi.org/10.7554/eLife.22169.001>
- Swart JC, Frank MJ, Maatta JI, Jensen O, Cools R, den Ouden HEM (2018) Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLoS Biol* 16:e2005979
- Thompson J, Thomas N, Singleton A, Piggott M, Lloyd S, Perry EK, Morris CM, Perry RH, Ferrier IN, Court JA (1997) D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. *Pharmacogenetics* 7:479–484
- Tunbridge EM (2010) The catechol-O-methyltransferase gene: its regulation and polymorphisms. *Int Rev Neurobiol* 95:7–27

- van Nuland AJ, Helmich RC, Dirkx MF, Zach H, Toni I, Cools R, den Ouden HEM (2020) Effects of dopamine on reinforcement learning in Parkinson's disease depend on motor phenotype. *Brain: J Neurol* 143:3422–3434
- Wang F, Simen A, Arias A, Lu QW, Zhang H (2013) A large-scale meta-analysis of the association between the ANKK1/DRD2 Taq1A polymorphism and alcohol dependence. *Hum Genet* 132:347–358
- Weinberger DR, Egan MF, Bertolino A, Callicott JH, Mattay VS, Lipska BK, Berman KF, Goldberg TE (2001) Prefrontal neurons and the genetics of schizophrenia. *Biol Psychiat* 50:825–844
- Wickens JR, Reynolds JN, Hyland BI (2003) Neural mechanisms of reward-related motor learning. *Curr Opin Neurobiol* 13:685–690
- Wickens JR, Budd CS, Hyland BI, Arbuthnott GW (2007) Striatal contributions to reward and decision making: making sense of regional variations in a reiterated processing matrix. *Ann N Y Acad Sci* 1104:192–212
- Wimber M, Schott BH, Wendler F, Seidenbecher CI, Behnisch G, Macharadze T, Bauml KH, Richardson-Klavehn A (2011) Prefrontal dopamine and the dynamic control of human long-term memory. *Transl Psychiatry* 1:e15
- Wolf ME, Roth RH (1990) Autoreceptor regulation of dopamine synthesis. *Ann N Y Acad Sci* 604:323–343
- Yao J, Pan YQ, Ding M, Pang H, Wang BJ (2015) Association between DRD2 (rs1799732 and rs1801028) and ANKK1 (rs1800497) polymorphisms and schizophrenia: a meta-analysis. *Am J Med Genet Part B, Neuropsychiatr Genet: off Publ Int Soc Psychiatr Genet* 168B:1–13
- Yavich L, Forsberg MM, Karayiorgou M, Gogos JA, Mannisto PT (2007) Site-specific role of catechol-O-methyltransferase in dopamine overflow within prefrontal cortex and dorsal striatum. *J Neurosci: off J Soc Neurosci* 27:10196–10209
- Zhang Y, Bertolino A, Fazio L, Blasi G, Rampino A, Romano R, Lee ML, Xiao T, Papp A, Wang D, Sadee W (2007) Polymorphisms in human dopamine D2 receptor gene affect gene expression, splicing, and neuronal activity during working memory. *Proc Natl Acad Sci USA* 104:20552–20557

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.