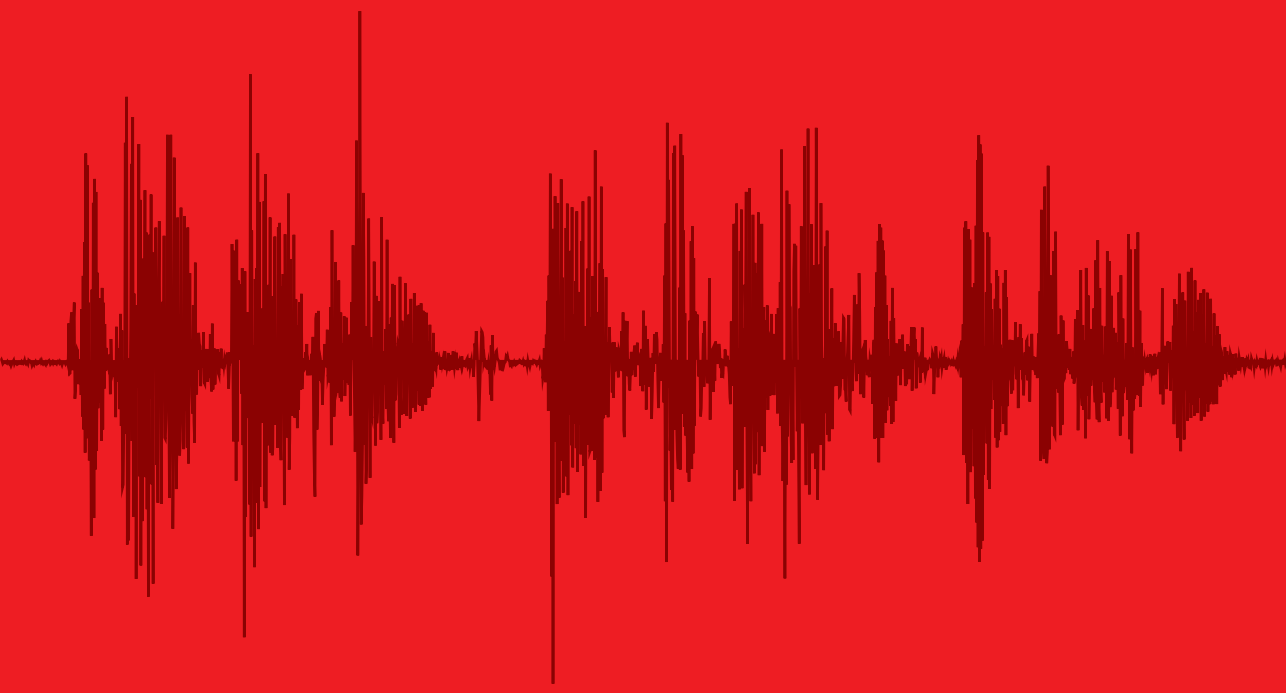
To go or not to go?

On motivational biases in decision-making



Jennifer C. Swart

DONDERS
SERIES

To go or not to go?

On motivational biases in decision-making

Jennifer C. Swart

ISBN: 978-94-6284-173-4

Design: www.proefschriftenprinten.nl
Printed by: Print Service Ede - www.proefschriftenprinten.nl

Support:
The work in this thesis was funded by a Research Talent grant (NWO-406-14-028) awarded to Jennifer C. Swart by the Dutch Organisation for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO).

Copyright © Jennifer C. Swart, 2019.
All rights are reserved. No part of this thesis may be reproduced, distributed, or transmitted in any form without prior written permission of the author.

To go or not to go?

On motivational biases in decision-making

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 29 maart 2019
om 10.30 uur precies

door

Jennifer Claudia Swart
geboren op 1 december 1989
te Zwijndrecht

Promotoren
Prof. dr. Roshan Cools
Prof. dr. Michael J. Frank (Brown University, Providence, Verenigde Staten)

Copromotoren
Dr. Hanneke E.M. den Ouden

Manuscriptcommissie
Prof. dr. Roy P.C. Kessels
Prof. dr. Peter Dayan
Prof. dr. Karin Roelofs

All men by nature desire knowledge.

- Aristotle

Contents

Chapter 1

General introduction

I am just a child who has never grown up. I still keep asking these 'how' and 'why' questions. Occasionally, I find an answer.
- Stephen Hawking

Partly published as:
Van Driel J*, Swart JC*, Egner T, Ridderinkhof KR, & Cohen MX (2015). (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. Cognitive, Affective, and Behavioral Neuroscience, 15:787-807. *Shared first authorship.

Motivation

Motivation is a hallmark of human cognition and strongly drives our behaviour. Some days we are highly motivated and achieve a lot (such as going to work, the gym, finishing household chores, ánd socializing with friends), whereas we seem to lack that motivation at other days and don't achieve much more than watching Netflix. In many cases motivation is beneficial for the realization of our desired goals, as it generally facilitates the behaviour that is required for achieving those goals. We often take these motivational drives for granted, but these drives actually depend on a delicate interplay of neural mechanisms that can go awry as well; motivation can drive our actions too much, as for example in addiction and impulsivity, or drive our actions too little, as for example in apathy and depression. The occurrence of motivational disorders has increased in modern society and is predicted to further increase during this century (www.volksgezondheidenzorg.info). These developments urge the field of cognitive neuroscience to address how our motivations drive our actions on the one hand, and address how we can overcome dysfunctional motivational drives on the other hand. In this thesis, I set out to address these questions and shed light on the neurocomputational mechanisms involved in driving vs. regulating motivated action. In the remainder of this chapter, I first focus on motivational control over actions and the associated neural underpinnings, and secondly focus on the neural mechanisms that allow us to reduce the impact of motivational drives when these drives become dysfunctional.

Motivational biasing of action

Day in, day out, we pursue a vast number of goals. These goals can be relatively short-term, such as obtaining a delicious piece of pie, or relatively long-term, such as completing this PhD thesis. When pursuing our goals, we generally try to act in a manner that leads to desirable outcomes, whereas we try to avoid acting in ways that lead to unwanted outcomes. These motivational drives have long been known to guide our actions accordingly; Anticipated rewards tend to facilitate taking action (Cools et al., 2011; Duffy, 1962; Estes, 1943; Estes and Skinner, 1941; Guitart-Masip et al., 2012), whereas anticipated punishments tend to facilitate holding back (Davis and Wright, 1979; Geurts et al., 2013; Huys et al., 2011). In other words, the anticipation of positive/rewarding and negative/punishing outcomes has an opposing effect on the activation of our behaviour, which we refer to as *a motivational bias (in action)* throughout this thesis (Figure 1). Remarkably, this motivational biasing of our behaviour even occurs while we have the feeling of being completely in control of our behaviour.

Figure 1. **Motivational biasing of action.**
The anticipated motivational valence (reward vs. punishment) strongly guides the activation of our behaviour (action vs. inaction), making it easy for example to approach and grab tasty looking food and to stay away from rotten food. These situations are indicated in colour above, and throughout this thesis I refer to these situations as 'motivationally congruent'. The coupling of action to reward and inaction to punishment is beneficial in many cases, yet for some situations it is more beneficial to do the opposite, such as taking out rotten food to avoid a smelly house and waiting (longer) to get the best lasagne. Our motivational biases make those situations (indicated by the grey-scale) harder, and therefore I refer to these situations as 'motivationally incongruent' throughout this thesis.

The coupling of action to reward and inaction to punishment is thought to reflect the statistics of the environment (Dayan et al., 2006), such that taking action generally leads to positive outcomes (e.g., when approaching tasty food) and holding back generally avoids negative outcomes (e.g., when staying away from rotten food). Therefore, the hardwired, or 'Pavlovian' shaping of our behaviour is thought to be highly functional as it can serve to reduce cognitive computational load while generally still promoting the appropriate level of (in)action. However, hardwired activation can by its inherent global nature not be informative of which *specific* action is most optimal, particularly not in an ever-changing environment. For such adaptive action selection, a more flexible and computational costly control system is required, namely the instrumental learning system. The instrumental learning system adaptively learns action-outcome contingencies by assigning obtained rewards and punishments to the specific actions that elicited the outcomes (Dickinson and Balleine, 1994; Rescorla and Wagner, 1972; Robbins and Everitt, 2007). Thus, the value of a specific action in a given environment (e.g., ordering red wine in your favourite restaurant) can become more positive if that action is followed by a rewarding outcome (you receive wine that turns out to be delicious), promoting the repetition of that action in the environment ('*I'll have the red wine again!*'). Conversely, the

value of ordering red wine would have decreased if it was followed by an aversive outcome, such as the wine having gone sour, making you less likely to repeat that action (*'I'll try a water instead!'*). By tracking and updating the value of actions in the environment, the instrumental system is adaptive, but also requires more experience and is computational costlier than the innate or 'hardwired' control system. What has not been studied before, however, is whether the instrumental system might also itself be biased in coupling action to reward and inaction to punishment in order to speed up the learning of likely action-outcome associations. I address this question in chapter 3 and 4, where we developed an experimental paradigm and mathematical models to disentangle motivational biases in the instrumental control system from motivational biases in the Pavlovian system.

Neural mechanisms underlying motivational biases in action

The key neural candidate for the motivational biasing of action is the striatum, a phylogenetically (or, 'evolutionary') old brain structure that lies deep in the brain (Figure 2), and particularly its modulation by the neurochemical dopamine (Boureau and Dayan, 2011; Cools et al., 2011). One of the first studies linking striatal dopamine to motivated action showed that increasing dopamine in the ventral part of the striatum (i.e., nucleus accumbens) enhanced the behavioural activation elicited by reward cues (Taylor and Robbins, 1984). This dopaminergic enhancement was attenuated after lesioning of the ventral striatum (Taylor and Robbins, 1986). Originally, these findings were often interpreted to specifically highlight a role for striatal dopamine in pure reward processing, whereas nowadays these results are additionally interpreted as reflecting a role for striatal dopamine in the coupling of reward to behavioural activation.

Next to promoting behavioural activation in the face of reward, dopamine is also strongly tied to reward-based *learning* as demonstrated in perhaps one of the most famous neuroscientific studies by Schulz and colleagues (1998, 1997). Schulz et al. recorded neural activity in midbrain dopamine neurons, which are known to project to the striatum among others. The activity of these dopamine neurons increased after unpredicted rewards, in other words, when events turned out better than expected. Once animals learned to associate predictive cues with reward outcomes, the activity of dopamine neurons shifted from the reward onset to the cue onset, as now these reward predicting cues became the first indication that events were better than expected. This work and subsequent theorizing (Montague et al., 1996) highlighted the role of midbrain dopamine in reward prediction errors, which are used for reward learning. To elaborate, when an outcome is the same as you predicted, there is no need to update your beliefs as your beliefs appropriately reflect reality. When the actual outcome, however, is better or worse than you expected, it is useful to update your beliefs with this discrepancy between your prediction and reality (i.e. the

reward prediction error), such that your prediction will better reflect reality the next time. The work of Schultz et al. implicated dopamine in learning from reward prediction errors, and showed that midbrain dopamine neurons become active towards both unexpected reward outcomes, and reward predicting cues. These midbrain neurons project to the striatum (called mesolimbic projections) and release dopamine in the striatum.

Importantly, the striatum does not function in isolation and is part of a set of nuclei working in concert, together called the basal ganglia (Figure 2). Current models of basal ganglia function (Collins and Frank, 2015a, 2014; Frank, 2005; Frank and O'Reilly, 2006; Lloyd and Dayan, 2016) provide the theoretical basis of our work and therefore I will provide a brief overview of the basal ganglia functioning in the remainder of this section and cover the basal ganglia models and predictions in the relevant chapters (2, 3, and 5) as well.

The basal ganglia are connected through a direct 'Go' and indirect 'NoGo' pathway (Figure 2), both projecting to (among others) the motor cortex, which is responsible for motor responses. The direct 'Go' pathway (through double inhibition) has a net effect of activating the motor cortex, thereby promoting behavioural activation (DeLong and Wichmann, 2007; Mink and Thach, 1991). The indirect 'NoGo' pathway contains an additional inhibitory step, creating an inhibitory net effect of the motor cortex, thereby promoting behavioural inhibition. Crucially, mesolimbic dopamine (i.e., the dopamine released by midbrain neurons in the striatum) modulates the activity of these direct and indirect pathways. Mesolimbic dopamine release potentiates activity in the direct pathway through D1 receptors (Hernandez-Lopez et al., 1997), whereas dips in mesolimbic dopamine release potentiate activity in the indirect pathway through D2 receptors (Hernandez-Lopez et al., 2000). Putting it all together, reward cues and outcomes elicit peaks in mesolimbic dopamine release, which potentiate the direct 'Go' pathway and thereby promote behavioural activation. In contrast, punishment cues and outcomes are known to elicit dips in mesolimbic dopamine release, which potentiate the indirect 'NoGo' pathway and thereby promote behavioural inhibition. Altogether, this complex interplay of neural projections provides a prime candidate mechanism for the motivational biasing of behavioural activation.

In this thesis, I present multiple experiments that allowed us to test several predictions derived from the basal ganglia dopamine model in relation to the motivational biases in action: We test whether the motivational biases are reduced in a human genetic dopamine (partial) knock-out population (chapter 2). We causally assess whether these biases are affected by direct stimulation of the human ventral striatum (chapter 5) and by a pharmacological dopaminergic challenge (chapter 3). Moreover, we theorize that cue- and outcome-based dopamine bursts provide a neural substrate for the cue-based, Pavlovian biases and outcome-based, instrumental learning biases in action respectively. To this end, we also test whether the dopamine challenge affects both the Pavlovian biases and instrumental learning biases (chapter 3).

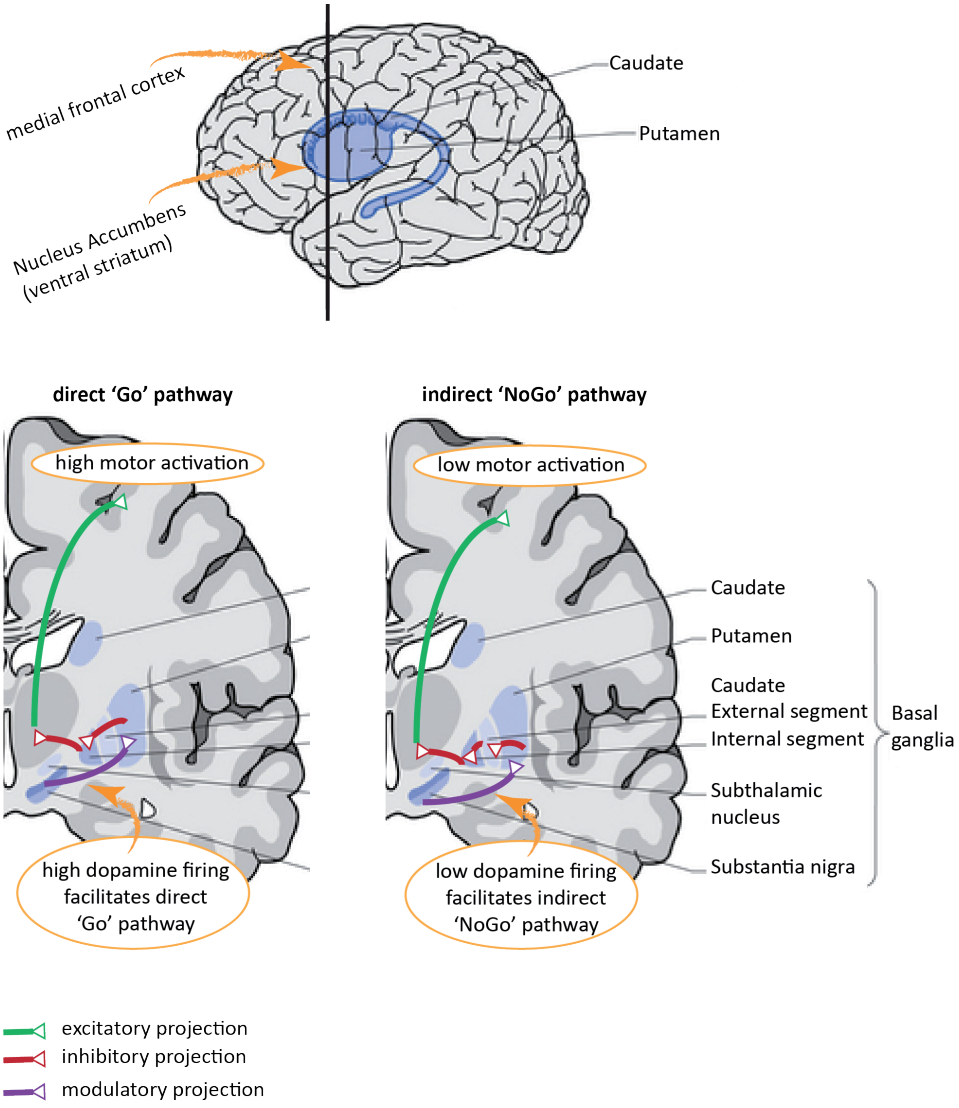


Figure 2. **Medial frontal cortex and the basal ganglia.**
Top: Sagittal view of a schematic brain. Arrows indicate the 'evolutionary younger' medial frontal cortex and the 'evolutionary older' ventral striatum. The black line indicates roughly the location of the coronal slices shown below. Bottom: The basal ganglia direct (left) and indirect (right) pathways displayed for one hemisphere. Other basal ganglia connections are discarded for simplicity. On the left side, the direct 'Go' pathway is presented. The Go pathway promotes behavioural activation. This pathway becomes stronger when much dopamine is released in the striatum, for example when expecting reward. On the right side, the indirect 'NoGo' pathway is presented. The NoGo pathway promotes behavioural inaction. This pathway 'becomes stronger' when little dopamine is released in the striatum, for example when anticipating punishment. The Go and NoGo pathways provide a neural candidate substrate for the motivational biasing of action. Adapted from iKnowledge (clinicalgate.com/the-basal-ganglia).

Control over maladaptive motivational biases

As indicated above, the observed motivational biases in our behaviour might be quite functional as these biases are thought to follow the statistics of our environment (Dayan et al., 2006). Put simply, rewards might generally require us to take action in order for us to receive these rewards, whereas we might miss out on the rewards when we hold back. Instead, punishments generally require us to withhold from taking action in order to avoid these negative outcomes, whereas taking action might increase the risk of encountering the punishment. As such, it is likely to be highly beneficial for the survival of a species to couple taking action to reward contexts and couple behavioural inhibition to aversive contexts.

Although it is often beneficial to take action for reward, it is certainly not the case that getting reward *always* requires us to taking action (Figure 1). Sometimes it can more beneficial to wait and withhold from responding, for example when investing on the stock market. Conversely, punishment contexts might sometimes require us to take action in order to avoid aversive outcomes, for example working harder to avoid getting fired. In those cases, our motivational biases actually work against us and make it harder to behave adequately. The persistence of the biases was demonstrated in a striking example using baby chicks (Hershberger, 1986). In the Hershberger experiment, food-deprived baby chicks were placed in a runway containing a food cup that moved with half vs. twice the speed of the chicks. Thus, when the food cup moved with half the speed, chicks needed to move forwards to catch up with the food cup; When the food cup moved with twice the speed, chicks needed to move backwards for the food cup to catch up with them. While the chicks were perfectly able to approach the rewarding food cup, they were completely unable to move away from the food cup when they needed to. This experiment demonstrates that once the environmental requirements conflict with our innate motivational biases, we need to exert control over our behaviour (which can be extremely difficult) in order to overcome the now suddenly *maladaptive* motivational biases. Fortunately, in contrast to the Hershberger chicks, humans are often able to overcome our motivational drives when these conflict with our environmental requirements.

The brain region that has primarily been linked to our ability to overcome dysfunctional motivational biases, is the medial frontal cortex (Cavanagh et al., 2013). The frontal cortex is a phylogenetically new brain region that is particularly well-developed in humans (Teffer and Semendeferi, 2012). The exact mechanisms by which the midfrontal cortex reduces maladaptive motivational biases, however, still need to be elucidated. Here, I propose that the regulation of maladaptive motivational biases might rely on similar neural mechanisms that are evident in the classic cognitive control literature, as described in my previous work (van Driel, Swart, et al., 2015):

"Cognitive control refers to a set of mental capacities devoted to optimize goal-directed behaviour in situations of multiple competing response alternatives (Botvinick et al., 2001; Ridderinkhof et al., 2011; Ridderinkhof et al., 2004). Neuroscience has tied these

adaptive control functions to processes in frontal brain networks (Fuster, 2000; Miller, 2000), where the medial frontal cortex (MFC) is thought to signal the need for control in response to challenging situations (Alexander and Brown, 2011; Botvinick et al., 2004; Ito et al., 2003), which is communicated to the dorsolateral prefrontal cortex (DLPFC; MacDonald et al., 2000). Both of these regions exert top-down influence over lower, task-related sensorimotor processing (Cohen et al., 2009; Danielmeier et al., 2011; Egner and Hirsch, 2005; Miller and D'Esposito, 2005), in order to adjust future behaviour (Kerns et al., 2004). Cognitive electrophysiology has provided compelling evidence of theta-band (3- to 8-Hz) oscillatory activity as the underlying "language" of communication within this network (see Cavanagh and Frank, 2014, and Cohen, 2014, for reviews), where the MFC has been proposed to be a "hub" for theta phase-synchronized information transfer (Cohen, 2011).

Cognitive control is a transient response, waxing and waning depending on the presence or absence of risks or demands such as response conflict. Indeed, because frontally mediated cognitive control is effortful, it is inefficient to recruit these mechanisms continuously (Ridderinkhof et al., 2004). Here, conflict is defined as the incongruence between a task-relevant learned response and a task-irrelevant stimulus feature, which results in slower and more error-prone behaviour relative to nonconflict (the "conflict effect"). […] Importantly, these trial-to-trial fluctuations in behavioural conflict effects have been shown to covary with trial-to-trial variability in midfrontal theta activity (Cohen and Cavanagh, 2011). […] Our general EEG results of increased frontal theta power as well as interregional phase synchrony after conflict are in accordance with a growing body of findings that have tied frontal theta-band activity to various cognitive control processes, including conflict adaptation (Cohen and Cavanagh, 2011; Pastötter et al., 2013), error processing (Luu et al., 2004; van Driel et al., 2012), task switching (Cunillera et al., 2012), and reinforcement learning (Cavanagh et al., 2010; van de Vijver et al., 2011). […]

From an anatomical perspective, mid–lateral frontal theta synchrony has been proposed to reflect MFC–DLPFC functional connectivity, which increases after conflict has been encountered (Cohen and Ridderinkhof, 2013). The current axiom in the cognitive control literature is that the MFC monitors for possible instances of conflict, and upon conflict detection, communicates the need for increased control to the DLPFC, which further implements control through top-down signals to motor and task-relevant sensory areas (Botvinick et al., 1999, 2004; Kerns et al., 2004; MacDonald et al., 2000; Ridderinkhof et al., 2011; Ridderinkhof et al., 2004; Ridderinkhof et al., 2004). However, direct regulatory top-down signals from MFC to guide behaviour in situations of conflict have also been observed (Cohen et al., 2009; Danielmeier et al., 2011; Kennerley et al., 2006; Ridderinkhof et al., 2004), suggesting a more integrative function of the MFC (Shenhav et al., 2013)."

To recapitulate, classic cognitive control models implicate the medial frontal cortex in detecting conflict between competing response options and signalling the need for control to a network of task-related regions. Accordingly, the decision threshold can be adjusted (Cavanagh et al., 2011) in order to prevent prepotent, impulsive responses, and allow for goal-directed responses to take over. Here I propose that the medial frontal cortex is similarly involved in detecting conflict between prepotent, Pavlovian responses and goal-directed, instrumental responses, and signalling the increased need for control to task-related sites. Accordingly, the decision threshold can be adjusted to prevent Pavlovian responses and allow for instrumental behaviour. We assess the presence of these hypothesized neural mechanisms in a healthy student population (chapter 4) and subsequently test whether reducing the Pavlovian biases with neuromodulation reduces the neural responses to Pavlovian conflict (chapter 5).

Aims and outline of this thesis

The aim of this thesis is twofold. The first aim is to provide a deeper understanding of the neurocomputational mechanisms that give rise to the motivational biasing of action, focusing particularly on dopamine function in the basal ganglia. The second aim is to provide a better understanding of the neurocomputational mechanisms involved in regulating motivational drives when they become dysfunctional, allowing for adequate behaviour, focusing particularly on medial frontal cortex functioning. To achieve these aims, we have conducted the four experiments presented in this thesis.

First we assessed whether the motivational biases in action are altered in a highly unique genetic population, namely carriers of a dopamine-related pathogenic genetic variant (or, 'mutation'), which putatively results in reduced dopamine function. We compared the motivational biases of this genetic group with a control group in **chapter 2**.

We continued to assess the role of dopamine in **chapter 3**, where we directly manipulated the dopamine system with a (non-selective) dopaminergic drug, namely methylphenidate ('Ritalin'). In this study, we first developed novel computational models that allowed us to disentangle hardwired, 'Pavlovian' biases from motivational biases in the learning system. We then assessed the effect of methylphenidate on these Pavlovian and instrumental biases. Moreover, we tested whether individual differences in the effects of methylphenidate could be predicted by proxy-measures of baseline dopamine function.

Once we established the Pavlovian and instrumental mechanisms driving motivational biases, we continued with uncovering neural mechanisms related to the reduction of *maladaptive* biases. In **chapter 4**, we assessed whether the medial frontal signals are related to reduction of these biases using electroencephalography (EEG). Specifically, we assessed whether the local, medial frontal activity might be related to the level of Pavlovian conflict, whereas the network-wide functional connectivity might be more directly related to the

suppression of Pavlovian response tendencies.

In **chapter 5** we assessed the causal role of the human ventral striatum in the motivational biasing of action. To this end, we directly stimulated the ventral striatum with deep brain stimulation (DBS) in a psychiatric population (obsessive-compulsive disorder patients) that receive DBS as part of their treatment. Here, we tested whether the disruption of striatal communication disrupts the motivational biases, and whether this disruption of the biases consequently reduces the conflict-related neural signatures over the medial frontal cortex.

Taken together, this thesis presents a state-of-the-art multidisciplinary approach, combining causal interventions (pharmacology, DBS), sophisticated behavioural paradigms, computational modelling, neuroimaging, genetics, and clinical work, to further our understanding of the neurocomputational mechanisms of motivational biases in decision-making. In **chapter 6** I provide a summary of the main findings, discuss and integrate the most relevant findings of this thesis, and highlight future directions.

References

Alexander, W.H., Brown, J.W., 2011. Medial prefrontal cortex as an action-outcome predictor. Nat. Neurosci. 14, 1338–1344. doi:10.1038/nn.2921

Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402, 179–181. doi:10.1038/46035

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. Psychol Rev. 108, 624–652. doi:10.1037/0033-295X.108.3.624

Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: An update. Trends Cogn. Sci. doi:10.1016/j.tics.2004.10.003

Boureau, Y.L., Dayan, P., 2011. Opponency revisited: Competition and cooperation between dopamine and serotonin. Neuropsychopharmacology. doi:10.1038/npp.2010.151

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., Frank, M.J., 2013. Frontal Theta Overrides Pavlovian Learning Biases. J. Neurosci. 33, 8541–8548. doi:10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. doi:10.1016/j.tics.2014.04.012

Cavanagh, J.F., Frank, M.J., Klein, T.J., Allen, J.J.B., 2010. Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. Neuroimage 49, 3198–3209. doi:10.1016/j.neuroimage.2009.11.080

Cavanagh, J.F., Wiecki, T.V, Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., Frank, M.J., 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat. Neurosci. 14, 1462–1467. doi:10.1038/nn.2925

Cohen, M.X., 2014. A neural microcircuit for cognitive conflict detection and signaling. Trends Neurosci. doi:10.1016/j.tins.2014.06.004

Cohen, M.X., 2011. Error-related medial frontal theta activity predicts cingulate-related structural connectivity. Neuroimage 55, 1373–1383. doi:10.1016/j.neuroimage.2010.12.072

Cohen, M.X., Cavanagh, J.F., 2011. Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. Front. Psychol. 2. doi:10.3389/fpsyg.2011.00030

Cohen, M.X., Ridderinkhof, K.R., 2013. EEG Source Reconstruction Reveals Frontal-Parietal Dynamics of Spatial Conflict Processing. PLoS One 8. doi:10.1371/journal.pone.0057293

Cohen, M.X., van Gaal, S., Ridderinkhof, K.R., Lamme, V.A.F., 2009. Unconscious errors enhance prefrontal-occipital oscillatory synchrony. Front. Hum. Neurosci. 3. doi:10.3389/neuro.09.054.2009

Collins, A.G.E., Frank, M.J., 2015. Surprise! Dopamine signals mix action, value and error. Nat. Neurosci. 19, 3–5. doi:10.1038/nn.4207

Collins, A.G.E., Frank, M.J., 2014. Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. 121, 337–66. doi:10.1037/a0037015

Cools, R., Nakamura, K., Daw, N.D., 2011. Serotonin and dopamine: Unifying affective, activational, and decision functions. Neuropsychopharmacology. doi:10.1038/npp.2010.121

Cunillera, T., Fuentemilla, L., Periañez, J., Marco-Pallarès, J., Krämer, U.M., Càmara, E., Münte, T.F., Antoni, R.F., 2012. Brain oscillatory activity associated with task switching and feedback processing. Cogn. Affect. Behav. Neurosci. 12, 16–33. doi:10.3758/s13415-011-0075-5

Danielmeier, C., Eichele, T., Forstmann, B.U., Tittgemeyer, M., Ullsperger, M., 2011. Posterior Medial Frontal Cortex Activity Predicts Post-Error Adaptations in Task-Related Visual and Motor Areas. J Neurosci 31, 1780–1789. doi:31/5/1780 [pii]\n10.1523/JNEUROSCI.4299-10.2011

Davis, H., Wright, J., 1979. procedural and parametric variability in studies of conditioned suppression. Bull. Psychon. Soc. 14, 149–150.

Dayan, P., Niv, Y., Seymour, B., D. Daw, N., 2006. The misbehavior of value and the discipline of the will. Neural Networks 19, 1153–1160. doi:10.1016/j.neunet.2006.03.002

DeLong, M.R., Wichmann, T., 2007. Circuits and circuit disorders of the basal ganglia. Arch. Neurol. 64, 20–24. doi:10.1001/archneur.64.1.20

Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18. doi:10.3758/BF03199951

Duffy, E., 1962. Activation and behavior. New York.

Egner, T., Hirsch, J., 2005. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat. Neurosci. 8, 1784–1790. doi:10.1038/nn1594

Estes, W.K., 1943. Discriminative conditioning. I. A discriminative property of conditioned anticipation. J. Exp. Psychol. 32, 150–155. doi:10.1037/h0058316

Estes, W.K., Skinner, B.F., 1941. Some quantitative properties of anxiety. J. Exp. Psychol. 29, 390–400. doi:10.1037/h0062283

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. 17, 51–72. doi:10.1162/0898929052880093

Frank, M.J., O'Reilly, R.C., 2006. A Mechanistic Account of Striatal Dopamine Function in Human Cognition: Psychopharmacological Studies With Cabergoline and Haloperidol. Behav. Neurosci. 120, 497–517. doi:10.1037/0735-7044.120.3.497

Fuster, J.M., 2000. Executive frontal functions. Exp. Brain Res. doi:10.1007/s002210000401

Geurts, D.E.M., Huys, Q.J.M., Den Ouden, H.E.M., Cools, R., 2013. Aversive Pavlovian control of instrumental behavior in humans. J. Cogn. Neurosci. 25, 1428–1441.

Guitart-Masip, M., Huys, Q.J.M., Fuentemilla, L., Dayan, P., Duzel, E., Dolan, R.J., 2012. Go and no-go learning in reward and punishment: Interactions between affect and effect. Neuroimage 62, 154–166. doi:10.1016/j.neuroimage.2012.04.024

Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A., Galarraga, E., 1997. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. J Neurosci 17, 3334–3342.

Hernandez-Lopez, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., Surmeier, D.J., 2000. D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLC[beta]1-IP3-calcineurin-signaling cascade. J. Neurosci. 20, 8987–8995. doi:20/24/8987 [pii]

Hershberger, W.A., 1986. An approach through the looking-glass. Anim. Learn. Behav. 14, 443–451. doi:10.3758/BF03200092

Huys, Q.J.M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R.J., Dayan, P., 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. PLoS Comput. Biol. 7. doi:10.1371/journal.pcbi.1002028

Ito, S., Stuphorn, V., Brown, J.W., Schall, J.D., 2003. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science (80-.). 302, 120–122. doi:10.1126/science.1087847

Kennerley, S.W., Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rushworth, M.F.S., 2006. Optimal decision making and the anterior cingulate cortex. Nat Neurosci 9, 940–7. doi:10.1038/nn1724

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior Cingulate Conflict Monitoring and Adjustments in Control. Science (80-.). 303, 1023–1026. doi:10.1126/science.1089910

Lloyd, K., Dayan, P., 2016. Safety out of control: dopamine and defence. Behav. Brain Funct. 12, 15. doi:10.1186/s12993-016-0099-7

Luu, P., Tucker, D.M., Makeig, S., 2004. Frontal midline theta and the error-related negativity: Neurophysiological mechanisms of action regulation. Clin. Neurophysiol. 115, 1821–1835. doi:10.1016/j.clinph.2004.03.031

MacDonald, a W., Cohen, J.D., Stenger, V. a, Carter, C.S., 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288, 1835–1838. doi:10.1126/science.288.5472.1835

Miller, B.T., D'Esposito, M., 2005. Searching for "the top" in top-down control. Neuron. doi:10.1016/j.neuron.2005.11.002

Miller, E.K., 2000. The prefrontal cortex and cognitive control. Nat. Rev. Neurosci. 1, 59–65. doi:10.1038/35036228

Mink, J.W., Thach, W.T., 1991. Basal ganglia motor control. III. Pallidal ablation: normal reaction time, muscle cocontraction, and slow movement. J. Neurophysiol. 65, 330–351.

Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic predictive Hebbian learning. J. Neurosci. 76, 1936–1947. doi:10.1.1.156.635

Pastötter, B., Dreisbach, G., Bäuml, K.T., 2013. Dynamic adjustments of cognitive control: oscillatory correlates of the conflict adaptation effect. J. Cogn. Neurosci. 25, 2167–78. doi:10.1162/jocn_a_00474

Rescorla, R. a, Wagner, a R., 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Class. Cond. II Curr. Res. Theory 21, 64–99. doi:10.1101/gr.110528.110

Richard Ridderinkhof, K., Forstmann, B.U., Wylie, S.A., Burle, B., van den Wildenberg, W.P.M., 2011. Neurocognitive mechanisms of action control: Resisting the call of the Sirens. Wiley Interdiscip. Rev. Cogn. Sci. 2, 174–192. doi:10.1002/wcs.99

Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S., 2004. The role of the medial frontal cortex in cognitive control. Science (80-.). 306, 443–447. doi:10.1126/science.1100301

Ridderinkhof, K.R., Van Den Wildenberg, W.P.M., Segalowitz, S.J., Carter, C.S., 2004. Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn. 56, 129–140. doi:10.1016/j.bandc.2004.09.016

Robbins, T.W., Everitt, B.J., 2007. A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). Psychopharmacology (Berl). doi:10.1007/s00213-006-0528-7

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science (80-.). 275, 1593–1599. doi:10.1126/science.275.5306.1593

Schultz, W., Tremblay, L., Hollerman, J.R., 1998. Reward prediction in primate basal ganglia and frontal cortex, in: Neuropharmacology. pp. 421–429. doi:10.1016/S0028-3908(98)00071-9

Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The expected value of control: An integrative theory of anterior cingulate cortex function. Neuron. doi:10.1016/j.neuron.2013.07.007

Taylor, J.R., Robbins, T.W., 1986. 6-Hydroxydopamine lesions of the nucleus accumbens, but not of the caudate nucleus, attenuate enhanced responding with reward-related stimuli produced by intra-accumbens d-amphetamine. Psychopharmacology (Berl). 90, 390–397. doi:10.1007/BF00179197

Taylor, J.R., Robbins, T.W., 1984. Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. Psychopharmacology (Berl). 84, 405–412. doi:10.1007/BF00555222

Teffer, K., Semendeferi, K., 2012. Human prefrontal cortex. Evolution, development, and pathology. Prog. Brain Res. 195, 191–218. doi:10.1016/B978-0-444-53860-4.00009-X

van de Vijver, I., Ridderinkhof, K.R., Cohen, M.X., 2011. Frontal oscillatory dynamics predict feedback learning and action adjustment. J. Cogn. Neurosci. 23, 4106–4121. doi:10.1162/jocn_a_00110

van Driel, J., Ridderinkhof, K.R., Cohen, M.X., 2012. Not All Errors Are Alike: Theta and Alpha EEG Dynamics Relate to Differences in Error-Processing Dynamics. J. Neurosci. 32, 16795–16806. doi:10.1523/JNEUROSCI.0802-12.2012

van Driel, J., Swart, J.C., Egner, T., Ridderinkhof, K.R., Cohen, M.X., 2015. (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. Cogn. Affect. Behav. Neurosci. 15, 787–807. doi:10.3758/s13415-015-0367-2

Chapter 2

Effects of carrying a pathogenic variant in the tyrosine hydroxylase gene on motivated action and valuation: A pilot study in family members with tyrosine hydroxylase deficiency

The very essence of instinct is that it's followed independently of reason.
- Charles Darwin

Contributing authors:
Swart JC*, Froböse MI*, Wassenberg T, den Ouden HEM, Willemsen MA, & Cools R. **Shared first authorship.*

Abstract

Catecholamines (particularly dopamine) have long been implicated in motivation, learning and behavioural activation. Benign variants in dopamine-regulating genes have widely been linked to these processes as well, yet the cognitive effects of carrying pathogenic variants in the gene coding for tyrosine hydroxylase, which transforms tyrosine into dopamine's direct precursor L-Dopa, have never been studied. Here, we assessed for the first time whether carriers of tyrosine hydroxylase deficiency (THD) show altered motivated action due to putative reductions in dopamine synthesis. To this end, we employed a motivational Go/NoGo learning task, which is sensitive to manipulations in dopamine function and compared 16 family members of THD patients with 20 education- and age-matched controls. In the first learning phase of this task, subjects learnt to make Go or NoGo responses to cues that predict reward vs. punishment. In the second transfer phase, the subjects were presented with pairs of cues and chose the one they preferred, in the absence of reinforcement. Cue valence strongly biased Go/NoGo responding in the learning phase, such that subjects made more Go responses to reward than punishment cues. The groups did not significantly differ in this motivational bias. However, the THD carriers exhibited a shift in preference from NoGo-to-Win to Go-to-Avoid cues relative to matched controls during the transfer phase. These results suggest that subjective valuation is altered in THD carriers, potentially due to catecholamine-dependent changes in reward expectations, whereas task performance was unaffected. This pilot study provides a first insight into the cognitive consequences of carrying pathogenic TH variants, focusing on alterations in the reward valuation system and motivational biases in action.

Introduction

The catecholamines (particularly dopamine) have long been known to play a role in motivational and cognitive functions (Brozoski et al., 1979; Schultz et al., 1997), such as motivation, learning, and behavioural activation and vigour (Berridge and Robinson, 1998; Cools et al., 2009; Frank et al., 2004; Robbins and Everitt, 2007; Salamone et al., 2005). Several hereditary neurometabolic disorders affecting synthesis, breakdown, and transport of the catecholamines have been described (Kurian et al., 2011), including tyrosine hydroxylase deficiency (THD). THD is an extremely rare autosomal recessive disorder in which tyrosine hydroxylase, i.e. the rate limiting step in catecholamine synthesis, is impaired (Bräutigam et al., 1999; Willemsen et al., 2010), see Figure 1. THD leads to neurological symptoms, ranging from mild motor distortions to severe, early onset encephalopathy and can be treated with L-Dopa supplementation (Willemsen et al., 2010). Although benign variants in dopamine-regulating genes, such as the dopamine transporter polymorphism (for which the functional consequences are less clear) have widely been linked to motivation, learning, and action (Frank and Fossella, 2011), the effect of carrying a pathogenic variant in the TH gene (for which the functional consequences are more severe)

on these processes has not been studied before. Here we assess for the first time the cognitive consequences of carrying pathogenic TH variants in relatives of THD patients.

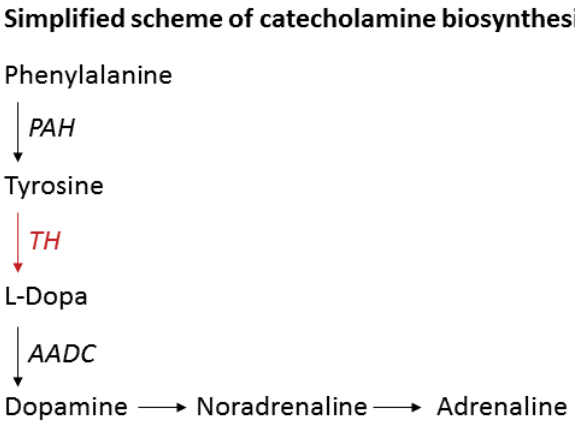


Figure 1. **Simplified scheme of the biosynthesis of the catecholamines dopamine, noradrenaline and adrenaline.**
Tyrosine hydroxylase deficiency (THD) affects the catecholamine synthesis by impairing enzymatic functioning of tyrosine hydroxylase (TH; marked in red) (Cansev and Wurtman, 2007), which catalyses the transformation of tyrosine into 3,4-dihydroxyphenylalanine (L-Dopa), the direct precursor of dopamine (Kurian et al., 2011). TH activity is the rate-limiting factor in catecholamine synthesis (Levitt et al., 1965), and consequently THD patients suffer from a critical reduction of catecholamine levels (Bräutigam et al., 1999; Willemsen et al., 2010). THD arises from an autosomal recessive pathogenic variant in both TH genes on chromosome 11p15.5 (OMIM #605407; Willemsen et al., 2010; Zafeiriou et al., 2009). Several pathogenic variants have been described (missense variants leading to partial loss of enzyme activity, deleterious variants leading to protein truncation, or pathogenic variants in the promoter region leading to reduced TH gene transcription) and patients can be homozygous or compound heterozygous (Willemsen et al., 2010). PAH = Phenylalanine; AADC = Aromatic amino acid decarboxylase.

As THD is an autosomal recessive disorder, heterozygous carriers of pathogenic variants in the TH gene ('THD carriers') are thought to be free of neurological symptoms (i.e., no overt cognitive, neurological and psychiatric impairments have been observed). However, given the clear parallels between THD and other monoamine neurotransmitter disorders, it is reasonable to assume that THD carriers express lower TH enzyme activity than non-carriers. For example, lower enzyme activity has been determined for first- and second-degree relatives of patients with aromatic L-amino acid decarboxylase (AADC) deficiency by Verbeek and colleagues (2007). Enzymatic activity analyses showed that the unaffected carriers had 35-40% lower AADC enzyme activity than healthy controls, usually in the absence of any clinical signs. AADC deficiency parallels TH deficiency because both are enzyme deficiency disorders affecting the catecholamine system, but each affects the biosynthesis of catecholamines at different stages (Figure 1) and AADC additionally affects the biosynthesis of serotonin (Willemsen et al., 2010). Accordingly, we hypothesised that carrying a pathogenic variant in the TH gene, which likely leads to decreased TH enzymatic activity and consequently decreased

dopamine biosynthesis, would be associated with subtle adaptations in motivated action and learning that surface only when probing behaviour using sophisticated catecholamine-sensitive experimental paradigms. Here we focus on a paradigm that has been previously established to be sensitive to manipulation of catecholamines, namely a motivational Go/NoGo learning paradigm (Guitart-Masip et al., 2014b; Swart et al., 2017).

Dopamine has been linked to behavioural activation in the context of reward (Taylor and Robbins, 1986, 1984), where enhanced dopamine facilitates instrumental activation in the context of reward conditioned cues (Wyvell and Berridge, 2000), and lowered dopamine levels reduce instrumental activation in the context of these cues (Dickinson et al., 2000; Hebart and Gläscher, 2015; Lex and Hauber, 2008). Conversely, punishment conditioned cues suppress instrumental responding (Davis and Wright, 1979; Huys et al., 2011), and striatal dopamine has been proposed to also contribute to such aversively motivated behaviour (Faure et al., 2008; Lloyd and Dayan, 2016). These motivational biases in action (i.e., behavioural activation and inhibition by reward and punishment cues respectively) is consistent with current accounts of striatal dopamine function (Collins and Frank, 2015b, 2014, Frank, 2006, 2005; Lloyd and Dayan, 2016), suggesting that dopamine bursts elicited by predicted rewards potentiate the basal ganglia direct 'Go' pathways, thereby promoting behavioural activation. Consequently, relatively enhanced dopamine responses would further facilitate behavioural activation. In contrast, dips in dopamine firing elicited by predicted punishments potentiate the basal ganglia 'NoGo' pathway, promoting behavioural inhibition. In this study, we hypothesized that THD carriers might show weaker motivational biases in action compared with controls, due to reduced dopamine function.

We set out to investigate the consequences of carrying a pathogenic TH genetic variant on motivational biases in action. To this end, we employed a motivational Go/NoGo learning task that requires subjects to learn to make Go or NoGo responses to cues in order to obtain reward or avoid punishment (cf. Guitart-Masip et al., 2011; Swart et al., 2017). The task quantifies the degree to which subjects are biased towards Go responding when pursuing reward, and NoGo responding when avoiding punishment. The task also allowed us to assess the valuation of these motivational Go and NoGo cues, by assessing explicit, subjective cue preferences after learning (Cavanagh et al., 2013). We contrasted THD carriers with an education- and age-matched control group in a between-subject design.

Methods

Subjects

For this study, all known Dutch families of a child with THD (n=8) were approached. We tested one group of THD carriers (n=16; sample size limited by the THD prevalence) and one education- and age-matched control group (n=20; see Table 1 for demographics). All subjects were native Dutch speakers. The THD carrier group consisted predominantly of the biological parents of TH

deficient children (8 mothers, 7 fathers), and of one other family member (aunt) who was a known THD carrier. Because THD is an autosomal recessive heritable disorder, both biological parents are obligated carriers of a pathogenic variant in one TH allele (Lüdecke et al., 1995), and genetic assessments confirmed the presence of a pathogenic variant in the TH gene. The THD carrier group was recruited via the treating child neurologists (MW and TW), and the control group via the Radboud University campus. Potential subjects received information prior to the testing day and signed informed consent prior to participation. THD carriers also signed a consent form that allowed us to request results of their genetic assessments at according hospitals to confirm their pathogenic TH variant. Subjects with abnormal vision (e.g., colour-blindness) were excluded from this study, resulting in the exclusion of one THD carrier (this subject did not complete the task). One other THD carrier could not complete the task due to prior medical reasons. Additional exclusion criteria for the control group were use of dopaminergic medication, (history of) neurological and psychiatric treatment, and alcohol or drug dependence. Subjects received a reimbursement for travel expenses and EUR8,- per hour for their time-investment.

Experimental procedure

The study contained one test session, including self-paced breaks. The test session took place at the Donders Institute or at the subjects' home. The test session consisted of a cognitive task battery (~95 min) and a neuropsychological assessment (~60 min). The THD carrier group additionally completed a neurological and psychiatric screening (~60 min). The cognitive task battery included a probabilistic reversal learning task (den Ouden et al., 2013), a delayed match-to-sample task (Fallon and Cools, 2014), a motivational Go/NoGo learning task (see below), and the Listening Span Test (Daneman and Carpenter, 1980). In this chapter, I focus on the motivational Go/NoGo learning task, but note that we intend to publish an overarching paper combining all independent assessments. The neuropsychological assessment consisted of i) neuropsychological tasks, namely the Dutch reading test (NLV; Schmand et al., 1991), Story Recall Test (Wechsler, 1997), Box Completion (Salthouse, 1994), Number Cancellation (Mesulam, 1985), Stroop task (Stroop, 1935), Verbal Fluency (Benton and Hamsher, 1983), and ii) of self-report questionnaires, namely the Barratt Impulsiveness Scale (BIS-II; Patton et al., 1995), Obsessive Compulsive Inventory revised (OCI-R; Foa et al., 2002), NEO personality inventory (NEO-FFI; Costa and McCrae, 1992), Need for Cognition Scale (NfC; Cacioppo et al., 1984), Perceived Stress Scale (PSS; Cohen et al., 1983), Beck Depression Inventory (BDI; Beck et al., 1996), and the Hospital Anxiety and Depression Scale (HADS; Zigmond and Snaith, 1983). The neurological and psychiatric screening consisted of a general health assessment, standardized neurological exam (Clarke et al., 2016), Unified Parkinson's Disease Rating Scale (UPDRS-III; Fahn and Elton, 1987), Fahn-Marsden Dystonia Movement Scale (FMDM; Burke et al., 1985), and the Mini International Neuropsychiatric Interview questionnaire (MINI; Sheehan et al., 1998). The study was approved by the local ethics committee (CMO / METC Arnhem Nijmegen: CMO2014/288), and in line with the Declaration of Helsinki.

	THD carriers (n=14)	Control group (n=20)	Group difference
Matching criteria			
Gender (women / men)	9 / 5	11 / 9	$p = .588$
Age (Mean (SD), range)	49.9 (8.8), 36 - 65	50.8 (9.5), 30 - 63	$p = .788$
Education (n)			$p = .946$
Lower education	8	11	
Higher education	5	8	
University	1	1	
NLV (Mean (SD))	79.6 (15.3)	80.1 (9.2)	$p = .914$
Control measures			
BDI (Mean (SD))	4.6 (4.1)	4.4 (6.0)	$p = .921$
HADS (Mean (SD))	7.1 (4.3)	7.3 (4.8)	$p = .910$
PSS (Mean (SD))	11.0 (5.2)	12.0 (5.2)	$p = .588$
Measures of interest			
BIS-II (Mean (SD))	55.4 (6.0)	62.9 (8.4)	$p = .007$*

Table 1. **Demographics for the THD carrier group (n=14) and the matched control group (n=20).** The control group was successfully matched to the THD carrier group in terms of gender, age, education, and verbal intelligence (NLV). We checked whether the THD carriers showed increased perceived stress (PSS; Perceived Stress Scale) and depressive or anxiety symptoms (BDI; Beck Depression Inventory. HADS; Hospital Anxiety and Depression Scale) as a potentially direct consequence of caring for a child with severe medical problems. The groups did not significantly differ on any of these control measures. Finally, we assessed whether the groups significantly differed in terms of trait impulsivity (BIS; Barratt Impulsiveness Scale), which has been linked to dopamine function with PET (Buckholtz et al., 2010; Kim et al., 2014; Lee et al., 2009; Reeves et al., 2012) and has commonly been used as a proxy variable for baseline dopamine function within our group (Frobose et al., 2017; Swart et al., 2017). The THD carriers had significantly lower trait impulsivity scores than the matched controls.

Motivational Go/NoGo learning task

We employed a motivational Go/NoGo learning task (similar to Guitart-Masip et al., 2011; Swart et al., 2017), in which cue valence (Win vs. Avoid cue) was orthogonal to the instrumental response (Go vs. NoGo). In this task, subjects needed to learn to make Go or NoGo responses in order to obtain rewards (Win cues) or avoid punishments (Avoid cues). Each cue had one correct response, which subjects needed to learn by trial-and-error based on feedback. See Figure 2 for an overview of the task.

Each trial started with a cue presentation (1.2s) during which subjects could either press the spacebar (Go response) or wait until the cue disappears (NoGo response). Each cue had a coloured edge indicating the cue valence. A green edge was indicative of a Win cue, which could only be followed by reward or a neutral outcome. Conversely, a red edge was indicative of an Avoid cue, which could only be followed by a punishment or a neutral outcome. The cue was followed by a fixation cross (0.5s), and response-dependent feedback (1s). More specifically, correct responses to Win cues were followed by reward 75% of the time, and by neutral outcomes otherwise. Similarly,

correct responses to Avoid cues were followed by neutral outcomes 75% of the time, and by punishment otherwise. For incorrect responses, these probabilities were reversed. In total there were four cue types for which cue valence was orthogonalised to the required action (Figure 2). Reward consisted of a green '+100' text and a flourish sound. Neutral outcomes consisted of a grey '000' text and neutral beep. Punishment outcomes consisted of a red '-100' text and a low buzz. Trials ended with a randomized inter-trial interval (1.25-2s) during which a fixation cross was presented.

The task was preceded by instructions, including two practice rounds. Subjects were instructed i) that each cue had one optimal response, ii) that each cue could be followed by either reward or punishment, and iii) about the probabilistic nature of the feedback. Subjects received a self-paced break halfway during the task. Each cue was presented 30 times in pseudorandomized order. The task was performed twice with independent, counterbalanced stimulus sets.

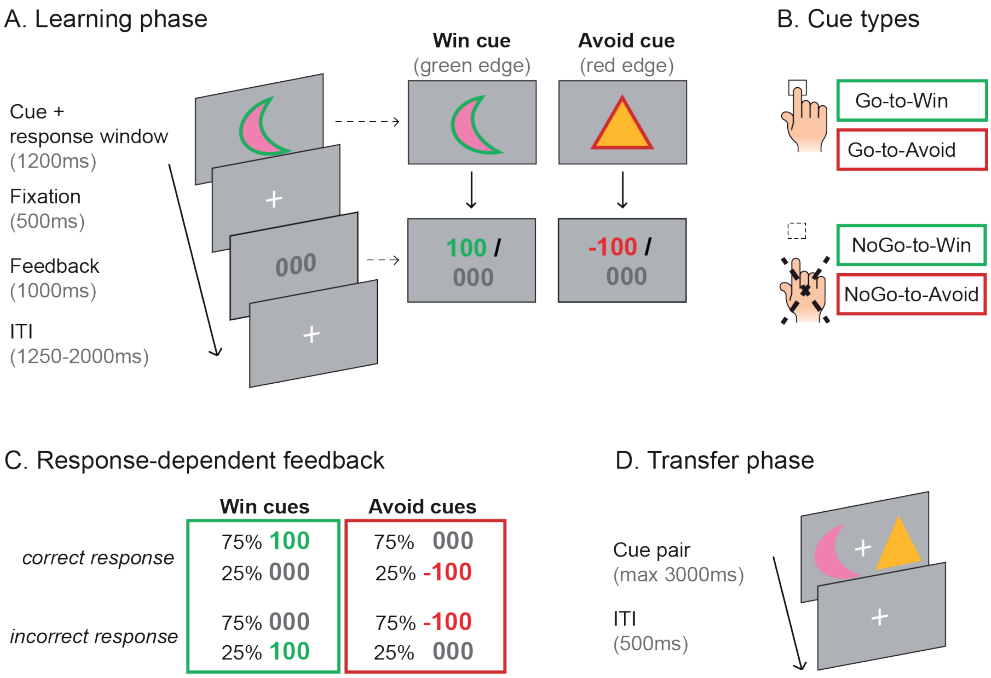


Figure 2. **Motivational Go/NoGo learning task.**
(**a**) Trials start with a cue, indicating the response window, followed by feedback. Win cues can be followed by reward, whereas Avoid cues can be followed by punishment. Image adapted from (Swart et al., 2017). (**b**) There are four cue-types for which cue valence (Win vs. Avoid) are orthogonalised to the required action (Go vs. NoGo). (**c**) Each cue has one correct response, which subjects need to learn by trial-and-error based on the feedback. Correct responses are followed by reward (Win cues) and neutral outcomes (Avoid cues) 75% of the time, or by neutral outcomes (Win cues) and punishment (Avoid cues) otherwise. These probabilities are reversed for incorrect responses. (**d**) The learning task is followed by a transfer phase. Cues are presented in pairs and subjects are instructed to select the most rewarding cue. The coloured cue edges are omitted during this phase.

After finishing the learning task, subjects completed a transfer phase (Cavanagh et al., 2013) in which we assessed the relative, subjective cue values. During this phase, cues from the last stimulus set were presented in pairs, and subjects were asked to select the most rewarding cue. This transfer phase allowed us to verify that subjects experienced the Win cues as more rewarding than the Avoid cues, but more importantly, whether subjects preferred the cues requiring active Go response over the cues requiring passive NoGo responses, as has been shown previously (Cavanagh et al., 2013; Swart et al., 2018). The transfer phase contained 48 trials. During this phase, the coloured cue edges signalling valence were omitted, in order to probe the learned relative preferences and minimize interference by the explicit cue valences.

Statistical analysis

In this study we investigated the consequences of carrying a pathogenic TH genetic variant on motivational biases in action. To this end, we first tested whether subjects adjusted Go/NoGo responding to the cue valence, which we refer to as the motivational bias, and then assessed whether the THD carriers showed a reduced motivational bias compared with the control group. We additionally assessed whether subjects adjusted Go/NoGo responding to the required action, in line with task learning, and whether the groups differed in terms of task learning. Accordingly, the statistical model for Go responses included the between-subject factor Group (THD carrier vs. control), and the within-subject factors Valence (Win vs. Avoid cue) and Required Action (Go vs. NoGo). We analysed reaction times (RTs) as a complementary measure of behavioural vigour. Here, we restricted the RT analysis to correct responses, i.e. to the Go cues, to reduce the model's effects structure and thereby increase statistical power. Thus, the RT model included the within subject factors Valence and the between subject factor Group. Given that we set out to test the hypothesis that the motivational biasing of action might be reduced in the THD carriers due to assumed dopamine depletion, we employed one-sided tests for the Valence x Group interactions. These one-sided tests are clearly indicated in the Results section.

We analysed trial-by-trial choices (RTs) with logistic (linear) mixed-effect models using lme4 in R (Bates et al., 2014; R Developement Core Team, 2015). The mixed-effect analysis has a clear advantage over ANOVA particularly for the RTs, as mixed-effect models take the number and consistency of RTs per subject into account, thereby accounting for within and between subject variability. RTs were log-transformed to improve normality and RTs<100ms were discarded from the analysis. The mixed models included all main effects and interactions, and a full random effects structure (Barr, 2013; Barr et al., 2013). We estimated effect sizes based on the corresponding repeated measures ANOVA performed within SPSS, given that there is no clear consensus on the estimation of effect size for mixed-models. We report partial eta squared (η_p^2) as a measure of effect size for all group

effects, where we interpret η_p^2>.14 as large effects, η_p^2>.06 as medium effects, and η_p^2>.01 as small effects, in line with (Cohen, 1992, 1988). Finally, we repeated all analyses including the control covariates age, gender and education to confirm that our conclusions remain the same.

Finally, we assessed whether the THD carrier group differed from the control group in their relative cue preferences during the transfer phase. To this end, we analysed how often each cue was chosen during the transfer phase relative to chance. We analysed the frequency data with repeated measures ANOVA in SPSS using the between-subject factor Group, and the within-subject factors Valence and Required Action.

Results

General task performance and subjective valuation

Before addressing differences between the THD carrier group and the control group, we established that expected task effects were present across groups. First, we assessed behaviour as a function of the required actions, related to task learning, and second, we assessed behaviour as a function of cue valence, related to motivational biases. Subjects made more Go responses to Go than NoGo cues (X^2_1=33.9, p<.001; Figure 3), indicating that subjects adjusted their responses to the instrumental requirements. Independent of the action requirements, subjects made more Go responses (X^2_1=44.7, p<.001; Valence x Required Action: X^2_1<1, p=.926) and faster Go responses (X^2_1=48.8, p<.001) to Win than Avoid cues, which we refer to as a motivational bias. Altogether, the current sample shows the commonly observed task effects related to task learning and motivational biases (Guitart-Masip et al., 2014a; Swart et al., 2018, 2017).

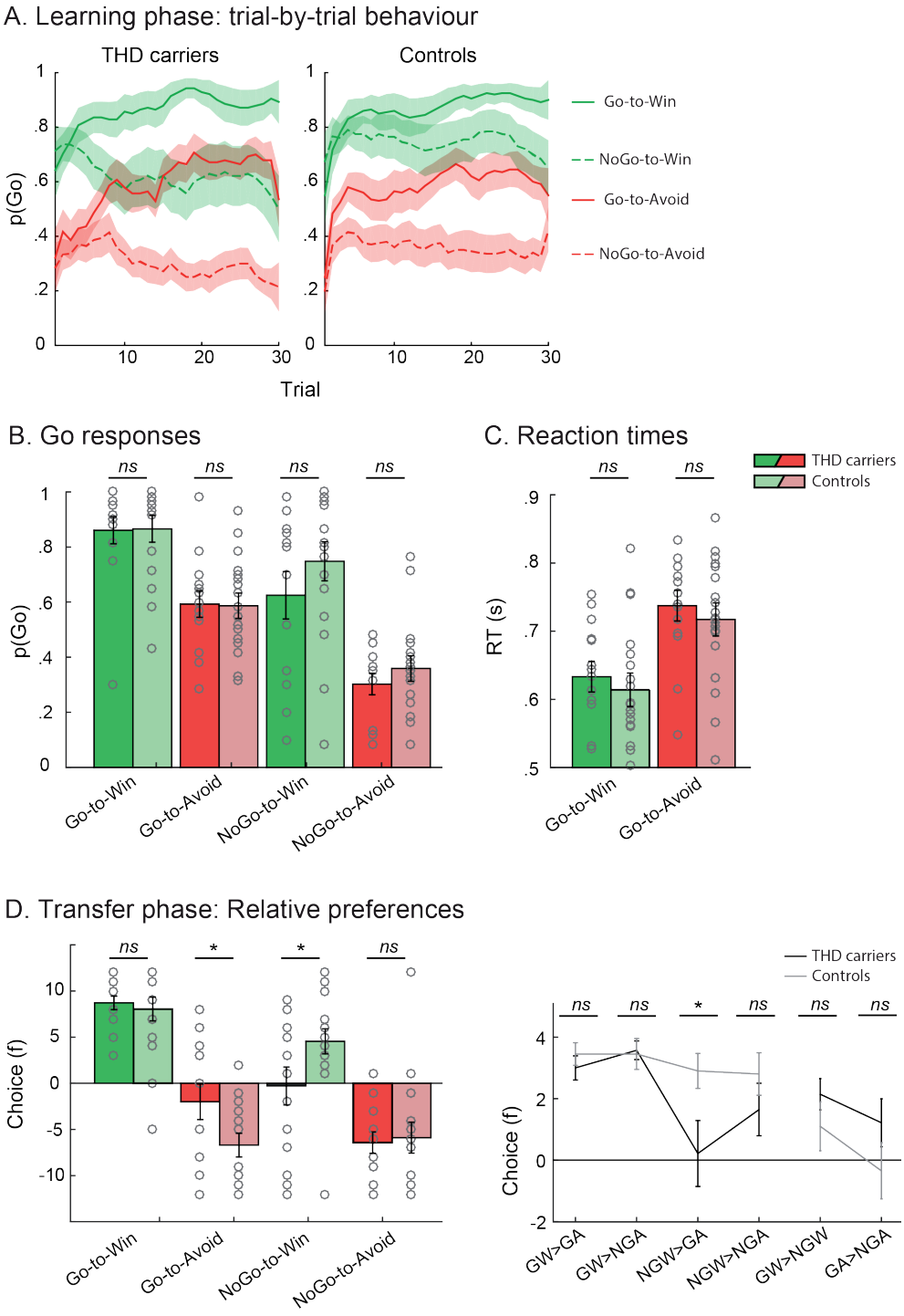
Before turning to the group differences, we assessed the choices during the transfer phase across groups. During the transfer phase, cue pairs were presented and subjects needed to select the most rewarding cue. Accordingly, subjects selected the Win cues more frequently than the Avoid cues ($F_{1,32}$=93.4, p<.001), indicating that subjects indeed considered Win cues more rewarding than the Avoid cues. Furthermore, subjects selected Go cues more often than NoGo cues ($F_{1,32}$=10.1, p=.003), which was particularly driven by the Win cues (simple effect of Required Action: $F_{1,32}$=19.6, p<.001), rather than the Avoid cues (simple effect of Required Action: $F_{1,32}$=1.3, p=.263; Required Action x Valence: $F_{1,32}$=7.5, p=.010). This pattern of results is also consistent with previous reports (Cavanagh et al., 2013; Swart et al., 2018), showing enhanced relative values for cues associated with Go responses relative to cues associated with NoGo responses.

Figure 3 (right page). **Behavioural performance.**
(**a**) Trial-by-trial responses for the THD carriers (left) and controls (right) using a sliding average of 5 trials. The shaded areas indicate the standard error of the mean. Subjects increased Go responses for Go cues and NoGo responses for NoGo cues over trials (p<.001), indicative of task learning. From the first trial onwards, subjects made more Go responses to Win than Avoid cues (p<.001), which we refer to as the motivational bias. (**b-c**) Average proportion Go responses and reaction times. Circles indicate individual subjects and error bars indicate the standard error of the mean. The THD carrier group and the control group did not show any significant differences in proportion Go responses and reaction times (p>.05). ns indicates p>.05. (**d**) Cue preferences as measured in the transfer phase relative to chance level. Left: Total choice frequency per cue. Right: Choice frequency per cue pair. The groups particularly differed in their relative preferences when comparing the NoGo-to-Win and the Go-to-Avoid cues (p=.013; all other pairs: p>.05), where the THD carrier group selected the NoGo-to-Win cue significantly less often than the control group. Remarkably, this reduced preference for the NoGo-to-Win cues was not explained by a lower reward history, as the THD carriers performed numerically better for the NoGo-to-Win cues (panel b). * indicates p<.05. (N)GW=(No)Go-to-Win; (N)GA=(No)Go-to-Avoid.

Altered subjective cue valuation, but not task performance, in THD carriers vs. matched controls

Having established the presence of common task effects across groups in both the learning and transfer phase, we continued with contrasting the THD carrier group with the control group. First, we addressed whether the groups differed in terms of their motivational bias, that is, the differential responding to Win and Avoid cues. The groups did not differ significantly in the proportion of Go responses for Win vs. Avoid cues (X^2_1<1, p=.224, *one-sided test*, η_p^2=.006) or in terms of RTs (X^2_1<1, p=.394, *one-sided test*, η_p^2<.001), see Figure 3. We only observed a small effect size for a reduced valence effect in the THD carrier group on the NoGo trials, which did not reach statistical significance in the current sample (Required Action x Valence x Group: X^2_1<1, p=.872, η_p^2=.013). Taking together, we did not observe a significant reduction in the valence-based biases in the THD carriers compared with the matched controls. The groups also did not significantly differ in the extent to which they adjusted their Go/NoGo responses to the required action (X^2_1<1, p=.358, η_p^2=.047), nor in the overall proportion of Go responses or RTs (Go: X^2_1=2.3, p=.127, η_p^2=.051; RT: X^2_1<1, p=.515, η_p^2=.019). We confirmed that these results remained unchanged when including age, education, and gender as control covariates in the models. Altogether, we did not observe significant differences in the task performance between the THD carrier group and the control group in the learning phase.

Second, we addressed group differences in the subjective cue preferences as measured in the transfer phase. The THD carrier group showed a marginally weaker preference for the Win vs. Avoid cues (Group x Valence: $F_{1,32}$=3.7, p=.064, η_p^2=.103), and showed a significantly stronger preference for the Go vs. NoGo cues compared with the matched controls (Group x Required Action: $F_{1,32}$=4.5, p=.042, η_p^2=.123; Group x Valence x Required Action: $F_{1,32}$<.1, p=.934, η_p^2<.001). These group differences specifically reflected a shift in the valuation of


A. Learning phase: trial-by-trial behaviour
THD carriers
Controls
p(Go)
Trial
Go-to-Win
NoGo-to-Win
Go-to-Avoid
NoGo-to-Avoid
B. Go responses
C. Reaction times
THD carriers
Controls
RT (s)
D. Transfer phase: Relative preferences
Choice (f)
GW>GA
GW>NGA
NGW>GA
NGW>NGA
GW>NGW
GA>NGA
ns
*

the incongruent cues (i.e. Go-to-Avoid and NoGo-to-Win cues), see Figure 3. The groups indeed differed significantly in their choices on the NoGo-to-Win vs. Go-to-Avoid cue pairing (t_{32}=2.6, p=.013), and not on the other cue pairings (all: p>.05). To elaborate, the THD carriers selected the NoGo-to-Win significantly less often than the control group when choosing between the NoGo-to-Win vs. Go-to-Avoid cue, suggesting that the THD carriers' preferences were less affected by valence and more by the associated action. Importantly, this shift in cue preferences was not explained by a differential outcome history for the groups (NoGo-to-Win – Go-to-Avoid: t_{32}=1.5, p=.155; NoGo-to-Win: t_{32}=1.4, p=.158; Go-to-Avoid: t_{32}<.1, p=.925). If anything, the THD carriers received (numerically) more rewards for the NoGo-to-Win cues relative to controls, yet significantly preferred this cue less. This group difference was also not purely explained by the group difference in trait impulsivity; trait impulsivity did not significantly relate to the choices on the NoGo-to-Win vs. Go-to-Avoid cue pairing (R=.26, p=.140), and the group difference in cue preference remained significant when correcting for impulsivity (t_{31}=2.1, p=.043). The transfer results also remained unchanged when correcting for age, education, and gender in the model. Together these results raise the hypothesis that THD carriership affects relative cue preferences in the context of motivationally incongruent cues, while leaving motivational biasing of action unaltered.

Discussion

Here we assessed the effects of carrying a pathogenic variant in the tyrosine hydroxylase (TH) gene on the motivational biasing of action by comparing family members of tyrosine hydroxylase deficient (THD) patients with education- and age-matched controls. In both the carriers and matched controls, cue valence strongly biased Go/NoGo responding, such that subjects made more Go responses when playing for reward (Win cues) than when trying to avoid punishment (Avoid cues). This motivational bias in Go/NoGo responding was not significantly reduced in the THD carriers relative to the controls. In contrast, the THD carriers differed from the matched controls in relative cue preferences. The carriers showed a reduced impact of valence on their subjective cue valuation, specifically in the context of incongruency between the action requirements and the valence. In other words, they liked NoGo-to-Win less, and Go-to-Avoid cues more, relative to controls.

This study is part of the first project addressing the neurocognitive consequences for heterozygous carriers of a pathogenic variant in the TH gene. In this chapter, we set out to specifically assess the consequences on the well-established motivational biases in action. We hypothesized that family members of THD patients have a mild reduction in dopamine synthesis that would result in reduced motivational biases in action, in line with current accounts of striatal dopamine function (Collins and Frank, 2015b, 2014; Frank, 2005; Frank

et al., 2004; Lloyd and Dayan, 2016) and prior results from our group showing that increases in catecholamine transmission with methylphenidate enhanced such motivational action biasing (Swart et al., 2017). We replicated commonly observed motivational biases in both active actions and response times (Guitart-Masip et al., 2014a; Swart et al., 2018, 2017), reflected in the increased proportion and speed of Go responses to Win relative to Avoid cues. This motivational biasing of Go responses and response times was not significantly reduced in the THD carrier group. These results might indicate that the motivational biasing of action is not sensitive to reductions in TH enzymatic activity. However, we did observe a small effect size for a reduced motivational bias on the NoGo cues, meaning that we cannot exclude the possibility that the lack of significance was due to a lack of statistical power. Indeed, patients with Parkinson's disease off dopaminergic medication (i.e., when striatal dopamine is severely depleted) have been shown to express enhanced NoGo-to-Win performance (Moustafa et al., 2008) and reduced willingness to exert effort for reward (Chong et al., 2015), which both normalize ON dopaminergic medication. Thus, our current sample sizes might have been too small to detect significant group differences.

On the other hand, the absence of a significant reduction in motivational biasing of action in the carriers might be due to a degree of evolutionarily preprogrammed redundancy in (and thus compensatory capacity of) monoamine synthesis enzymatic activity (Wassenberg et al., 2012). Put simply, although heterozygous state must result in lower tyrosine hydroxylase enzymatic activity in carriers, this might have no clinical significance on catecholamine levels. We argue that this is less likely given that benign variants in the monoamine pathways without known functional effects on protein level have long been thought to be associated with subtle motivational and/or cognitive deficits, and to contribute to several neuropsychiatric disorders (Haavik et al., 2008).

The absence of significant group differences in the motivational biases might raise the question whether dopamine function is indeed altered in THD carriers. More direct measures of dopamine function, for example dopamine synthesis capacity or turnover, are required to conclusively answer this question, yet various aspects of the data support the assumptions that dopamine function is altered. First, the THD carriers displayed significantly lower trait impulsivity scores than the matched controls, and trait impulsivity has been linked to dopamine function with PET (Buckholtz et al., 2010; Kim et al., 2014; Lee et al., 2009; Reeves et al., 2012). Thus, the significant group difference in trait impulsivity is consistent with the assumptions that THD carriers express altered dopamine function. Second, the THD carriers show modulated relative cue preferences, which we will cautiously link to dopamine function below. Finally, carriers of a pathogenic variant in the related AADC gene also express significantly lower AADC enzyme activity (Verbeek et al., 2007). Considering the parallels between AADC and TH deficiency (also described in the introduction of this chapter), we expect a similar decrease in TH enzyme activity

(and consequently dopamine synthesis) in THD carriers. Taken together, we stand by our initial assumption that dopamine function is altered in THD carriers, yet we acknowledge that future studies measuring TH enzyme activity or dopamine function more directly are needed to verify this assumption.

At the end of the learning task, subjects were asked to select the most rewarding cue out of presented cue pairs. During this transfer phase, the carrier group showed altered relative cue preferences compared with the matched control group, as evidenced by a shift in preference from the NoGo-to-Win cues towards the Go-to-Avoid cues. The altered cue preferences in the absence of altered motivated action is particularly remarkable when considering that motivation and valuation typically go hand in hand (Niv et al., 2007). Yet, it has been shown that these processes can be dissociated (Miller et al., 2014), suggesting that these processes might rely on differential mechanisms. Although the transfer phase was not the primary measure of the current study, we will discuss potential explanations for the observed group differences in the following section.

First of all, the altered relative preferences in the carriers might reflect changes in the outcome predictions, presumably due to reduced dopamine function. Dopamine has classically been linked to reward prediction (Schultz et al., 1998, 1997), and has been linked to the prediction of hedonic pleasure as well (Sharot et al., 2009), even though the role of dopamine in instant hedonic pleasure or 'liking' has been disputed (Berridge, 2009; Berridge et al., 2009). Sharot and colleagues showed that administration of a dopamine-enhancing drug (L-Dopa) increased subjective estimations of future hedonic pleasure to positive future life events. Similarly, administration of L-Dopa enhanced the optimism bias (for a review see Sharot, 2011), as L-Dopa reduced negative expectations about the future (Sharot et al., 2012). Consistent with these findings, the assumed dopamine reduction in the THD carriers might have led to attenuated expectations of reward outcomes and associated hedonic pleasure. Such a dopamine-dependent attenuation in reward expectations would explain why the carriers indicated the Win cues less often as rewarding compared with the matched controls.

Notably, the attenuated preference of Win cues in the carrier group was specific to the context where cue valences were motivationally incongruent with the action requirements (i.e. NoGo-to-Win vs. Go-to-Avoid cues). In other words, the group difference in value-based preferences only surfaced when these preferences were inconsistent with action-based preferences. To elaborate, subjects expressed an overall relative preference of the Go cues over the NoGo cues, in line with previous studies (Cavanagh et al., 2013; Swart et al., 2018). In general, approach and avoid behaviour are known to respectively increase positive and negative valuation of novel stimuli (e.g. Huijding et al., 2011; Laham et al., 2014; Woud et al., 2013, 2008), and approach-avoid training is even used to retrain approach tendencies of harmful consumption behaviour, such as alcohol use and unhealthy eating (Kakoschke

et al., 2017). Similarly, freely chosen options tend to enhance relative preferences, whereas discarded options tend to decrease relative preferences ('choice bias'; e.g., Cockburn et al., 2014; Sharot et al., 2010, 2009). Here, the active Go and inactive NoGo responses might have influenced the affective valuation of the cues in a similar manner. This action-based affective valuation was clearly present in the carrier group, and was enhanced relative to the control group. Thus, the relative contribution of action-based and valence-based cue valuation were shifted in the carrier group, with an increased relative contribution of the associated action (or, 'the actor') and a decreased relative contribution of the cue valence (or, 'the critic'). A shift in the relative contribution of the associated action and valence values would explain why the THD carriers expressed a shift in relative preferences only when the action- and valence-based preferences were incongruent. Although i) reduced reward expectations are consistent with current views of reduced dopamine function, and ii) unaffected action-based contribution would be consistent with the unaffected task performance in the learning phase, future research is needed to disentangle the absolute changes in the contribution of these complementary mechanisms.

Alternative to reflecting affected reward expectations, the attenuated relative preferences for Win and Avoid cues in the carrier group could reflect a disruption in value-based *learning*, which also has been widely linked to dopamine function (Collins and Frank, 2014; Frank et al., 2004; Montague et al., 2004; Schultz et al., 1997; Wise, 2004). If either the valuation or learning of reward and punishment outcomes is disrupted, that would explain why the THD carrier group indicated the Win cues as relatively less rewarding. Although one might have expected a disruption in value-based learning or decision-making based on current views of dopamine function, such a disruption seems unlikely given that the carrier group did not perform significantly differently from the control group during the learning phase. We cannot rule out compensation strategies (e.g., enhanced contribution of working memory or additional prefrontal functions), or enhanced engagement in the THD carriers (particularly given the personal relevance of the study), and that the carrier group thereby could compensate for a disruption in value-based learning, yet such a combined account is less parsimonious.

We have linked the group differences in relative preferences to changes in dopamine function above, yet it should be noted that the TH enzyme does not only affect dopamine synthesis, but affects catecholamine synthesis in general (Levitt et al., 1965), as dopamine is the precursor for noradrenaline (Kurian et al., 2011; Figure 1). Thus, even though these results are consistent with altered dopamine function, we cannot exclude the possibility that the other catecholamines contributed to the observed group differences.

Finally, given that the groups particularly differed in their relative preferences for the motivationally incongruent cues (i.e., NoGo-to-Win and Go-to-Avoid), one might wonder whether the carriers showed enhanced discounting for exerting control (cf. Cavanagh et

al., 2014). On the incongruent trials, instrumental requirements conflict with prepotent, Pavlovian response tendencies elicited by the cue valence, and this Pavlovian conflict is thought the require increased levels of control over behaviour (Cavanagh and Frank, 2014; Swart et al., 2018) and to be inherently aversive (Cavanagh et al., 2014). Although the carriers indeed preferred the NoGo-to-Win cues less than the controls, they preferred the Go-to-Avoid cues *more* than the matched controls, which argues against enhanced discounting of conflict in the THD carrier group.

Conclusion

We set out to assess for the first time the cognitive consequences of carrying a pathogenic variant in the TH gene. We specifically assessed the impact of being of carrier of THD on the well-established motivational biases in action on the one hand and subjective valuation on the other. In both the THD carriers and matched controls, anticipated rewards and punishment elicited Go and NoGo responses respectively. While the groups did not significantly differ in this motivational biasing of Go responding, the groups strikingly differed in their relative valuation of the cues. The THD carrier group valued the NoGo-to-Win cues less than the matched control group, while preferring the Go-to-Avoid cues more. Our results suggest that motivational biases in action are unaffected in THD carriers, whereas subjective cue valuation is altered relative to matched controls, potentially due to catecholamine-dependent changes in reward expectations. This pilot study provides a first insight into the subtle cognitive changes in a highly unique and hitherto unstudied genetic population involving the catecholamine system.

References

Barr, D.J., 2013. Random effects structure for testing interactions in linear mixed-effects models. Front. Psychol. 4, 328. doi:10.3389/fpsyg.2013.00328

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4. R Packag. version. doi:citeulike-article-id:7112638

Beck, A.T., Steer, R.A., Ball, R., Ranieri, W., 1996. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. J. Pers. Assess. 67, 588–597. doi:10.1207/s15327752jpa6703_13

Benton, A., Hamsher, K., 1983. Multilingual Aphasia Examinationle, AJA Associates. doi:10.1076/clin.15.1.13.1911

Berridge, K.C., 2009. "Liking" and "wanting" food rewards: Brain substrates and roles in eating disorders. Physiol. Behav. 97, 537–550. doi:10.1016/j.physbeh.2009.02.044

Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Res. Rev. doi:10.1016/S0165-0173(98)00019-8

Berridge, K.C., Robinson, T.E., Aldridge, J.W., 2009. Dissecting components of reward: "liking", "wanting", and learning. Curr. Opin. Pharmacol. doi:10.1016/j.coph.2008.12.014

Bräutigam, C., Steenbergen-Spanjers, G.C., Hoffmann, G.F., Dionisi-Vici, C., van den Heuvel, L.P., Smeitink, J. a, Wevers, R. a, 1999. Biochemical and molecular genetic characteristics of the severe form of tyrosine hydroxylase deficiency. Clin. Chem. 45, 2073–8.

Brozoski, T., Brown, R., Rosvold, H., Goldman, P., 1979. Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. Science (80-.). 205, 929–932. doi:10.1126/science.112679

Buckholtz, J.W., Treadway, M.T., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., Baldwin, R.M., Schwartzman, A.N., Shelby, E.S., Smith, C.E., Kessler, R.M., Zald, D.H., 2010. Dopaminergic Network Differences in Human Impulsivity. Science (80-.). 329, 532. doi:10.1126/science.1185778

Burke, R.E., Fahn, S., Marsden, C.D., Bressman, S.B., Moskowitz, C., Friedman, J., 1985. Validity and reliability of a rating scale for the primary torsion dystonias. Neurology 35, 73–7. doi:10.1212/WNL.35.1.73

Cacioppo, J.T., Petty, R.E., Kao, C.F., 1984. The efficient assessment of need for cognition.: EBSCOhost. J. Pers. Assess. 48, 306–307. doi:10.1207/s15327752jpa4803_13

Cansev, M., Wurtman, R.J., 2007. Aromatic Amino Acids in the Brain. Handb. Neurochem. Mol. Neurobiol. 59–97. doi:10.1007/978-0-387-30373-4_4

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., Frank, M.J., 2013. Frontal Theta Overrides Pavlovian Learning Biases. J. Neurosci. 33, 8541–8548. doi:10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. doi:10.1016/j.tics.2014.04.012

Cavanagh, J.F., Masters, S.E., Bath, K., Frank, M.J., 2014. Conflict acts as an implicit cost in reinforcement learning. Nat. Commun. 5, 5394. doi:10.1038/ncomms6394

Chong, T.T.J., Bonnelle, V., Manohar, S., Veromann, K.R., Muhammed, K., Tofaris, G.K., Hu, M., Husain, M., 2015. Dopamine enhances willingness to exert effort for reward in Parkinson's disease. Cortex 69, 40–46. doi:10.1016/j.cortex.2015.04.003

Clarke, C., Howard, R., Rossor, M., Shorvon, S., 2016. Neurology: A Queen Square textbook: Second edition, Neurology: A Queen Square Textbook: Second Edition. doi:10.1002/9781118486160

Cockburn, J., Collins, A.G.E., Frank, M.J., 2014. A Reinforcement Learning Mechanism Responsible for the Valuation of Free Choice. Neuron 83, 551–557. doi:10.1016/j.neuron.2014.06.035

Cohen, S., Kamarck, T. and Mermelstein, R., 1983. A Global Measure of Perceived Stress. J. Health Soc. Behav. 24, 386–396. doi:10.2307/2136404

Cohen, J., 1992. A power primer. Psychol. Bull. 112, 155–159. doi:10.1037/0033-2909.112.1.155

Cohen, J., 1988. Statistical power analysis for the behavioral sciences. Stat. Power Anal. Behav. Sci. doi:10.1234/12345678

Collins, A.G.E., Frank, M.J., 2015. Surprise! Dopamine signals mix action, value and error. Nat. Neurosci. 19, 3–5. doi:10.1038/nn.4207

Collins, A.G.E., Frank, M.J., 2014. Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. 121, 337–66. doi:10.1037/a0037015

Cools, R., Frank, M.J., Gibbs, S.E., Miyakawa, A., Jagust, W., D'Esposito, M., 2009. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J. Neurosci. 29, 1538–43. doi:10.1523/JNEUROSCI.4467-08.2009

Costa, P.T., McCrae, R.R., 1992. Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI). Odessa FL Psychol. Assess. Resour. 3, 101. doi:10.1037//1040-3590.4.1.5

Daneman, M., Carpenter, P. a., 1980. Individual differences in working memory and reading. J. Verbal Learning Verbal Behav. 19, 450–466. doi:10.1016/S0022-5371(80)90312-6

Davis, H., Wright, J., 1979. procedural and parametric variability in studies of conditioned suppression. Bull. Psychon. Soc. 14, 149–150.

den Ouden, H.E.M., Daw, N.D., Fernandez, G., Elshout, J.A., Rijpkema, M., Hoogman, M., Franke, B., Cools, R., 2013. Dissociable Effects of Dopamine and Serotonin on Reversal Learning. Neuron 80, 1090–1100. doi:10.1016/j.neuron.2013.08.030

Dickinson, A., Smith, J., Mirenowicz, J., 2000. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav. Neurosci. 114, 468–483. doi:10.1037/0735-7044.114.3.468

Fahn, S., Elton, R., 1987. Unified Parkinson's Disease Rating Scale (UPDRS), in: Recent Developments in Parkinson's Disease, Vol 2. pp. 153–163.

Fallon, S.J., Cools, R., 2014. Reward Acts on the pFC to Enhance Distractor Resistance of Working Memory Representations. J. Cogn. Neurosci. 26, 2812–2826. doi:10.1162/jocn_a_00676

Faure, A., Reynolds, S.M., Richard, J.M., Berridge, K.C., 2008. Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. J. Neurosci. 28, 7184–7192. doi:28/28/7184 [pii] 10.1523/JNEUROSCI.4961-07.2008

Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., Salkovskis, P.M., 2002. The Obsessive-Complusive Inventory: Development and validation of a short version. Psychol. Assess. 14, 485–495. doi:10.1037//1040-3590.14.4.485

Frank, M.J., 2006. Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. Neural Networks 19, 1120–1136. doi:10.1016/j.neunet.2006.03.006

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. 17, 51–72. doi:10.1162/0898929052880093

Frank, M.J., Fossella, J. a, 2011. Neurogenetics and pharmacology of learning, motivation, and cognition. Neuropsychopharmacology 36, 133–152. doi:10.1038/npp.2010.96

Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science (80-.). 306, 1940–1943. doi:10.1126/science.1102941

Frobose, M.I., Swart, J.C., Cook, J.L., Geurts, D.E., Ouden, H.E. den, Cools, R., 2017. Catecholaminergic modulation of the avoidance of cognitive control. bioRxiv 191015. doi:10.1101/191015

Guitart-Masip, M., Duzel, E., Dolan, R., Dayan, P., 2014a. Action versus valence in decision making. Trends Cogn. Sci. doi:10.1016/j.tics.2014.01.003

Guitart-Masip, M., Economides, M., Huys, Q.J.M., Frank, M.J., Chowdhury, R., Duzel, E., Dayan, P., Dolan, R.J., 2014b. Differential, but not opponent, effects of l-DOPA and citalopram on action learning with reward and punishment. Psychopharmacology (Berl). 231, 955–966. doi:10.1007/s00213-013-3313-4

Guitart-Masip, M., Fuentemilla, L., Bach, D.R., Huys, Q.J., Dayan, P., Dolan, R.J., Duzel, E., 2011. Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. J Neurosci 31, 7867–7875. doi:10.1523/JNEUROSCI.6376-10.2011

Haavik, J., Blau, N., Thöny, B., 2008. Mutations in human monoamine-related neurotransmitter pathway genes. Hum. Mutat. doi:10.1002/humu.20700

Hebart, M.N., Gläscher, J., 2015. Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. Psychopharmacology (Berl). 232, 437–451. doi:10.1007/s00213-014-3682-3

Huijding, J., Muris, P., Lester, K.J., Field, A.P., Joosse, G., 2011. Training children to approach or avoid novel animals: Effects on self-reported attitudes and fear beliefs and information-seeking behaviors. Behav. Res. Ther. 49, 606–613. doi:10.1016/j.brat.2011.06.005

Huys, Q.J.M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R.J., Dayan, P., 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. PLoS Comput. Biol. 7. doi:10.1371/journal.pcbi.1002028

Kakoschke, N., Kemps, E., Tiggemann, M., 2017. Approach bias modification training and consumption: A review of the literature. Addict. Behav. doi:10.1016/j.addbeh.2016.08.007

Kim, J.-H., Son, Y.-D., Kim, H.-K., Lee, S.-Y., Kim, Y.-B., Cho, Z.-H., 2014. Dopamine D(2/3) receptor availability and human cognitive impulsivity: a high-resolution positron emission tomography imaging study with [^{11}C] raclopride. Acta Neuropsychiatr. 26, 35–42. doi:10.1017/neu.2013.29

Kurian, M.A., Gissen, P., Smith, M., Heales, S.J.R., Clayton, P.T., 2011. The monoamine neurotransmitter disorders: An expanding range of neurological syndromes. Lancet Neurol. doi:10.1016/S1474-4422(11)70141-7

Laham, S.M., Kashima, Y., Dix, J., Wheeler, M., Levis, B., 2014. Elaborated contextual framing is necessary for action-based attitude acquisition. Cogn. Emot. 28, 1119–1126. doi:10.1080/02699931.2013.867833

Lee, B., London, E.D., Poldrack, R. a, Farahi, J., Nacca, A., Monterosso, J.R., Mumford, J. a, Bokarius, A. V, Dahlbom, M., Mukherjee, J., Bilder, R.M., Brody, A.L., Mandelkern, M. a, 2009. Striatal dopamine d2/d3 receptor availability is reduced in methamphetamine dependence and is linked to impulsivity. J. Neurosci. 29, 14734–40. doi:10.1523/JNEUROSCI.3765-09.2009

Levitt, M., Spector, S., Sjoerdsma, A., Udenfriend, S., 1965. Elucidation of the Rate-Limiting Step in Norepinephrine Biosynthesis in the Perfused Guinea-Pig Heart. J. Pharmacol. Exp. Ther. 148, 1–8.

Lex, A., Hauber, W., 2008. Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. Learn. Mem. 15, 483–491. doi:10.1101/lm.978708

Lloyd, K., Dayan, P., 2016. Safety out of control: dopamine and defence. Behav. Brain Funct. 12, 15. doi:10.1186/s12993-016-0099-7

Lüdecke, B., Dworniczak, B., Bartholomé, K., 1995. A point mutation in the tyrosine hydroxylase gene associated with Segawa's syndrome. Hum. Genet. 95, 123–125. doi:10.1007/BF00225091

Mesulam, M.M., 1985. Mental state assessment of young and elderly adults in behavioral neurology. In: Principles of behavioral neurology. Davis Company, Philadelphia.

Miller, E.M., Shankar, M.U., Knutson, B., McClure, S.M., 2014. Dissociating motivation from reward in human striatal activity. J. Cogn. Neurosci. 26, 1075–1084. doi:10.1162/jocn_a_00535

Montague, P.R., Hyman, S.E., Cohen, J.D., 2004. Computational roles for dopamine in behavioural control. Nature 431, 760–767. doi:10.1038/nature03015

Moustafa, A.A., Cohen, M.X., Sherman, S.J., Frank, M.J., 2008. A Role for Dopamine in Temporal Decision Making and Reward Maximization in Parkinsonism. J. Neurosci. 28, 12294–12304. doi:10.1523/JNEUROSCI.3116-08.2008

Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: Opportunity costs and the control of response vigor. Psychopharmacology (Berl). 191, 507–520. doi:10.1007/s00213-006-0502-4

Patton, J.H., Stanford, M.S., Barratt, E.S., 1995. Factor structure of the Barratt impulsiveness scale. J. Clin. Psychol. 51, 768–774. doi:10.1002/1097-4679(199511)51:6<768

R Developement Core Team, 2015. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. 1, 409. doi:10.1007/978-3-540-74686-7

Reeves, S.J., Polling, C., Stokes, P.R.A., Lappin, J.M., Shotbolt, P.P., Mehta, M.A., Howes, O.D., Egerton, A., 2012. Limbic striatal dopamine D2/3 receptor availability is associated with non-planning impulsivity in healthy adults after exclusion of potential dissimulators. Psychiatry Res. - Neuroimaging 202, 60–64. doi:10.1016/j.pscychresns.2011.09.011

Robbins, T.W., Everitt, B.J., 2007. A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). Psychopharmacology (Berl). doi:10.1007/s00213-006-0528-7

Salamone, J.D., Correa, M., Mingote, S.M., Weber, S.M., 2005. Beyond the reward hypothesis: Alternative functions of nucleus accumbens dopamine. Curr. Opin. Pharmacol. doi:10.1016/j.coph.2004.09.004

Salthouse, T.A., 1994. The Nature of the Influence of speed on adult age differences in cognition. Dev. Psychol. 30, 240–259. doi:10.1037/0278-7393.20.6.1486

Schmand, B., Bakker, D., Saan, R., Louman, J., 1991. [The Dutch Reading Test for Adults: a measure of premorbid intelligence level]. Tijdschr. Gerontol. Geriatr. 22, 15–19.

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science (80-.). 275, 1593–1599. doi:10.1126/science.275.5306.1593

Schultz, W., Tremblay, L., Hollerman, J.R., 1998. Reward prediction in primate basal ganglia and frontal cortex, in: Neuropharmacology. pp. 421–429. doi:10.1016/S0028-3908(98)00071-9

Sharot, T., 2011. The optimism bias. Curr. Biol. 21. doi:10.1016/j.cub.2011.10.030

Sharot, T., De Martino, B., Dolan, R.J., 2009. How Choice Reveals and Shapes Expected Hedonic Outcome. J. Neurosci. 29, 3760–3765. doi:10.1523/JNEUROSCI.4972-08.2009

Sharot, T., Guitart-Masip, M., Korn, C.W., Chowdhury, R., Dolan, R.J., 2012. How dopamine enhances an optimism bias in humans. Curr. Biol. 22, 1477–1481. doi:10.1016/j.cub.2012.05.053

Sharot, T., Shiner, T., Brown, A.C., Fan, J., Dolan, R.J., 2009. Dopamine Enhances Expectation of Pleasure in Humans. Curr. Biol. 19, 2077–2080. doi:10.1016/j.cub.2009.10.025

Sharot, T., Velasquez, C.M., Dolan, R.J., 2010. Do decisions shape preference? Evidence from blind choice. Psychol. Sci. 21, 1231–1235. doi:10.1177/0956797610379235

Sheehan, D. V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10, in: Journal of Clinical Psychiatry. pp. 22–33. doi:10.1016/S0924-9338(99)80239-9

Stroop, J.R., 1935. Stroop color word test. J. Exp. Physiol. 643–662. doi:10.1007/978-0-387-79948-3

Swart, J.C., Frank, M.., Määttä, J.I., Jensen, O., Cools, R., den Ouden, H.E.M., 2018. Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. PLoS Biol.

Swart, J.C., Froböse, M.I., Cook, J.L., Geurts, D.E.M., Frank, M.J., Cools, R., den Ouden, H.E.M., 2017. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in) action. Elife 6. doi:10.7554/eLife.22169

Taylor, J.R., Robbins, T.W., 1986. 6-Hydroxydopamine lesions of the nucleus accumbens, but not of the caudate nucleus, attenuate enhanced responding with reward-related stimuli produced by intra-accumbens d-amphetamine. Psychopharmacology (Berl). 90, 390–397. doi:10.1007/BF00179197

Taylor, J.R., Robbins, T.W., 1984. Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. Psychopharmacology (Berl). 84, 405–412. doi:10.1007/BF00555222

Verbeek, M.M., Geurtz, P.B.H., Willemsen, M.A.A.P., Wevers, R.A., 2007. Aromatic l-amino acid decarboxylase enzyme activity in deficient patients and heterozygotes. Mol. Genet. Metab. 90, 363–369. doi:10.1016/j.ymgme.2006.12.001

Wassenberg, T., Monnens, L.A., Geurtz, B.P., Wevers, R.A., Verbeek, M.M., Willemsen, M.A., 2012. The paradox of hyperdopaminuria in aromatic L-amino Acid deficiency explained. JIMD Rep. 4, 39–45. doi:10.1007/8904_2011_84 [doi]

Wechsler, D., 1997. Wechsler memory scale—Third edition. Psychol. Corp. San Antonio, TX.

Willemsen, M.A., Verbeek, M.M., Kamsteeg, E.J., De Rijk-Van Andel, J.F., Aeby, A., Blau, N., Burlina, A., Donati, M.A., Geurtz, B., Grattan-Smith, P.J., Haeussler, M., Hoffmann, G.F., Jung, H., De Klerk, J.B., Van Der Knaap, M.S., Kok, F., Leuzzi, V., De Lonlay, P., Megarbane, A., Monaghan, H., Renier, W.O., Rondot, P., Ryan, M.M., Seeger, J., Smeitink, J.A., Steenbergen-Spanjers, G.C., Wassmer, E., Weschke, B., Wijburg, F.A., Wilcken, B., Zafeiriou, D.I., Wevers, R.A., 2010. Tyrosine hydroxylase deficiency: A treatable disorder of brain catecholamine biosynthesis. Brain 133, 1810–1822. doi:10.1093/brain/awq087

Wise, R.A., 2004. Dopamine, learning and motivation. Nat. Rev. Neurosci. 5, 483–494. doi:10.1038/nrn1406

Woud, M.L., Becker, E.S., Rinck, M., 2008. Implicit evaluation bias induced by approach and avoidance. Cogn. Emot. 22, 1187–1197. doi:10.1080/02699930801906843

Woud, M.L., Maas, J., Becker, E.S., Rinck, M., 2013. Make the manikin move: Symbolic approach-avoidance responses affect implicit and explicit face evaluations. J. Cogn. Psychol. 25, 738–744. doi:10.1080/20445911.2013.817413

Wyvell, C.L., Berridge, K.C., 2000. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. J. Neurosci. 20, 8122–8130. doi:http://www.jneurosci.org/content/21/19/7831

Zafeiriou, D.I., Willemsen, M.A., Verbeek, M.M., Vargiami, E., Ververi, A., Wevers, R., 2009. Tyrosine hydroxylase deficiency with severe clinical course. Mol. Genet. Metab. 97, 18–20. doi:10.1016/j.ymgme.2009.02.001

Zigmond, a S., Snaith, R.P., 1983. The hospital anxiety and depression scale (HADS). Acta Psychiatr. Scand. 67, 361–370. doi:10.1016/S0016-5085(01)83173-5

Chapter 3

Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated action

Politics is for the present, but an equation is something for eternity.
- Albert Einstein

Published as:
Swart JC, Froböse MI, Cook JL, Geurts DEM, Frank MJ, Cools R*, & den Ouden HEM* (2017). Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. eLife 2017; 6:e22169. *Shared last authorship.

Abstract

Catecholamines modulate the impact of motivational cues on action. Such motivational biases have been proposed to reflect cue-based, 'Pavlovian' effects. Here, we assess whether motivational biases may also arise from asymmetrical instrumental *learning* of active and passive responses following reward and punishment outcomes. We present a novel paradigm, allowing us to disentangle the impact of reward and punishment on instrumental learning from Pavlovian response biasing. Computational analyses showed that motivational biases reflect both Pavlovian and instrumental effects: reward and punishment cues promoted generalized (in)action in a Pavlovian manner, whereas outcomes enhanced instrumental (un) learning of chosen actions. These cue- and outcome-based biases were altered independently by the catecholamine enhancer melthylphenidate. Methylphenidate's effect varied across individuals with a putative proxy of baseline dopamine synthesis capacity, working memory span. Our study uncovers two distinct mechanisms by which motivation impacts behaviour, and helps refine current models of catecholaminergic modulation of motivated action.

eLife digest

When we see a threat, we tend to hold back. When we see a reward, we have a strong urge to approach. Most of the time, these hardwired tendencies – or biases – are the right thing to do. However, our behaviour is not all hardwired; we can also learn from our previous experiences. But might this learning be biased too? For example, we might be quicker to believe that an action led to a reward, because actions often do bring rewards. Conversely, we might be less likely to attribute a punishment to having held back, because holding back usually helps us to avoid punishments.

Swart et al. have now tested whether rewards and punishments influence our actions solely via hardwired behavioural tendencies, or whether they also bias our learning. That is, are we biased to learn that taking action earns us rewards, while holding back spares us punishments? Previous work has shown that chemical messengers in the brain called catecholamines help us to take action when we anticipate a reward. Swart et al. therefore also examined whether catecholamine levels contribute to any bias in learning.

One hundred young healthy adults twice performed a task in which they could earn rewards and avoid losses by taking or withholding action. By using a mathematical model to work out what influenced the choices made by the volunteers, Swart et al. found that rewards and punishments did indeed bias learning. Moreover, this learning bias became stronger when the volunteers took methylphenidate (also known as Ritalin), a drug that increases catecholamine levels and which is used to treat ADHD and narcolepsy. The volunteers varied markedly in how strongly methylphenidate affected their choices. This emphasises how important it is to account for differences between people when evaluating the effects of medication.

Motivations are what get us going and keep us going. The findings of Swart et al. mean that we now have a better understanding of how motivations, such as desired rewards or unwanted punishments, influence our behaviour. A future challenge is to understand how we can overcome these motivations when they work against us, such as in addiction or obesity.

Introduction

Catecholamine (i.e. dopamine and noradrenaline) transmission has long been implicated in key aspects of adaptive behaviour, including learning, action, and motivation. Deficits in these aspects of adaptive behaviour are observed in a wide range of neuropsychiatric disorders, such as attention deficit hyperactivity disorder, Parkinson's disease, and addiction (Dagher and Robbins, 2009; Prince, 2008; Skolnick, 2005), and many of those deficits can be treated with catecholaminergic drugs (Faraone and Buitelaar, 2010; Wigal et al., 2011). While overwhelming evidence implicates catecholamines in both motivated activation and motivated learning of behaviour (Bromberg-Martin et al., 2010; Robbins and Everitt, 1996; Wise, 2004), their respective contributions are still highly debated. In this study, we dissect the contribution of catecholamines to motivational biases in behavioural activation and learning.

The neuromodulator dopamine has been linked particularly strongly to behavioural activation in the context of reward (Taylor and Robbins, 1986, 1984), putatively by amplifying the perceived benefits of action over their costs (Collins and Frank, 2014; Niv et al., 2007). This behavioural activation to reward-predicting cues is likely to be, at least partly, Pavlovian in nature, with the conditioned cues eliciting innately specified responses (Figure 1A). The Pavlovian nature of these motivational biases has been demonstrated using Pavlovian-instrumental transfer (PIT) paradigms (Estes, 1943; Estes and Skinner, 1941). In PIT, conditioned cues elicit innately specified responses that may potentiate (or interfere with) instrumental responding, e.g. appetitive cues promote active responding (appetitive PIT), whereas aversive cues increase behavioural inhibition (aversive PIT; Davis and Wright, 1979; Huys et al., 2011). Enhanced dopamine increases appetitive PIT (Wyvell and Berridge, 2000), while appetitive PIT is lowered when striatal dopamine is reduced (Dickinson et al., 2000; Hebart and Gläscher, 2015; Lex and Hauber, 2008). Striatal dopamine has also been linked to controlling aversively motivated behaviour (Faure et al., 2008; Lloyd and Dayan, 2016). Together, these results show that appetitive cues promote activation and aversive cues promote inhibition in a Pavlovian manner, mediated by the dopamine system.

While Pavlovian response biases can be helpful in reducing computational load by shaping our actions in a hardwired manner, they are inherently limited because of their general nature (Dayan et al., 2006). In order to select the best action in a specific environment, instrumental systems allow organisms to adaptively learn action-outcome contingencies, by assigning value to actions that in the past have led to good outcomes, while reducing value of actions that led to negative outcomes (Dickinson and Balleine, 1994; Rescorla and Wagner, 1972; Robbins and Everitt, 2007).

Pavlovian and instrumental learning are often presented as a dichotomy, whereby cue-based, Pavlovian effects are solely responsible for motivational biases, while adaptive 'rational' choice results from instrumental learning. For example, multiple recent studies showing that reward or punishment cues bias action, eliciting appetitive activation and/or aversive inhibition, have been interpreted specifically in terms of a Pavlovian response bias (for a review see Guitart-Masip et al., 2014a).

We hypothesised that these motivational biases of behavioural activation may also arise from asymmetrical, or biased, instrumental learning (Figure 1B), in addition to Pavlovian response biases. Such biases in learning, like response biases, may reflect predominant statistics of the environment. For example, we might be quicker to believe that an action led to a reward, because actions often cause rewards. However, we may not attribute a punishment to having held back, because holding back usually helps avoid a punishment. Such an instrumental learning bias may arise from a circuitry where reinforcers are more potent at driving learning following active 'Go' than inactive 'NoGo' actions. This means that Go responses (relative to NoGo responses) are easier to learn and unlearn following reward and punishment respectively. This instrumental learning bias would predict that Go responses that elicited a reward are more likely to be repeated (i.e. better learned) than NoGo responses that elicited a reward. Similarly, Go responses that elicited a punishment are relatively less likely to be repeated (i.e. better unlearned) than NoGo responses that elicited a punishment. These instrumental learning and Pavlovian response biasing accounts of motivated (in)action could not be dissociated in earlier studies (Cavanagh et al., 2013; Guitart-Masip et al., 2014b, 2012), because they allowed for only a single Go response: With only one response option, general activation of action cannot be disentangled from facilitated learning of a specific response. In our proposed framework, motivational biases in behavioural (in) activation are likely the result of a combination of Pavlovian response biasing plus an asymmetry in instrumental learning of Go and NoGo responses (Figure 1).

At the neurobiological level, this hypothesis arises from current theorizing about the mechanism of action of reinforcement-related changes in dopamine. Specifically, a potential substrate for this proposed learning asymmetry could be provided by the striatal dopamine system, which is notably involved in instrumental learning via modulation of synaptic plasticity (Collins and Frank, 2014 for review and models). In this framework, dopamine bursts elicited by better than expected outcomes reinforce the actions that led to these outcomes (Montague et al., 2004; Schultz et al., 1998, 1997) via long-term potentiation (LTP) in the striatal direct 'Go' pathway (Frank et al., 2004). The temporal specificity of the phasic dopamine bursts allows for assigning credit to the most recent action, by potentiating the recently active striatal neurons. Due to the LTP in the 'Go' pathway, rewards may be more effective in reinforcing neurons coding for active Go responses than NoGo responses. Conversely, dopamine dips elicited by worse-than-expected outcomes (Matsumoto and Hikosaka, 2009; Tobler et al., 2005) lead to long-term depression (LTD) of the 'Go' pathway and LTP in the 'NoGo' pathway, making it less likely that the same cue would

elicit an active than inactive response next time. In short, the striatal system is biased to attribute rewards and punishments to active Go responses, which ecologically may be more commonly the cause of observed outcomes. The implication of this is that is easier to learn to take action based on reward, but easier to withhold making an action based on punishment.

A key additional prediction of this model is that prolonging the presence of dopamine, e.g. by blocking dopamine reuptake with methylphenidate, would lead to a spread of credit assignment (Figure 1C). Here, credit is assigned to striatal neurons that were recently active, due to recent actions that did not actually lead to the current reward and phasic dopamine burst ("spread of effect"; Thorndike, 1933). In this framework, the dopamine system can produce biased motivated behaviour due to i) direct Pavlovian biases (e.g. predicted rewards potentiate the Go pathway during action selection), and ii) disproportionate potentiation of instrumental learning of Go actions that (recently) led to reward. Put more simply, i) dopamine bursts prompted by reward-predicting cues can potentiate activation of the Go pathway, giving rise to the cue-based, Pavlovian activation, and ii) dopamine bursts prompted by reward outcomes can potentiate plasticity within the Go pathway, making rewards more effective in reinforcing Go responses than NoGo responses.

In this study, we aimed to assess whether biases in instrumental learning contribute to the pharmaco-computational mechanisms subserving well-established reward/punishment biases of motivated (in)action. To dissociate biased instrumental learning from Pavlovian response biases, we developed a novel experimental paradigm including multiple active response options (Figure 2), and combined this task with a catecholamine challenge (catecholamine reuptake blocker methylphenidate - MPH). We tested the following hypotheses: i) cue-valence (appetitive vs. aversive cues) biases action in a Pavlovian manner, whereas outcome-valence (reward vs. punishment) biases instrumental learning of Go vs. NoGo actions; ii) blocking the catecholamine reuptake with MPH enhances the strength of the Pavlovian response bias as a result of prolonged dopamine release to reward cues; iii) MPH reduces the specificity of credit assignment to specific actions that elicited rewards, as the prolonged DA release to reward outcomes would spread credit to non-chosen active actions (Figure 1).

Finally, MPH prolongs the effects of catecholamine release by blocking the reuptake of catecholamines, without stimulating release or acting as a receptor (ant)agonist (e.g. Volkow et al., 2002). Accordingly, it is likely that the effect of MPH on catecholamine-dependent function is a function of dopamine synthesis capacity and release. Simply put, if there is no release, there is no reuptake to block. To assess these potential sources of individual variability in MPH effects, we took into account two measures that have been demonstrated with PET to relate to dopamine baseline function: working memory span for its relation to striatal dopamine synthesis capacity (Cools et al., 2008; Landau et al., 2009) and trait impulsivity for its relation to dopamine (auto) receptor availability (Buckholtz et al., 2010; Kim et al., 2014; Lee et al., 2009; Reeves et al., 2012), and collected a large sample (N=106) to expose individual differences.

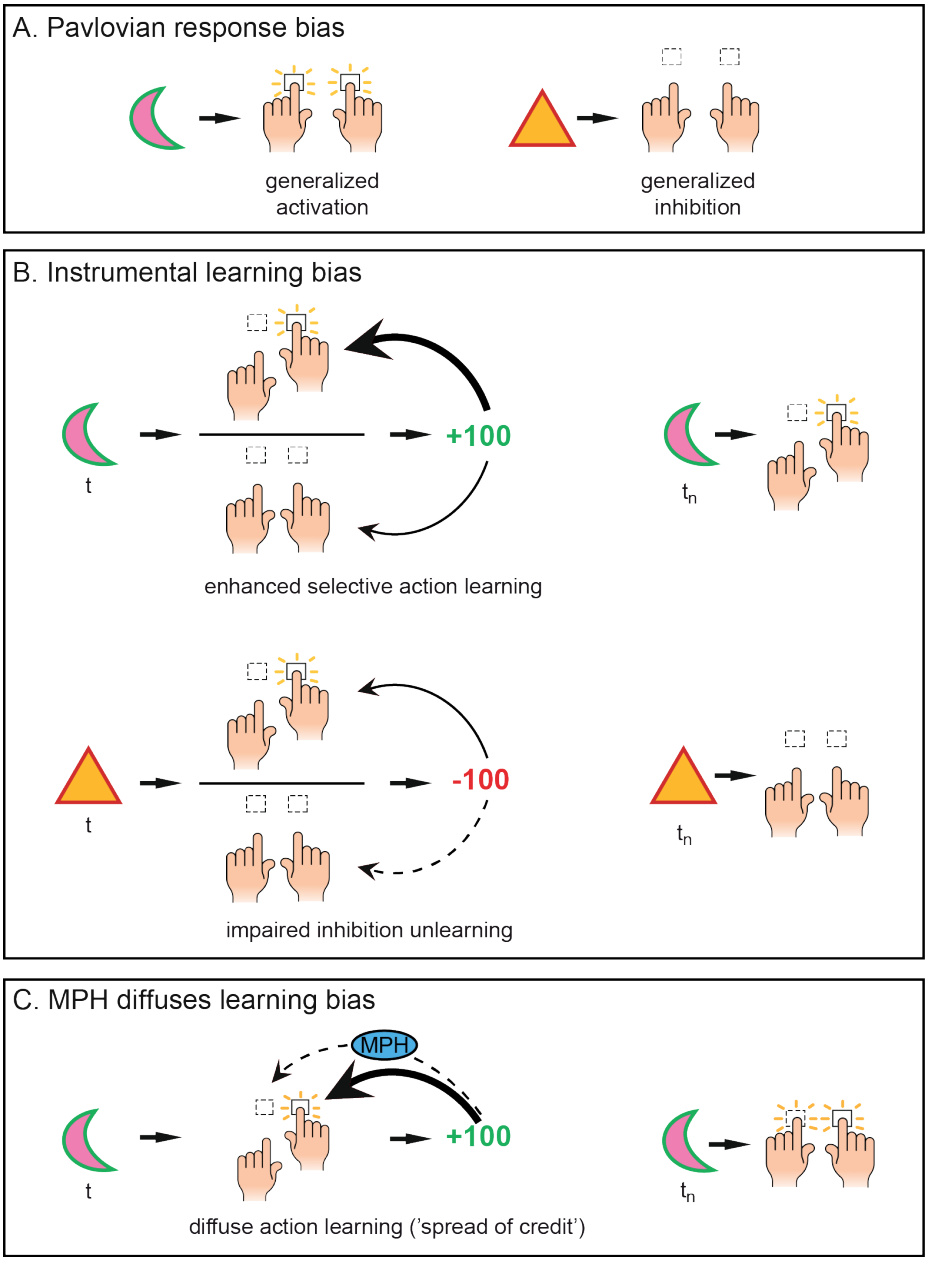


Figure 1. **Distinct mechanisms by which motivational valence may bias behavioural activation.** (**a**) Pavlovian response bias: appetitive cues (green edge) elicit generalized behavioural activation ('Go'), whereas aversive cues (red edge) elicit behavioural inhibition ('NoGo'). This Pavlovian response bias is introduced in model M3a as the parameter π (c.f. Figure 3). (**b**) Instrumental learning bias: rewarding outcomes (upper panel) facilitate learning of action ('Go', thick arrow) relative to inaction ('NoGo', thin arrow). Thus, learning effects at the individual trials *t* will result in a cumulative selective increase of the rewarded action on later trials t_n. Punishment outcomes (lower panel) hamper the unlearning of inaction

('NoGo', dashed arrow) relative to action ('Go', solid arrow), resulting in sustained inaction. Neutral outcomes are equally well associated with actions and inactions, and are not illustrated here. The instrumental learning bias is introduced as the parameter κ in model M3b (c.f. Figure 3). We assess whether these two mechanisms (i) act in parallel, and (ii) are modulated by the catecholamine system. To test the latter, we administered methylphenidate (MPH), which prolongs the effects of catecholamine release via blockade of the catecholamine receptors. We first assess whether MPH affects the strength of the Pavlovian response bias, introduced as the parameter π_{MPH} in model M5a, and instrumental learning bias, implemented as the parameter $\kappa_{MPH\text{-}selective}$ in model M5b (c.f. Figure 5). (**c**) We hypothesise that prolonged effects of dopamine release following reward outcomes will reduce (temporal) specificity, leading to spread of credit: Credit is assigned to other recent actions (thin arrow), in addition to the performed (and rewarded) Go response (thick arrow), resulting in additional learning of the alternative (not-performed) Go response. This MPH-induced diffuse learning bias is implemented by the parameter $\kappa_{MPH\text{-}diffuse}$ in model M5c (c.f. Figure 5).

Results

Healthy participants performed a motivational Go/NoGo learning task, in which cue valence (Win vs. Avoid cue) is orthogonalized to the instrumental response (Go vs. NoGo). During this task, subjects need to learn for each of 8 cues to make a Go or NoGo response, and by making the correct response subjects are rewarded for Win cues (green edge) and avoid punishment for the Avoid cues (red edge) in 80% of the time. Crucially, in contrast to task designs in previous studies (Guitart-Masip et al., 2014a), in this novel paradigm subjects could make either of two Go responses (press left vs. right) or withhold responding (NoGo; Figure 2A-D). Including two Go response options enabled us to tease apart general activation/inhibition related to the Pavlovian response bias and specific action learning related to the instrumental learning bias using computational models and behavioural analyses.

Motivational Valence affects (in)correct action

Subjects successfully learned this difficult task, in which they needed to identify the correct response out of 3 options (Go-left/Go-right/NoGo) for 8 different cues, as evidenced by increased Go responding to cues indicating the need to Go vs. NoGo (Required Action: X^2_1=624.3; p<.001; Figure 2E,F). In other words, subjects were able to adjust Go responding to the required action. As expected, cue valence also influenced Go responding (Valence: X^2_1=40.0; p<.001), reflecting a motivational bias in responding. Overall subjects made more Go responses for Win than Avoid cues. The effect of cue valence was highly significant for both Go and NoGo cues (Go cues: X^2_1=37.5, p<.001; NoGo cues: X^2_1=13.3, p<.001), though marginally stronger for the Go cues (Required Action x Valence: X^2_1=3.1; p=.076). Because each Go cue was associated with only one correct Go response, we confirmed that this motivational bias was present for both correct and incorrect Go responses. Subjects made more Go responses to Win than avoid cues for both correct (Valence: X^2_1=26.1, p<.001) and incorrect (Valence: X^2_1=25.6, p<.001) Go responses. Next, we tested the hypothesis that this motivational bias arose from a combination of a Pavlovian response bias and biased instrumental learning (Figure 1A-B).

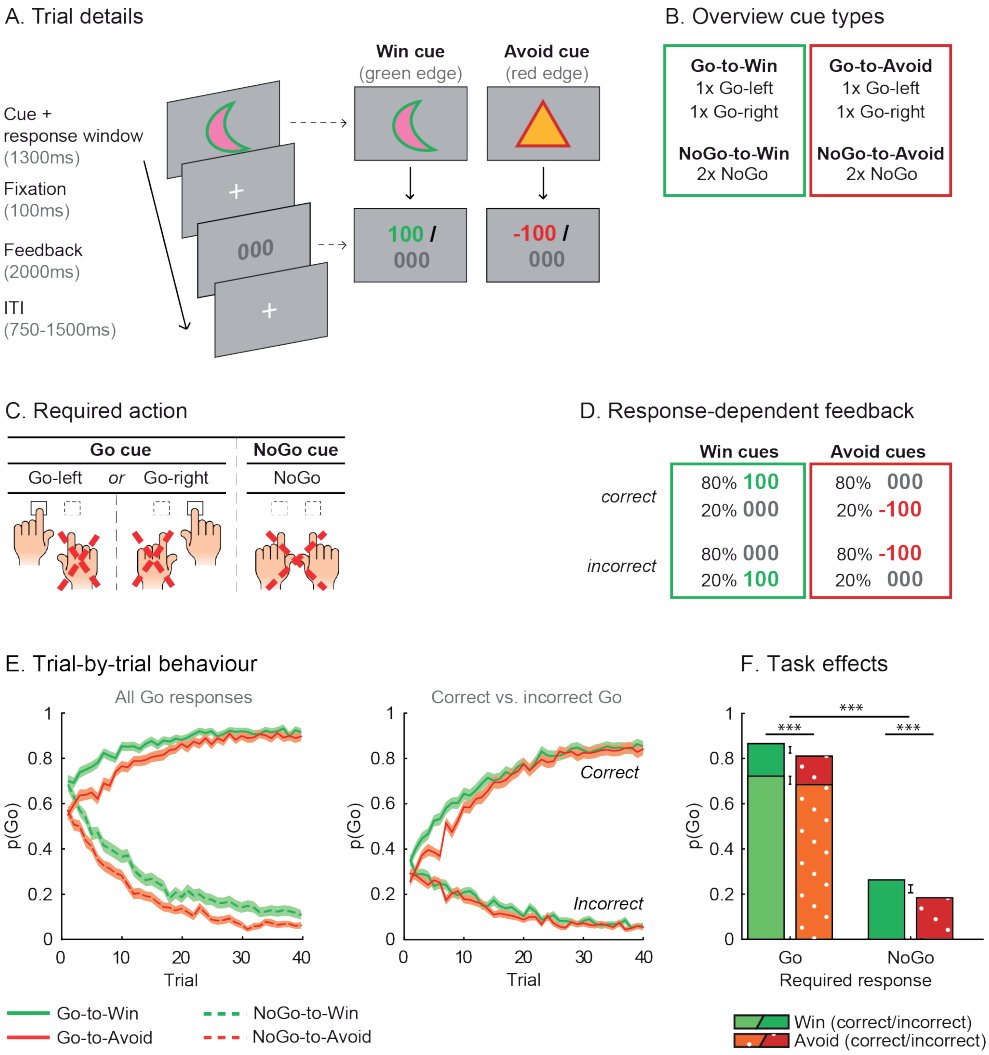


Figure 2. **Motivational Go/NoGo learning task and performance.**
(**a**) On each trial, a Win or Avoid cue appears on screen. Subjects can respond during cue presentation. Response-dependent feedback follows. (**b**) In total eight cues are presented for which the correct response needs to be learned. (**c**) Each cue has only one correct response (Go-left, Go-right, or NoGo), which subjects can learn from the feedback. (**d**) Feedback is probabilistic. Correct responses are followed by reward (Win cues) or a neutral outcome (Avoid cues) in 80% of the time and by a neutral outcome (Win cues) or punishment (Avoid cues) otherwise. For incorrect responses, these probabilities are reversed. (**e**) Trial-by-trial proportion of Go responses (±SEM) for Go cues (solid lines) and NoGo cues (dashed lines), collapsed over Placebo and MPH. Left: All cue types. From the first trial onwards, subjects made more Go responses to Win vs. Avoid cues (i.e. green lines are above red lines), reflecting the motivational bias. Additionally, subjects clearly learn whether to make a Go response or not (proportion of Go responses increases for Go cues and decreases for NoGo cues). Right: Go cues only. For the Go cues, a Go response could be either correct or incorrect. The motivational bias is present in both correct and incorrect Go responses, but incorrect Go responses are unlearnt. Note that the total p(Go) in this plot sums up to the solid lines in the left plot. (**f**) Mean (±SED) proportion Go responses. Proportion Go responses is higher for

Go vs. NoGo cues, indicative of task learning. Additionally, subjects made more correct and incorrect Go responses to Win vs. Avoid cues.
Figure supplement:

- *Figure 2–figure supplement 1 (p.73).* Individual traces (black lines) and group average (coloured lines) of correct and incorrect Go responses using a sliding average of 5 trials.

Computational modelling: disentangling Pavlovian response bias and instrumental learning bias

We used a computational modelling approach to quantify latent processes that we hypothesised to underlie the behavioural performance. Specifically, our first aim was to disentangle the contribution of Pavlovian response biases and instrumental learning biases to the observed valence effect in behaviour. To this end we extended a simple reinforcement learning model using hierarchical Bayesian parameter estimation. We developed five nested base models (M1, M2, M3a, M3b, M4) with increasing complexity to assess whether additional parameters explained the observed data better, while penalizing for increasing complexity.

In all models, the probability of each response is estimated based on computed action weights. In the simplest model (M1) the action weights are fully determined by the learned action values (*Q*-values). Action values are updated with the prediction error, i.e. the deviation of the observed outcome from the expected outcome (standard "delta-rule" learning; Rescorla and Wagner, 1972). M1 contains two free parameters: a learning rate (ε) scaling the impact of the prediction-error, and feedback sensitivity (ρ) scaling the outcome value. Next, to allow for a non-selective bias in Go responses unrelated to valence, a go bias parameter (*b*) is added to the action weights of Go responses in M2. This parameter simply captures how likely people are to make a 'Go' response overall.

In this task, we explicitly instructed the cue valence, by colouring the edge of each cue, where green signalled that subjects could win a reward, while red signalled they had to avoid a punishment (Figure 2A). As a consequence, we observed an effect of the instructed cue valence on Go responses already from the first trial onwards (Figure 2E), implying a motivational bias before learning could occur, which is therefore likely Pavlovian in nature. To assess this Pavlovian response bias, cue values are added to the action weights in M3a. In this model positive (negative) Pavlovian values increase (decrease) the action weight of Go responses, where π scales the weight of the Pavlovian values (Cavanagh et al., 2013; Guitart-Masip et al., 2014b, 2012). Thus, the Pavlovian bias parameter increases the probability of all Go responses for Win cues and decreases the probability of all Go responses for Avoid cues.

In M3b we assessed whether a motivational learning bias affects behaviour. Specifically, we included an instrumental learning bias parameter (κ), to assess whether reward is more effective in reinforcing Go responses than NoGo responses, whereas

punishment is less effective in unlearning NoGo responses than Go responses. This biased learning parameter indexes the degree to which the *specific* Go response that elicited a reward would be relatively more likely to be repeated in subsequent trials, resulting in increased instrumental learning of Go responses for reward. Note that earlier studies used only a single Go response and could thus not dissociate this specific learning vs. Pavlovian bias account. In addition to this effect on learning from rewards, κ indexes the degree to which punishment is biased to potentiate activity in the NoGo versus Go pathway, thus biasing unlearning to be more effective after Go responses than after NoGo responses, (i.e., making punishment-based avoidance learning of NoGo responses more difficult than punishment-based avoidance learning of Go responses; Figure 1B). Because the Pavlovian and instrumental learning bias might explain similar variance in the data, we tested model M4, where we included both π and κ to test whether there was evidence for the independent presence of both the instrumental learning bias and the Pavlovian response bias.

Stepwise addition of the go bias (Appendix 5), Pavlovian response bias and instrumental learning bias parameter improved model fit, as quantified by Watanabe-Akaike Information Criteria (WAIC; Figure 3; Table 1). The Pavlovian bias parameter estimates (π) of the winning model M4 were positive across the group (96.4% of posterior distribution > 0). The Pavlovian bias estimates were modest across the group (Figure 3; Table 1), and showed strong individual variability (Figure 3 – Figure Supplement 2; Figure 3 – Figure Supplement 3). This strong inter-individual variability is consistent with previous reports, e.g. Cavanagh et al., (2013), who show that differences in the strength of the Pavlovian bias is inversely predicted by EEG mid-frontal theta activity during incongruent relative to congruent cues, putatively reflecting the ability to suppress this bias on incongruent trials. The further improvement of model fit due to the instrumental learning bias parameter (M3a vs. M4) provides clear evidence for the contribution of biased action learning on top of the Pavlovian response bias described in previous studies. The biased instrumental learning parameter estimates were also positive across the group (100% of posterior distribution > 0). In other words, in the winning model, the motivational bias, as reflected by an increase in Go responses to Win relative to Avoid cues, is explained by the presence of both a Pavlovian response bias and biased instrumental learning. Figure 3 and accompanying Figure supplements illustrate the model predictions and parameter estimates.

	Base models				
	M1	**M2**	**M3a**	**M3b**	**M4**
WAIC	71014	69038	67678	67602	66987
ρ	42.7 [19.3 79.8]	41.6 [18.7 72.4]	35.2 [15.8 66.4]	33.4 [13.9 59.8]	32.5 [14.9 56.4]
ε_0	0.013 [0.008 0.059]	0.015 [0.008 0.054]	0.017 [0.009 0.064]	0.022 [0.010 0.070]	0.021 [0.010 0.063]
b		−0.25 [−0.45 0.04]	−0.25 [−0.46 0.04]	.01 [−0.33 0.27]	−0.03 [−0.29 0.19]
π			0.47 [0.02 1.00]		0.12 [−0.29 0.70]
$\varepsilon_{\text{rewarded Go}}$ $(\varepsilon_0+\kappa)$				0.037 [0.016 0.122]	0.034 [0.016 0.109]
$\varepsilon_{\text{punished NoGo}}$ $(\varepsilon_0-\kappa)$				0.006 [0.002 0.014]	0.008 [0.003 0.022]

Table 1. **Base models.**
Median [25–75 percentile] of subject-level parameter estimates in model space. See Appendix 5 for subject-level / top-level parameters in sampling space (i.e. untransformed). Absolute WAIC is reported at the top as the estimate of model evidence, where a smaller WAIC indicates higher evidence.

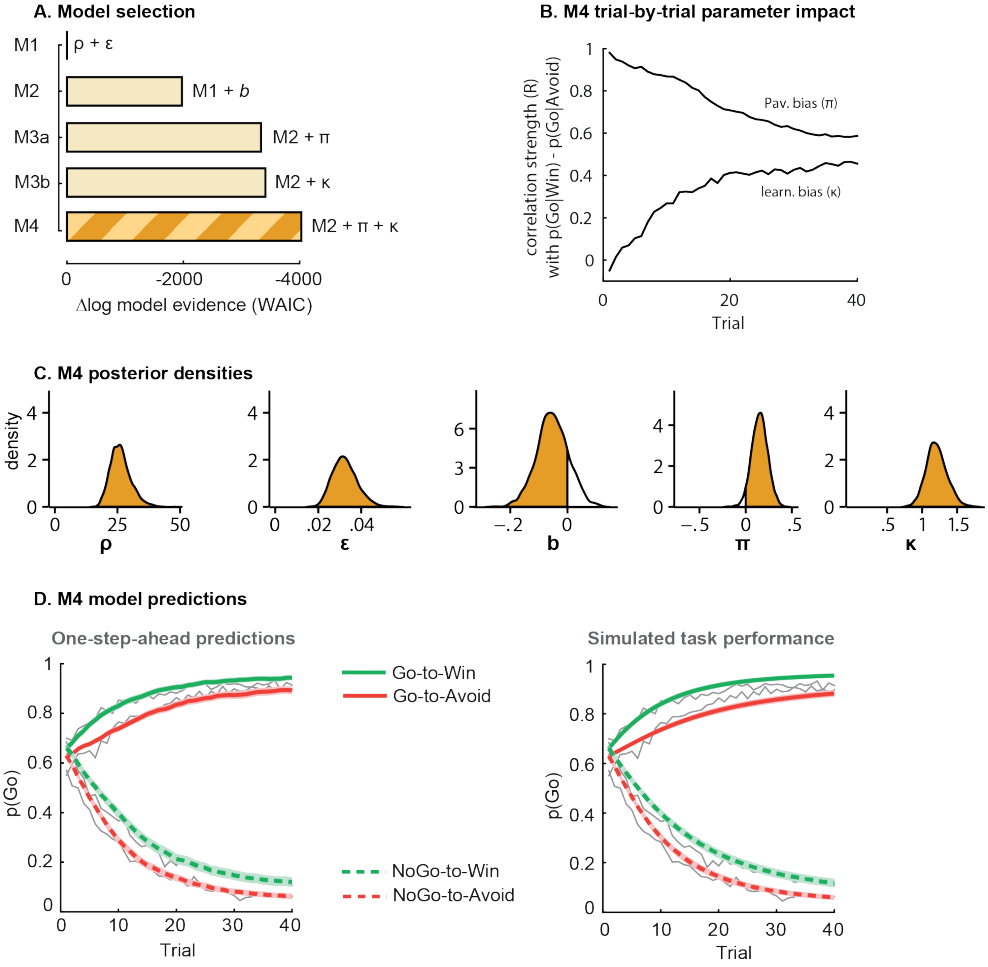


Figure 3. **Model evidence and parameter inference of base models.**
(**a**) Model evidence, relative to simplest model M1, clearly favours M4. The simplest model M1 contains a feedback sensitivity (ρ) and learning rate (ε) parameter. Stepwise addition of the go bias (*b*), Pavlovian bias (π; Figure 1A), and instrumental learning bias (κ; Figure 1B) parameter improves model fit, quantified by WAIC (estimated log model evidence). Lower (i.e. more negative) WAIC indicates better model fit. (**b**) Temporal dynamics of the correlation between the motivational bias parameters (M4) and the predicted motivational bias, i.e. probability to make a Go response to Win relative to Avoid cues. The impact of the Pavlovian bias (π) on choice decreases over time (although, importantly, the parameter itself remains constant). This is because the instrumental values of the actions are learnt and thus will increasingly diverge. As a result, π is less and less 'able' to tip the balance in favour of the responses in direction of the motivational bias (i.e. it can no longer overcome the difference in instrumental action values). In contrast, the impact of κ on choice increases over time, reflecting the cumulative impact of biased learning (also Figure 3—figure supplement 2). (**c**) Posterior densities of the winning base model M4. Appendix 5 shows posterior densities for all models. (**d**) One-step-ahead predictions and posterior predictive model simulations of winning base model M4 (coloured lines), to assess whether the winning model captures the behavioural data (grey lines). Both absolute model fit methods use the fitted parameters to compute the choice probabilities according to the model. The one-step-ahead predictions compute probabilities based on the history of each subject's actual choices and outcomes, whereas the simulation method

generates new choices and outcomes based on the response probabilities (see Materials and methods for details). Both methods capture the key features of the data, i.e. responses are learnt (more 'Go' responding for 'Go' cues relative to 'NoGo' cues) and a motivational bias (more Go responding for Win relative to Avoid cues). We note that the model somewhat underestimates the initial Pavlovian bias (i.e. difference in Go responding between Win and Avoid trials is, particularly trial 1–2), while it overestimates the Pavlovian bias on later trials. This is likely the result from the fact that while the modelled Pavlovian bias parameter (π) is constant over time, the impact of the Pavlovian stimulus values weakens over time, as the subjects' confidence in the instrumental action values increases. Interestingly, notwithstanding the constancy of the Pavlovian bias parameter, we do capture some of these dynamics as Figure 3B shows that the impact of the Pavlovian bias on choice decreases over time.

Figure supplements:

- *Figure 3 – figure supplement 1 (p.73).* Subject traces of model M4 (green/red) overlaid on observed behaviour (black).
- *Figure 3 – figure supplement 2 (p.74).* Illustration of the behavioural effects associated with the Pavlovian bias and instrumental learning bias parameters.
- *Figure 3 – figure supplement 3 (p.74).* M4 subject-level parameters in model spac (i.e. untransformed).

MPH enhances effect of cue valence proportional to working memory span

Next, we asked whether acute administration of MPH altered the motivational bias. As noted above, the effects of dopaminergic drugs often depend on baseline dopamine function. We therefore used two neuropsychological measures that have been shown to predict baseline dopamine function using PET: working memory span, predictive of baseline dopamine synthesis capacity (Cools et al., 2008; Landau et al., 2009), and trait impulsivity, predictive of D2 autoreceptor availability (Buckholtz et al., 2010; Kim et al., 2014; Lee et al., 2009; Reeves et al., 2012). Importantly, both working memory span and trait impulsivity predict dopaminergic drugs effects on various cognitive functions (Clatworthy et al., 2009; Cools et al., 2009, 2007; Frank and O'Reilly, 2006; Gibbs and D'Esposito, 2005; Kimberg et al., 1997; van der Schaaf et al., 2013).

MPH enhanced the effect of cue valence on Go responding proportional to working memory span (Valence x Drug x Listening Span: X^2_1=5.9; p=.016; Figure 4B), in the absence of a Valence x Drug effect across the group (Valence x Drug: X^2_1=1.5; p=.221; Figure 4A). While high-span subjects showed a drug-induced increase in motivational bias (MPH versus placebo increased Go responding to Win vs. Avoid cues), low-span subjects showed a drug-induced decrease in motivational bias. This span-dependent bias emerged under MPH (X^2_1=4.6, p=.032), and was not significant under placebo (X^2_1=.9, p=.335; Figure 4 – Figure supplement 1).

A break-down of this effect into correct and incorrect responses revealed that it was driven by *incorrect* Go responses (Valence x Drug x Listening Span: X^2_1=11.9, p<.001; Figure 4C). MPH did not significantly affect the *correct* Go responses (Valence x Drug x Listening Span: X^2_1=2.0, p=.152). In other words, higher span subjects were more likely to make Go responses to Win cues under MPH, but this Go response was more likely to be incorrect. We reasoned that an enhanced learning bias would manifest primarily in increased *correct* Go responses to Win cues (i.e. the correct responses are better learned), while an enhanced Pavlovian bias or diffusion of

credit assignment would manifest in increased *correct* and *incorrect* Go responses to Win cues (due to overall action invigoration and potentiation respectively). Thus, we expected that the altered valence effect on *incorrect* Go responses under MPH can best be attributed to MPH alteration of Pavlovian response bias or diffusion of credit assignment, which we formally test using computational modelling (see below).

In contrast to listening span, trait impulsivity did not significantly predict the effect of MPH on the motivational bias (all $p>.05$; see Appendix 3 for an overview of the mixed model effects). We confirmed that the MPH effects were not explained by session effects, i.e. whether MPH was received on the first or second testing day ($X^2_2=2.1$, $p=.349$), nor did the factor Testing day improve model fit ($X^2_1=2.0$, $p=.162$). Finally, we confirmed that including nuisance variables Gender and NLV scores (measuring for verbal intelligence), did not improve model fit either ($X^2_2=0.4$, $p=.815$).

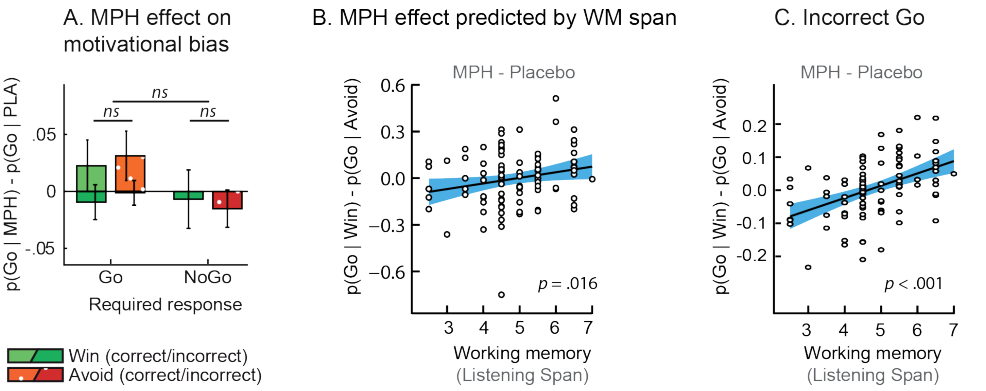


Figure 4. **MPH-induced changes in motivational bias (i.e. proportion of Go responses to Win relative to Avoid cues).**
(**a**) Mean (±SED) proportion Go responses under MPH relative to Placebo. MPH did not significantly alter the motivational bias across the group ($p=0.22$; *ns* indicates $p>0.05$). (**b**) MPH increased the motivational bias in high span subjects, yet decreased it in low span subjects (R=0.21; $p=0.016$). (**c**) MPH altered the motivational bias particularly for *incorrect* Go proportional to working memory span (incorrect Go: $p<0.001$; *correct* Go: $p=0.152$).
Figure supplement:

- *Figure 4 – figure supplement 1 (p.75)*. Simple effects of MPH-induced changes in motivational bias.

Computational modelling: dissociable effects of MPH on Pavlovian response bias and biased instrumental learning

Continuing our modelling approach, we next assessed whether the MPH-induced motivational bias could be attributed to an altered Pavlovian response bias and/or instrumental learning bias. To this end we extended the winning base model M4 into competing models. In M5a we included an MPH-induced Pavlovian bias parameter (π_{MPH}), to assess whether MPH altered the

Pavlovian response bias. Here π_{MPH} alters the individual's Pavlovian bias parameter under MPH. In M5b we included an MPH-induced instrumental learning bias ($\kappa_{MPH\text{-}selective}$). Thus, M5b tests whether MPH affects the strength of the instrumental learning bias in individuals. We further tested whether MPH might make the learning bias more diffuse, because of its mechanisms of action. Because MPH blocks reuptake, it prolongs dopamine release, such that reinforcement and synaptic potentiation might not be attributed only to the temporally coincident neurons that code for the recently selected action, but could be spread to other actions (*diffuse learning*). To test this hypothesis, M5c contains a MPH-induced diffuse learning bias ($\kappa_{MPH\text{-}diffuse}$), where $\kappa_{MPH\text{-}diffuse}$ is a learning rate that alters the value of all Go responses following a reward, under MPH (Figure 1C) by scaling the prediction error following all rewarded Go responses.

Model fit improved markedly when extending the winning base model M4 with the MPH-induced Pavlovian bias parameter π_{MPH} (M5a; Figure 5; Table 2). Extending M4 with the MPH-induced selective learning bias parameter $\kappa_{MPH\text{-}selective}$ (M5b) only slightly increased model fit. Conversely, the MPH-induced diffuse learning bias parameter $\kappa_{MPH\text{-}diffuse}$ (M5c) also strongly improved model fit relative to base model M4. This observation is in line with our earlier prediction that the MPH effects are predominantly driven by changes in the proportion of *incorrect* Go responses. Confirming the model comparison results, the MPH modulation of Pavlovian bias and diffuse learning parameters both covaried with Listening Span (π_{MPH}: R=.25, p=.013; $\kappa_{MPH\text{-}diffuse}$: R=.28, p=.006), while the MPH selective learning bias did not ($\kappa_{MPH\text{-}selective}$: R=-.01, p=.9). In other words, $\kappa_{MPH\text{-}selective}$ did not explain our effect of interest and improved model fit relatively weakly.

To assess whether π_{MPH} and $\kappa_{MPH\text{-}diffuse}$ explained unique Listening Span-dependent effects of MPH (i.e. whether there was evidence for both of these effects), we constructed a composite model (M6) containing both effects. Model comparison showed that indeed this composite model explained the data best (*Figure 5*). In this model, both parameters again significantly varied proportional to Listening Span (π_{MPH}: R=.24, p=.020; $\kappa_{MPH\text{-}diffuse}$: R=.22, p=.032; Figure 5).

Taken together, these modelling results attribute the MPH-induced motivational bias partly to an altered Pavlovian response bias (π_{MPH}), and partly to a reward-driven diffusion of credit during instrumental learning ($\kappa_{MPH\text{-}diffuse}$). In other words, MPH i) alters the impact of *cue* valence on action, which is present and persists from the first trial onward, and ii) alters the impact of rewarding *outcomes* on the learning of actions, which fully depends on and evolves with experience. Following a reward, the effect of $\kappa_{MPH\text{-}diffuse}$ is to increase the value of incorrect Go responses in addition to the correct Go response.

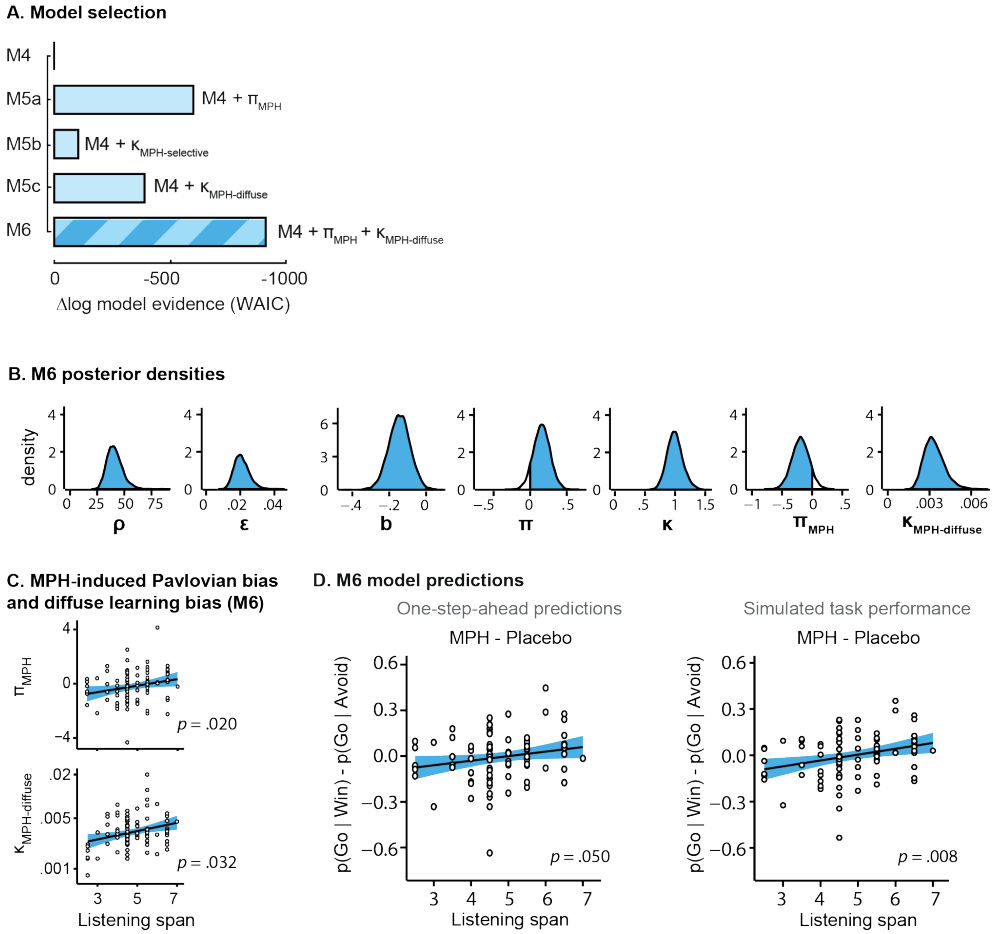


Figure 5. **Model evidence and parameter inference of extended MPH models.**
(**a**) Model evidence (WAIC) relative to winning base model M4. We tested whether MPH alters the strength of the Pavlovian response bias (π_{MPH}; M5a), the instrumental learning bias ($\kappa_{MPH\text{-}selective}$; M5b), or has a diffuse effect on the learning bias ($\kappa_{MPH\text{-}diffuse}$; M5c; Figure 1C). Model selection favoured the composite model M6, including the π_{MPH} and $\kappa_{MPH\text{-}diffuse}$ parameters. (**b**) Posterior densities of the top-level parameters of M6. (**c**) Subject-level estimates of MPH-induced Pavlovian bias parameter (upper) and the MPH-induced diffuse learning bias parameter (lower; logistic scale) correlated significantly with Listening Span. (**d**) One-step-ahead model predictions and posterior predictive model simulations of M6 using subject-level parameter estimates. The model predictions and simulations echo the observed data, i.e. that the motivational bias correlates positively with working memory span (Figure 4B), confirming the winning model M6 captures the MPH-induced increase in Go responses to Win vs. Avoid cues.
Figure supplements:

- *Figure 5 – figure supplement 1 (p.75).* Illustration of the behavioural effects of MPH related to the Pavlovian bias and diffuse learning bias parameters.
- *Figure 5 – figure supplement 2 (p.76).* M6 subject-level parameters in model space (i.e. untransformed).

	Extended MPH models			
	M5a	**M5b**	**M5c**	**M6**
WAIC	66383	66883	66595	66069
ρ	31.2 [14.7 53.6]	31.6 [15.6 57.0]	55.8 [19.6 104.8]	51.9 [20.6 98.7]
ε_0	0.022 [0.010 0.067]	0.021 [0.011 0.061]	0.011 [0.006 0.051]	0.012 [0.006 0.055]
b	−0.04 [−0.33 0.18]	−0.05 [−0.34]	−0.10 [−0.37 0.13]	−0.14 [−0.42 0.10]
π	0.27 [−0.50. 71]	0.15 [−0.28. 70]	0.05 [−0.46. 61]	0.27 [−0.47. 74]
π (MPH)	0.20 [−0.38. 71]			−0.05 [−0.70. 50]
$\varepsilon_{\text{rewarded Go}}$	0.037 [.017. 116]	0.030 [.018. 103]	0.018 [.009. 082]	0.019 [.009. 085]
$\varepsilon_{\text{rewarded Go}}$ **(MPH)**		0.031 [.016. 104]		
$\varepsilon_{\text{punished NoGo}}$	0.009 [.004. 030]	0.009 [.003. 021]	0.004 [.002. 013]	0.005 [.002. 017]
$\varepsilon_{\text{punished NoGo}}$ **(MPH)**	0.009 [.004. 030]	0.008 [.002. 021]	0.004 [.002. 013]	0.005 [.002. 017]
$\varepsilon_{\text{diffuse}}$ **(MPH)**			0.002 [.002. 004]	0.003 [.002. 004]

Table 2. **MPH models.** Median [25–75 percentile] of subject-level parameter estimates in model space. Absolute WAIC is reported as the estimate of model evidence, where a smaller WAIC indicates higher evidence. Biased instrumental learning rate for rewarded Go and punished NoGo responses as computed by $\varepsilon_0 \pm \kappa$ under placebo and by $\varepsilon_0 \pm (\kappa + \kappa_{MPH})$ under MPH. (MPH) indicates the value of that parameter under MPH.

Finally, we tested whether our best fitting model was sufficient to reproduce the key features of the data. This is important because model selection only provides relative, but not absolute evidence for the winning model (e.g., Nassar and Frank, 2016). We used two approaches to compute the post hoc absolute model fit, namely data simulation and "one-step-ahead" model predictions. In the simulation method, the first choice is simulated based on the initial values; the corresponding outcome used for learning; the next choice is simulated based on the updated, learned values; and so on. Thus, this simulation method ignores any subject-specific sequential/history effects to determine the current choice probability. Therefore, this can result in choice/outcome sequences that diverge completely from the subjects' actual experiences. Violating the subject-specific choice and outcome history will change the learning effects, making this method less robust in generating the exact learning effects compared to experience-*in*dependent effects. We therefore included a second absolute model fit method that does take into account the subjects' choice and outcome histories: the post-hoc absolute fit method (also known as 'one-step-ahead prediction'; Pederson et al., 2016; Steingroever and Wagenmakers, 2014). Here, the initial choice probabilities are determined based on the initial values. For each subsequent trial, the choice probabilities are determined based on the learned values

using the actual (subject's) choices and outcomes on all preceding trials. We used both methods as the strongest test providing converging evidence that the models could capture the observed results.

Using both absolute model fit methods, we simulated choices for each individual, using model M6 with each individual's parameter estimates. Both methods confirmed that M6 can capture the observed effects, replicating the Listening Span dependent effect of MPH on choice, where MPH increased Go responses to Win vs. Avoid cues more in higher span subjects (simulations: R=.27, p=.008; one-step-ahead: R=.20, p=.050; Figure 5). These simulations echo the results reported above, demonstrating the MPH-induced Pavlovian bias parameter π_{MPH} and diffuse learning bias $\kappa_{MPH\text{-}diffuse}$ are sufficient to both explain and predict the span-dependent MPH-induced increase in Go responses to Win vs. Avoid cues. Figure 5 and accompanying Figure supplements illustrate the model predictions and parameter estimates.

Discussion

Motivational biases of behaviour are well established: Reward biases towards action, punishment towards inaction. In this study, we had two goals. First, we aimed to assess whether these motivational biases arise from biases in instrumental learning in addition to Pavlovian response biases. Second, given the strong link between catecholamine transmission and motivated action, we aimed to assess effect of catecholaminergic manipulation on these biases. To this end, a large sample of participants (N=106) performed a novel motivational Go/NoGo learning task twice, once under a catecholamine challenge (methylphenidate - MPH) and once on placebo. Based on previous literature of dopaminergic drug effects (Cools & D'Esposito, 2011, and Frank & Fossella, 2011 for reviews), we hypothesized that MPH effects on motivated action would covary with measures scaling with baseline dopamine function, namely working memory span (Cools et al., 2008) and trait impulsivity (Buckholtz et al., 2010). Our findings are threefold: First, cue valence elicits behavioural activation in a Pavlovian manner, whereas outcome valence biases the learning of action vs. inhibition (Figure1A,B). Second, MPH modulates Pavlovian biasing, while also altering the reward-driven diffusion of credit assignment during instrumental learning. Third, the direction of the effect of MPH covaries with individual differences in working memory span, but not trait impulsivity.

Dissociable effects of cue and outcome valence on behavioural activation and instrumental learning

Cue valence affected activation versus inhibition of behaviour, consistent with previous reports (Geurts et al., 2013; Guitart-Masip et al., 2012). Even though cue valence was orthogonal to what subjects *should* be doing, subjects made more Go responses when pursuing reward, and fewer Go responses when trying to avoid punishment. We and others have previously

suggested that this motivational asymmetry in behavioural activation entails Pavlovian control over instrumental behaviour (Cavanagh et al., 2013; Geurts et al., 2013; Huys et al., 2011). Here we challenge this initial idea, and argue that motivational valence may also bias instrumental learning. To disentangle the hypothesised contribution of a Pavlovian response bias from biased instrumental learning, we extended existing paradigms by incorporating multiple Go response options. For the cues requiring active responses, only one response option was considered correct, enabling us to disentangle general activation from specific action learning. For cues where subjects had to activate responding ('Go' cues), they increased both correct and incorrect Go responses when pursuing reward compared with when avoiding punishment. Thus, the increased activation towards reward was in part beneficial, and in part detrimental.

We used computational models to formalise our hypothesis regarding a dissociable contribution of Pavlovian activation and biased instrumental learning. We then fitted competing models to the subjects' choices, and compared the performance of all models. We demonstrate that cue valence shapes behavioural activation/inhibition in a Pavlovian manner, and additionally that outcome valence biases instrumental learning of activation/inhibition: reward enhances the learning of specific active actions, and punishment suppresses the unlearning of inactions. In short, we are quicker to believe that an action led to a reward, but reluctant to attribute a punishment to having held back.

Current views of striatal dopamine function (Collins and Frank, 2015b, 2014, Frank, 2006, 2005; Lloyd and Dayan, 2016) suggest that the striatal architecture is well suited to implement the Pavlovian asymmetry in behavioural activation. Appetitive (aversive) conditioned cues elicit peaks (dips) in mesolimbic dopamine release in the striatum (Cohen et al., 2012; Day et al., 2007; Matsumoto and Hikosaka, 2009; Tobler et al., 2005). Increased striatal dopamine levels activate the direct D1 ("Go") pathway (Hernandez-Lopez et al., 1997), which promotes behavioural activation (DeLong and Wichmann, 2007; Mink and Thach, 1991), whereas decreased striatal dopamine levels activate the indirect D2 ("NoGo") pathway (Hernandez-Lopez et al., 2000), promoting behavioural inhibition. In striatal dopamine models, increased dopamine biases action selection to be driven more by the potential rewards of alternative actions encoded in D1 neurons and less by the costs encoded in D2 neurons (Collins and Frank, 2014; see also recent optogenetic experiment supporting this notion; Zalocusky et al., 2016), but this can also be manifest in terms of Pavlovian biases. Taken together, the striatal (in)direct pathways provide a neural mechanism for implementing Pavlovian activation to appetitive vs. aversive cues.

In parallel with our behavioural findings, the same striatal pathways may also generate the asymmetry in action learning. Here, dopamine bursts elicited by reward prediction errors (Montague et al., 2004; Schultz et al., 1998, 1997) during the outcome, enhance long-term potentiation (LTP) of the corticostriatal synapses associated with the just-performed response (Frank et al., 2004). Importantly, enhancing LTP in the "Go" pathway should

promote learning of active responses, relative to learning the inhibition of actions. Recent experiments show temporally and spatially selective enhancement of corticostriatal spines given glutamatergic input (putatively representing the selected action) and followed closely in time by dopaminergic bursts (Yagishita et al., 2014). Thus, prolonged release of DA (e.g. after DAT blockade) might reduce this selectivity, and diffuse the specificity of credit assignment. Conversely, striatal dopamine dips following negative prediction errors can drive avoidance by promoting long-term depression (LTD) in the "Go" pathway and LTP in the "NoGo" pathway (Beeler et al., 2012; Frank, 2005; Shen et al., 2008). Indeed, transient optogenetic inhibition of DA induces behavioural avoidance of recently selected actions (Danjo et al., 2014; Hamid et al., 2015), an effect that depends on D2 receptors (Danjo et al., 2014). D2 neurons are excited in response to losses (Zalocusky et al., 2016); their activation during losses induces subsequent avoidance learning (Kravitz et al., 2012; Zalocusky et al., 2016), and their disruption prevents avoidance learning (Hikida et al., 2010). While LTP in the NoGo pathway would be beneficial for unlearning to perform actions, LTP in the NoGo pathway would be detrimental in case of unlearning to make NoGo responses (i.e. attributing a punishment to a NoGo response). To summarize, the dopamine peaks following positive reinforcement can enhance learning of actions by enhancing LTP in the striatal "Go" pathway. Conversely, the dopamine dips following negative outcomes can disrupt learning to initiate responses by increasing LTD in the "Go" pathway and LTP in the NoGo pathway.

Methylphenidate modulates Pavlovian activation and spreads credit assignment of rewarded actions

Blocking the reuptake of catecholamines with MPH altered the extent to which subjects were influenced by the cue and outcome valence. This effect of MPH was highly variable between individuals, and depended on working memory span. In high relative to low span subjects, MPH enhanced the influence of valence, such that subjects made even more active responses when pursuing reward and displayed more inhibition when avoiding punishment. This effect was driven particularly by changes in the proportion of incorrect Go responses that subjects made. Formal modelling showed that this effect was due to MPH affecting both generalized Pavlovian activation and a diffusion of credit assignment. Specifically, MPH induced a spread of credit assignment following rewarded active responses, rather than magnifying the selective instrumental learning bias.

We argue that both of these effects can be understood as reflecting prolonged catecholamine presence in the synaptic cleft with MPH. Blocking catecholamine reuptake with MPH extends the duration of dopamine presence in the synaptic cleft (Dreyer and Hounsgaard, 2012). This prolonged dopamine presence (i.e. reduced temporal specificity) would be less selective in potentiating the actions that were selected immediately prior to rewards (e.g. Yagishita et al., 2014). This would reduce credit assignment of specific active

actions, but still bias reinforcement of actions more generally (e.g. Collins and Frank, 2015; Syed et al., 2015). This account explains why MPH modulates the strength of the Pavlovian activation (which is inherently global) but not of the specific instrumental learning bias (which is inherently selective). Our results indeed provided evidence for this diffusing effect of MPH on the instrumental learning bias, such that reward potentiates actions globally. The data were best explained by a combination of this diffuse instrumental learning and Pavlovian response bias modulation. Thus, on the one hand MPH modulated the impact of the cue valence on behavioural activation, which surfaces already before any learning has taken place. On the other hand, MPH spread credit assignment following rewarded responses to all Go responses, which is an experience-dependent effect.

Our results are highly consistent with those predicted from current models of dopamine in the basal ganglia, suggesting that the effects of MPH are due to modulation of striatal dopamine. Of course, the present study does not allow us to exclude the possibility that (part of) the effects were mediated by extra-striatal, e.g. prefrontal regions (Spencer et al., 2015), or by the noradrenaline system (Arnsten and Dudley, 2005). Future studies are needed to investigate directly the site of the presently observed effects of MPH, e.g. with fMRI, and dopamine dependence and selectivity, e.g. with selective dopamine antagonists.

MPH effects predicted by individual differences in working memory span

Individuals vary strongly in the extent to which MPH increases extracellular dopamine (Volkow et al., 2002). We therefore anticipated that the effect of MPH would covary with measures relating to baseline dopamine function. We assessed whether MPH effects were predicted by (i) working memory span, given its known relation to dopamine synthesis capacity (Cools et al., 2008; Landau et al., 2009), and (ii) trait impulsivity, for its known relation to D2 (auto)receptor availability (Buckholtz et al., 2010; Kim et al., 2014; Lee et al., 2009; Reeves et al., 2012). MPH affected choice behaviour proportional to working memory span, but not trait impulsivity. Subjects with higher working memory span, linked to higher striatal synthesis capacity, showed a relative increase in both Pavlovian response bias and spread of credit assignment under MPH. This finding that transporter blockade has stronger effects in those individuals with putatively higher baseline dopamine is in line with the observation that MPH increases dopamine levels more in individuals with higher dopamine cell activity (van der Schaaf et al., 2013; Volkow et al., 2002). Indeed, baseline dopamine cell activity is a better predictor of effects of MPH than either D2 auto-receptor availability or DAT occupancy under MPH (Volkow et al., 2002). Together this may explain why the observed MPH effects covary with working memory span but not trait impulsivity.

The finding that drug effects depend on working memory is highly consistent with the hypothesis that they reflect modulation of striatal dopamine (c.f. Frank and Fossella, 2011). However, we need to be cautious in our interpretation. First, both striatal and prefrontal

dopamine are known to contribute to working memory performance (updating and maintenance respectively; e.g. Cools and D'Esposito, 2011). The Listening Span task does not dissociate between working memory updating and maintenance, and thus a contribution of modulation of prefrontal dopamine cannot be excluded. Another possibility raised by the finding that drug effects depend on span, is that they reflect modulation of working memory itself, rather than reflecting dependence on baseline dopamine synthesis capacity. However, we argue that this is unlikely, because there was no significant effect of baseline working memory on motivational bias under placebo conditions. Rather, this relationship was induced by MPH. For future studies, it would be of interest to also include other measures related to baseline dopamine levels, such as eyeblink rates. More broadly, further research is required to identify the optimal combination of the various proxy measures of individual variability in the dopamine system in order to account for the large inter-individual variability in dopaminergic drug response. This is one of the major aims of our ongoing work.

Across subjects, MPH increased subjective experiences of positive affect and alertness, and decreased calmness (Appendix 2). In contrast to the MPH-induced Pavlovian response bias and diffuse learning bias, these non-specific mood changes did not covary with working memory span. In other words, the MPH-induced mood changes are orthogonal to our effect of interest. Therefore, the MPH effect on Pavlovian activation and biased instrumental learning cannot be attributed to MPH-induced changes in mood.

Conclusion

This study elucidates two distinct mechanisms by which motivational valence can bias behaviour. Cue valence promotes activation/inhibition in a Pavlovian manner, whereas outcome valence affects action/inhibition learning. Blocking the reuptake of catecholamines with methylphenidate altered the Pavlovian response bias, and had a diffuse, rather than selective, effect on biased learning. The effect of methylphenidate on the Pavlovian bias and biased learning was predicted by working memory span, such that methylphenidate enhanced Pavlovian activation and biased learning proportional to working memory span. These results help bridge the study of motivational biasing of action and instrumental learning, and help refine current models of catecholamines in motivated action.

The present observations suggest that we need to add a new dimension to the suggested dichotomy of the role of dopamine in learning versus performance. Our study brings together two literatures that emphasise the role of (midbrain) dopamine in reward (prediction-error) based learning on the one hand (Collins and Frank, 2014; Frank et al., 2004; Schultz et al., 1997), and motivation-driven performance and behavioural activation on the other (Beierholm et al., 2013; Berridge, 2007; Robbins and Everitt, 2007; Shiner et al., 2012; Smittenaar et al., 2012). Our results suggest that these two interact, resulting in biased learning of action-reward and inaction-punishment links, putatively via the same striatal mechanism that drive motivational

Pavlovian response biases. Like motivational response tendencies, such biased learning would allow us to optimally profit from stable environmental statistics, as this instrumental learning bias supports rapid learning of likely action-outcome associations (e.g. that an action caused a reward), while avoiding learning unlikely, spurious, associations (e.g. that inhibition caused a punishment).

Materials and Methods

General procedure and pharmacological manipulation

The study consisted of two test sessions with an interval of one week to two months. The first test day started with informed consent, followed by a medical screening. Participation was discontinued if subjects met any of the exclusion criteria (Appendix 1). On both test days, subjects first completed baseline measures. Next subjects received a capsule containing either 20 mg MPH (Ritalin®, Novartis) or placebo, in a double-blind, placebo-controlled, cross-over design. MPH blocks the dopamine and noradrenaline transporters, thereby diminishing the reuptake of catecholamines. When administered orally, MPH has a maximal plasma concentration after 2 hours and a plasma half-life of 2-3 hours (Kimko et al., 1999). After an interval of 50 minutes, subjects started with the task battery containing the motivational Go/NoGo learning task. See Appendix 2 for an overview of the task battery. On average the motivational Go/NoGo learning task was performed 2 hours after capsule intake, well within the peak of plasma concentration. Both test days lasted approximately 4.5 hours, which subjects started at the same time (maximum difference of 45 minutes). Blood pressure, mood and potential medical symptoms were monitored three times each day: before capsule intake, upon start of the task battery and after finishing the task battery. Subjects were told to abstain from alcohol and recreational drugs 24h prior to testing and from smoking and drinking coffee on the days of testing. Subjects completed self-report questionnaires at home between (but not on) test days. Upon completion of the study, subjects received a monetary reimbursement or study credits for participation. The study was in line with the local ethical guidelines approved by the local ethics committee (CMO / METC Arnhem Nijmegen: protocol NL47166.091.13), pre-registered (trial register NTR4653, http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=4653), and in accordance with the Helsinki Declaration of 1975. Baseline measures, self-report questionnaires, mood- and medical symptom-ratings are reported in Appendix 2.

Subjects

As individual differences were a main focus of the study, we collected a large sample of 106 native Dutch volunteers (aged 18 – 28 years, mean (SD)=21.5 (2.3); 53 women; 84 right-handed; sample size calculation reported in CMO protocol NL47166.091.13). Four subjects dropped

out after the first test day (due to too much delay between test days, loss of motivation, nausea, and mild arrhythmia). Two subjects dissolved the capsules before swallowing and are discarded because of uncertainty in the pharmacodynamics. One subject did not sufficiently engage in the task (only 13/2% Go responses on day 1/2) and was discarded as well. We repeated the analyses with these subjects included to confirm that this did not alter the conclusions (Appendix 3). Of the resulting 99 subjects, 48 subjects received MPH on the first day. Exclusion criteria comprised a history of psychiatric, neurological or endocrine disorders. Appendix 1 presents a complete overview of the exclusion criteria.

Motivational Go/NoGo learning task

Each trial started with the on-screen presentation of a cue (Figure 2A). During cue presentation subjects could decide to press a button (*Go response*) or not (*NoGo response*). Subjects could either press the left (*Go-left*) or right (*Go-right*) button on a button box. Subjects received feedback based on their response.

Each cue had a red or green edge. Cues with a red edge (*Avoid cues*) were followed by neutral feedback or punishment. Cues with a green edge (*Win cues*) were followed by reward or neutral feedback. Subjects were informed about these contingencies. Note that the explicit cue valence is in contrast to previous studies where subjects needed to learn the cue valence during the task (e.g. Cavanagh et al., 2013; Guitart-Masip et al., 2012). The rationale of explicit cue valence was to directly observe effects of cue valence on choice and minimize individual differences in learning the cue valence. Punishment consisted of the display of the red text '-100', accompanied by a low buzz, reward of the green text '+100' together with a flourish sound, and the neutral feedback of the grey text '000' together with a short beep. All cues had unique shapes and colours well distinguishable from the red and green edge. Cue colour and shape were randomized over cue types. Two separate stimulus sets were used for the two test days to prevent transfer effects, and set order was counterbalanced across subjects.

For each cue, there was one correct response (Go-left, Go-right or NoGo; Figure 2C), which subjects had to learn by trial and error. Feedback validity was 80%, i.e. correct (incorrect) responses were followed by the desirable outcome 80%(20%) of the time (Figure 2D). There were 8 cues in total (Figure 2B). The number of Go and NoGo cues was kept equal to prevent reinforcing an overall Go bias.

The order of cue presentation was pseudorandom, as cues could be repeated once at most. Each cue was presented 40 times. The task lasted approximately 30 minutes, including instructions and a self-paced break halfway. The instructions were presented on screen. Subjects were informed about the probabilistic nature of the feedback and that each cue had one optimal response. At the end of the task the total number of points won or lost was displayed on screen and subjects were informed beforehand that these points would be converted to a monetary bonus at the end of the study (mean=EUR2.90, SD=1.49).

Listening span test

Working memory span was assessed with the Listening Span Test (Daneman and Carpenter, 1980; Salthouse and Babcock, 1991), which was also used in two FMT PET studies showing positive correlations with striatal dopamine synthesis capacity (Cools et al., 2008; Landau et al., 2009). Subjects completed the Listening Span Test on day 2 prior to capsule intake. The Listening Span Test consists of sets of pre-recorded sentences, increasing from 2 to 7 sentences. Subjects are presented with the sentences, and required to simultaneously answer written verification questions regarding the content of each sentence. At the end of each set, subjects recalled the final word of each sentence in the order of presentation. The Listening Span reflects the set size of which the subject correctly recalled the final words on at least two out of three trials. Listening span increased with half a point, when only one trial of the next level was correct.

Barratt impulsiveness scale

Trait impulsivity was assessed with the Barratt Impulsiveness Scale (BIS-11) (Patton et al., 1995). The BIS-11 is a self-report questionnaire, consisting of 30 questions tapping in common (non) impulsive behaviours and preferences. The BIS-11 total impulsivity scores reflect the tendency towards impulsivity. Subjects completed the questionnaire at home between test days.

Statistical analyses

To assess the influence of motivational valence on behavioural activation, we first analysed Go vs. NoGo responses (irrespective of Go-left vs. Go-right). Second we tested whether effects on Go responses were explained by correct or incorrect Go responses. We were specifically interested how MPH altered Go/NoGo responding to Win vs. Avoid cues as a function of Listening Span and Impulsivity.

To account for both between and within subject variability, choice data were analysed with logistic mixed-level models using the lme4 package in R (Bates et al., 2014; R Developement Core Team, 2015). Reflecting our objectives, the mixed models included the within subject factors Drug (MPH vs. placebo), Valence (Win vs. Avoid cue), and Required Action (Go vs. NoGo), and the between subject factors Listening Span and Impulsivity. The analysis of correct and incorrect Go responses included only the Go cues; hence this analysis did not include the factor Required Action. Models included all main effects and interactions, except for the interactions between Listening Span and Impulsivity. All models contained a full random effects structure (Barr, 2013; Barr et al., 2013). We performed control analyses using a model comparison approach, where we tested whether the following factors improved model fit: Drug Order, Testing Day, Gender, and NLV (a measure for verbal intelligence). For completeness, we analysed reaction times (RTs) as a measure of behavioural vigour (Appendix 4).

Computational modelling – Pavlovian response bias and instrumental learning bias

In all models, action weights (*w*) are estimated for each response option (*a*) for all trials (*t*) per cue (*s*). Based on these action weights choice probabilities are computed using a softmax function, as follows:

$$p(a_t|s_t) = \left[\frac{\exp(w(a_t, s_t))}{\sum_{a'} \exp(w(a', s_t))}\right] \quad \text{Eq. 1}$$

In the simplest model (M1) the action weights are fully determined by the learned action values (*Q*-values). To compute the action values, we used standard delta-rule learning with two free parameters; a learning rate (ε) scaling the update term, and feedback sensitivity (*ρ*) scaling the outcome value (comparable to the softmax temperature).

$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \varepsilon\left(\rho r_t - Q_{t-1}(a_t, s_t)\right) \quad \text{Eq. 2}$$

Here outcomes are reflected by *r*, where r ε{-1,0,1}. In the current paradigm cue valence is instructed, by means of the green and red cue edges. Therefore, the initial expected outcome is 0.5 for Win cues and -0.5 for Avoid cues. Initial *Q*-values (Q_0) are set accordingly to *ρ**0.5 for Win cues and *ρ**-0.5 for Avoid cues.

In M2 a go bias parameter (*b*) is added to the action weights of Go responses. We then explored the influence of Pavlovian biases that modulate Go responding according to predicted reward value. Pavlovian values (*V*) contribute to the action weights in M3a, increasing (decreasing) the weight of Go responses for positive (negative) Pavlovian values respectively.

$$w(a_t, s_t) = \begin{cases} Q(a_t, s_t) + \pi V(s) + b & if\ a = Go \\ Q(a_t, s_t) & else \end{cases} \quad \text{Eq. 3}$$

Here the weight of the Pavlovian values is determined by the parameter π. Pavlovian values are fixed at 0.5 for Win cues and at -0.5 for Avoid cues, again because cue valence is instructed.

In M3b we included the instrumental learning bias parameter (κ) instead of the Pavlovian bias, to assess whether the motivational bias can be explained in terms of enhanced learning of Go following a reward, and disrupted learning from punishment following NoGo.

$$\varepsilon = \begin{cases} \varepsilon_0 + \kappa & if\ r_t = 1\ \&\ a = go \\ \varepsilon_0 - \kappa & if\ r_t = -1\ \&\ a = nogo \\ \varepsilon_0 & else \end{cases} \quad \text{Eq. 4}$$

In model M4, we included both the Pavlovian bias parameter and the instrumental learning bias parameter.

We used a sampling method for hierarchical Bayesian estimation of group-level and subject-level parameters. The group-level parameters (*X*) serve as priors for the individual-level parameters (*x*), such that $x \sim N(X,\sigma)$. The hyperpriors for σ are specified by a half-Cauchy (Gelman,

2006) with a scale of 2. The hyperpriors for X are centered around 0 (with the exception of X_ρ) and weakly informative: $X_\rho \sim \mathcal{N}(2,3)$, $X_{\varepsilon,\kappa} \sim \mathcal{N}(0,2)$, $X_{b,\pi} \sim \mathcal{N}(0,3)$. All parameters are unconstrained, with the exception of ρ (positivity constraint; exponential transform) and ε ([0 1] constraint; inverse logit transform). To ensure that the effect of κ on ε (Eq.4) was symmetrical in model space (i.e. after sigmoid transformation to ensure [0 1] constraint), ε was computed as:

$$\varepsilon = \begin{cases} \varepsilon_0 = \text{inv.logit}(\varepsilon) \\ \varepsilon_{\text{punished NoGo}} = \text{inv.logit}(\varepsilon - \kappa) \\ \varepsilon_{\text{rewarded Go}} = \varepsilon_0 + (\varepsilon_0 - \varepsilon_{\text{punished NoGo}}) \end{cases} \quad \text{Eq. 5}$$

Model estimations were performed using Stan software in R (RStan) (Stan Development Team, 2016). Stan provides full Bayesian inference with Markov chain Monte Carlo (MCMC) sampling methods (Metropolis et al., 1953). The number of Markov chains was set at 4, with 200 burn-in iterations and 1000 post burn-in iterations per chains (4000 total). Model convergence was considered when the potential scale reduction factor $\hat{R} < 1.1$ for all parameters (Gelman and Rubin, 1992). In case model convergence was not reached, both (post) burn-in samples were increased to 1500. Not all models reached convergence at this point. Therefore, we repeated model estimation while excluding the subjects (N=5) for whom initially $\hat{R} > 1.1$ in any one of the models, resulting in model convergence for all models. We report model evidence including all subjects in Appendix 5, showing that model selection and parameter inference remains the same when excluding these subjects. Model comparison was evaluated using the Watanabe-Akaike Information Criteria (WAIC) (Watanabe, 2010). WAIC is an estimate of the likelihood of the data given the model parameters, penalized for the effective number of parameters to adjust for overfitting. Lower (i.e. more negative) WAIC values indicate better model fit. As WAIC is reported on the deviance scale (Gelman et al., 2014), a difference in WAIC value of 2-6 is considered positive evidence, 6-10 strong evidence, and >10 very strong evidence (Kass and Raftery, 1995).

Computational modelling – Effects of methylphenidate

Having established the mechanisms by which motivational valence may affect instrumental learning and activation, we extended the winning model to test which of these mechanisms are affected by MPH, putatively driven by a prolonged striatal presence of catecholamines (dopamine) following reward, due to reuptake inhibition by MPH.

In M5 we tested whether MPH altered the Pavlovian response bias. This model includes a parameter allowing for an MPH-induced change in the Pavlovian weight (π_{MPH}):

$$\pi = \begin{cases} \pi_0 & if\ placebo \\ \pi_0 + \pi_{MPH} & if\ MPH \end{cases} \quad \text{Eq. 6}$$

Next, we tested two mechanisms by which MPH might alter the bias in instrumental learning (κ). In M5b we tested whether MPH simply enhanced or reduced the learning bias parameter, estimating an additive effect of $\kappa_{MPH\text{-}selective}$:

$$\kappa = \begin{cases} \kappa_0 & if\ placebo \\ \kappa_0 + \kappa_{\mathrm{MPH-selective}} & if\ MPH \end{cases} \quad \text{Eq. 7}$$

Alternatively, the prolonged presence of catecholamines following reward under MPH could induce a more diffuse credit assignment, rather than a selective learning bias effect. To test this hypothesis, in M5c we included a MPH-induced learning bias parameter ($\kappa_{\text{MPH-diffuse}}$), which was used to update *both* Go responses, on all trials where any active Go response was followed by reward, in addition to the regular learning update for the chosen Go response:

$$if\ MPH, r_t = 1, \&\ a_{chosen} = Go:$$

$$\begin{aligned} Q_t\left(a_{chosenGo,t}, s_t\right) &= Q_{t-1}\left(a_{chosenGo,t}, s_t\right) + (\varepsilon + \kappa_0 + \kappa_{\mathrm{MPH-diffuse}}) \cdot PE \\ Q_t\left(a_{unchosenGo,t}, s_t\right) &= Q_{t-1}\left(a_{unchosenGo,t}, s_t\right) + \kappa_{\mathrm{MPH-diffuse}} \cdot PE \end{aligned} \quad \text{Eq. 8}$$

Where *PE* is the prediction error following the rewarded Go response: $PE = \rho r_t - Q_{t-1}(a_t, s_t)$. Thus where $\kappa_{\text{MPH-selective}}$ enhances the learning of the selected Go response after reward, $\kappa_{\text{MPH-diffuse}}$ induces learning of all Go responses when a Go response elicited reward.

To test whether MPH affected both the Pavlovian response bias and instrumental learning bias, M6 include π_{MPH} parameter as well as the winning model of the two learning bias mechanisms (M5c - $\kappa_{\text{MPH-diffuse}}$). For completeness, we report the composite model including the parameters π_{MPH} and $\kappa_{\text{MPH-selective}}$ in Appendix 5. The hyperpriors are again centered around 0 and weakly informative: $X_{\kappa\text{mph}} \sim N(0,2)$ and $X_{\pi\text{mph}} \sim N(0,3)$, where only $X_{\kappa\text{mph-diffuse}}$ is constrained ([0 1] constraint; inverse logit transform).

Having established the winning model, we used two absolute model fit approaches to confirm that the winning model captures the effects of interest; the post-hoc absolute-fit approach (also called *one-step-ahead prediction*) and posterior predictive model simulation approach (Steingroever and Wagenmakers, 2014). The posterior predictive model simulations simply 'play' the task, using the estimated parameters. This approach, however, ignores sequential/history effects of actually observed choices and outcomes. The 'one-step-ahead' prediction fits parameters to trials t_1 - t_{n-1}, and then predicts the choice on trial t_n. Taking these sequential effects into account is particularly important to assess effects of the parameters that estimate the effect of previous choice/outcome combinations, i.e. the learning rate parameters, relative to the constant parameters like the Pavlovian and go biases. For both the one-step-ahead predictions and model simulations, we computed action probabilities for all subjects on all trials using the sampled combinations of all individual-level parameter estimates. For the one-step-ahead predictions the observed choices and outcomes were used to update the action probabilities. For the model simulations choices were simulated depending on the computed action probabilities, and outcomes were determined according to the ground-truth outcome probabilities (i.e. a correct response would lead to the desired outcome 80% of the time). Subsequently, outcomes corresponding to the simulated choices were used to update the action probabilities. The one-step-ahead prediction and simulations were repeated for all sampled parameter combinations

(4000 times), and action probabilities were averaged over repetitions. Averaging over repetitions also minimizes effects of randomness due to the stochastic nature of the choice simulation.

Figure supplements

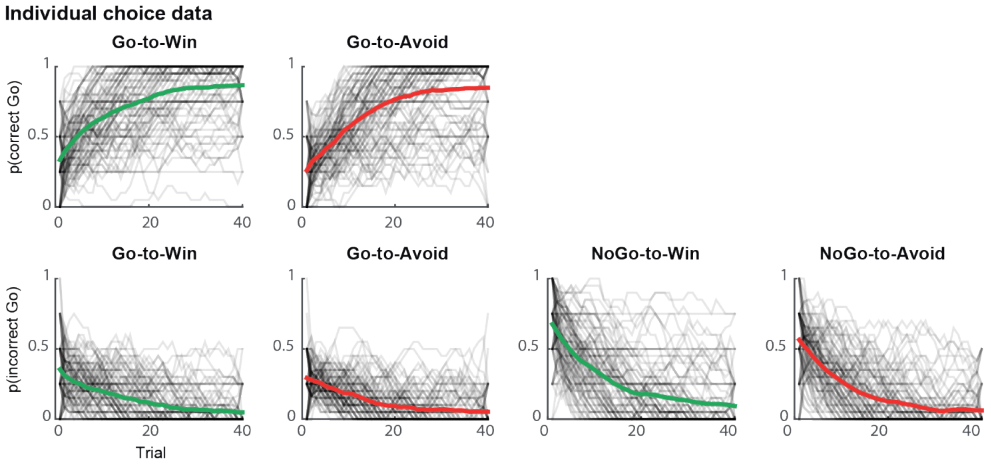


Figure 2 – figure supplement 1. **Individual traces (black lines) and group average (coloured lines) of correct and incorrect Go responses using a sliding average of 5 trials.**
Traces are averaged within cue types and over sessions. Individual traces are semi-transparent, so that darker areas reflect more overlaying subjects. Across trials, subjects increased correct Go responses (top) and decreased incorrect Go responses (bottom). Subjects performed at ceiling level more rapidly for the Go-to-Win cues (top-left) than Go-to-Avoid cues (top-right).

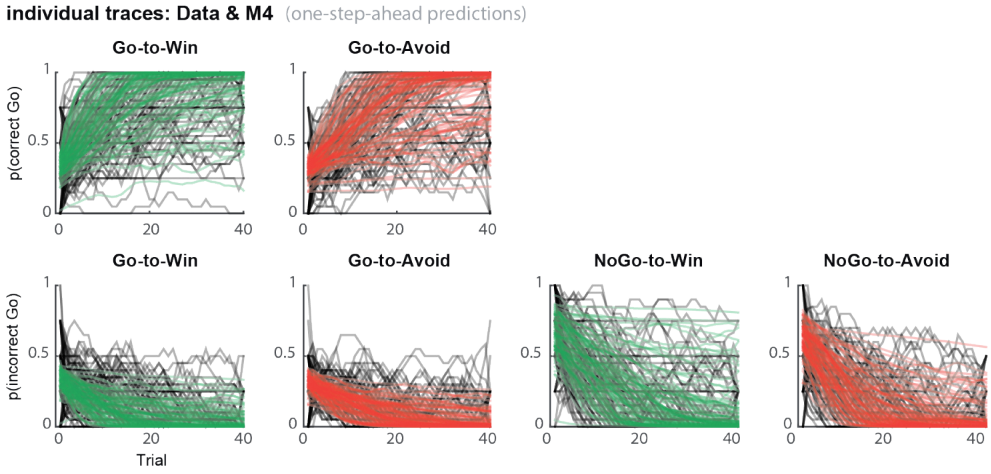


Figure 3 – figure supplement 1. **Subject traces of model M4 (green/red) overlaid on observed behaviour (black).**
M4 one-step-ahead predictions capture the individual variability in task performance.

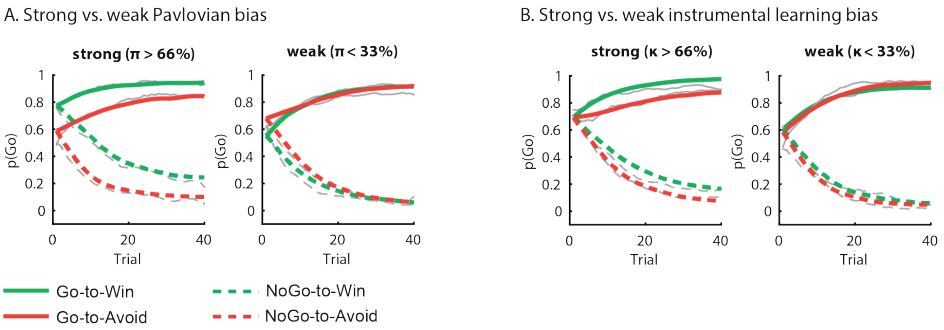


Figure 3 – figure supplement 2. **Illustration of the behavioural effects associated with the Pavlovian bias and instrumental learning bias parameters.**
Model M4 one-step-ahead predictions (coloured) overlaid on real data (grey) for the subjects with the upper versus lower tertile of parameter estimates. (**a**) Effects of Pavlovian bias (π). A strong Pavlovian bias (top 33% of π estimates) predicts higher Go responding for the Win than Avoid cues from the first trial onward, vice versa for a weak Pavlovian bias (33% lowest π estimates). (**b**) Effects of instrumental learning bias (κ). A strong instrumental learning bias (33% highest κ estimates) predicts steeper Go-to-Win learning and shallower Go-to-Avoid learning, vice versa for a weak instrumental learning bias (33% lowest κ estimates). See also Figure 3B for the temporal dynamics of the parameter-behaviour correlations.

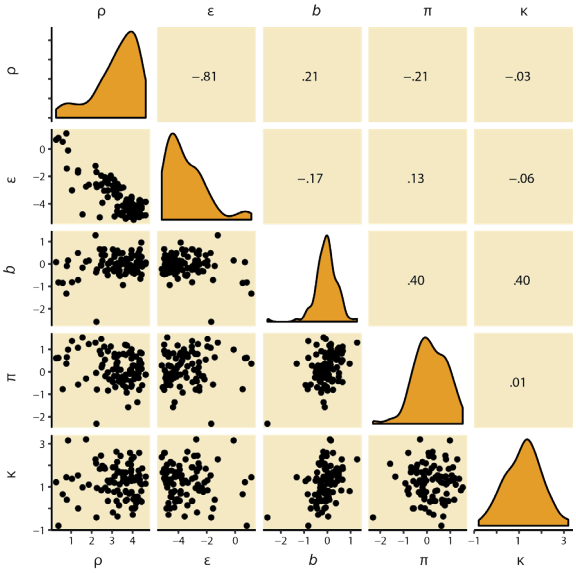


Figure 3 – figure supplement 3. **M4 subject-level parameters in model space (i.e. untransformed).**
The diagonal panels contain the posterior densities for the subject-level parameter means. The off-diagonal panels show the correlation over subjects in mean parameter estimates. Importantly, the two key parameters, Pavlovian bias (π) and instrumental learning bias (κ) are not correlated to any of the other parameters. We do note that the feedback sensitivity parameter (ρ) is anti-correlated with the learning rate (ε), such that the impact of high feedback sensitivity estimates is restricted by low learning rates. This correlation is not problematic, because independent estimation of learning rate and feedback sensitivity is no direct interest to the questions we ask.

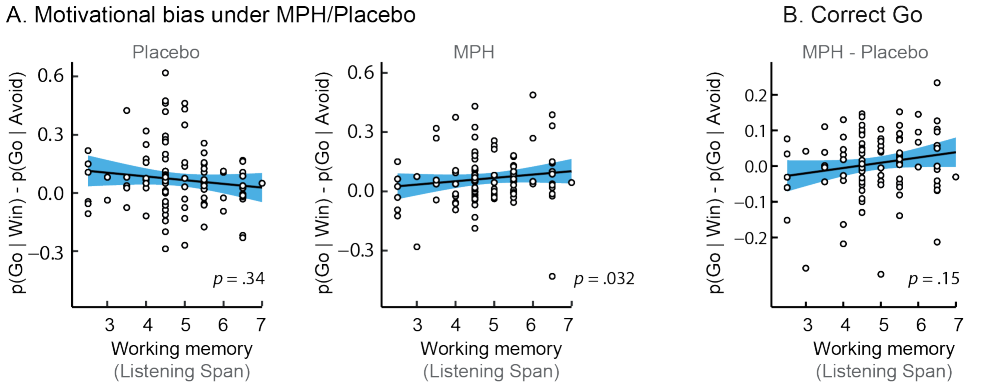


Figure 4 – figure supplement 1. **Simple effects of MPH-induced changes in motivational bias.**
(**a**) The span-dependent motivational bias emerged under MPH (right; p=0.032), and was not significant under placebo (left; p=0.34). (**b**) MPH did not significantly alter the motivational bias proportional to working memory span for *correct* Go responses (*correct* Go: p=0.15).

M6 one-step-ahead predictions | parameter estimates

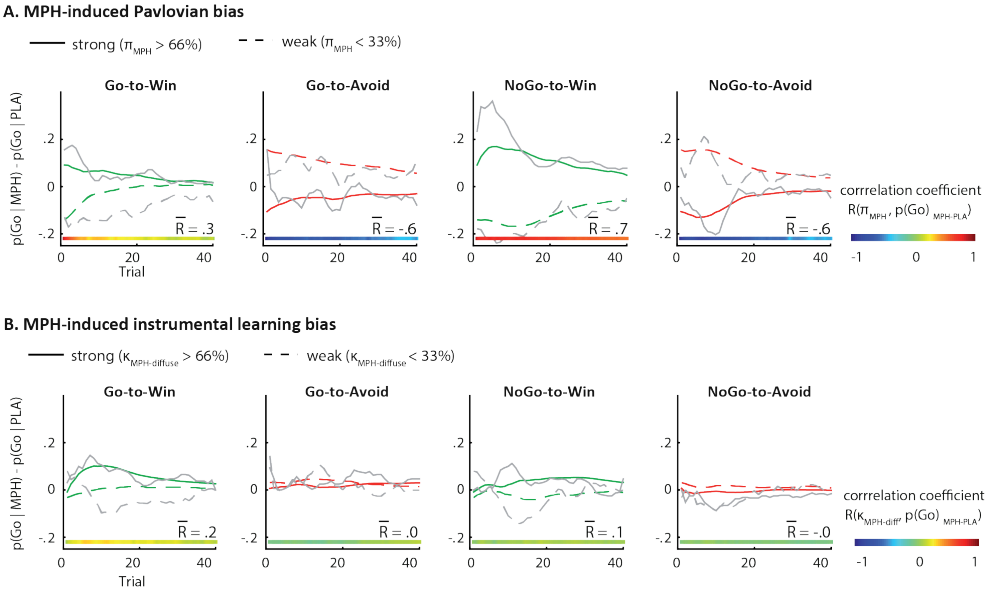


Figure 5 – figure supplement 1. **Illustration of the behavioural effects of MPH related to the Pavlovian bias and diffuse learning bias parameters.**
Model M6 one-step-ahead predictions (coloured) overlaid on real data (grey) for the subjects with the 33% strongest vs. weakest parameter estimates. The coloured bars at the bottom indicate the trial-by-trial correlation across all subjects, of the parameter estimate with the effect of MPH on Go responding per cue. The R value indicates the average correlation. (**a**) The effect of MPH on Pavlovian bias (π_{MPH}). Strong π_{MPH}estimates predict that MPH increases the motivational bias (increased Go to Win cues and decreased Go to Avoid cues), and vice versa for weak π_{MPH} estimates. The influence of π_{MPH} is present from the first trial onward and decreases over time as indicated by the correlation coefficients. (**b**) Effect of MPH on diffuse learning bias ($\kappa_{MPH\text{-}diffuse}$). Strong $\kappa_{MPH\text{-}diffuse}$ estimates predict that MPH increases the motivational bias for Win cues specifically, whereas this effect is diminished for subjects with relatively weak $\kappa_{MPH\text{-}}$

$_{diffuse}$ estimates. The effect of $\kappa_{MPH\text{-}diffuse}$ is experience-dependent and evolves over time. These one-step-ahead predictions illustrate how each parameter results in an increased motivational bias under MPH, but with unique temporal dynamics, even though the parameter themselves are constant.

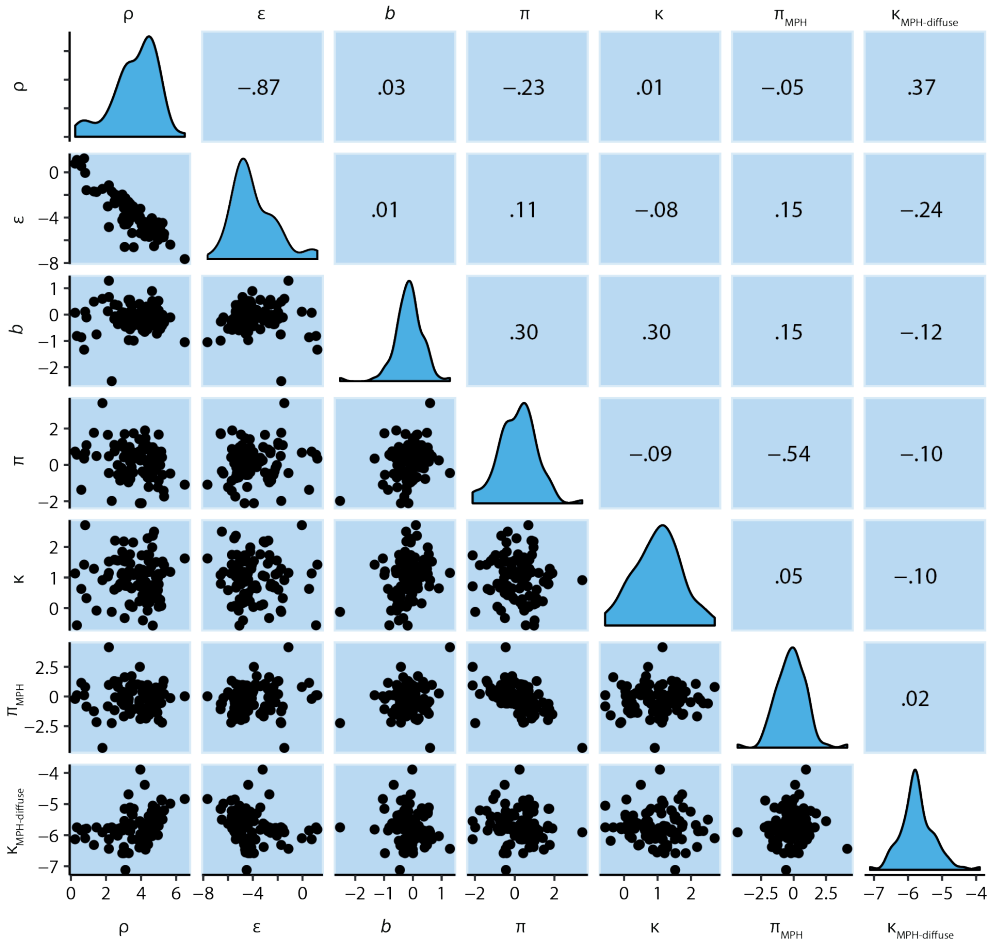


Figure 5 – figure supplement 2. **M6 subject-level parameters in model space (i.e. untransformed).** The diagonal panels contain the posterior densities for the subject-level parameter means. The off-diagonal panels contain the parameter correlations over subjects. Importantly, the parameters estimating the effects of MPH on Pavlovian bias (π_{MPH}) and diffuse learning bias ($\kappa_{MPH\text{-}diffuse}$) are not correlated to any of the other parameters.

Supplementary Files

Appendix 1 - Exclusion criteria

Exclusion criteria comprised a history of psychiatric, neurological or endocrine disorders. Further exclusion criteria were autonomic failure, hepatic, cardiac, obstructive respiratory, renal, cerebrovascular, metabolic, ocular or pulmonary disease, epilepsy, substance abuse,

suicidality, hyper/hypotension, diabetes, pregnancy/breastfeeding, lactose intolerance, abnormal hearing or (uncorrected) vision (e.g. colour blindness), irregular sleep/wake rhythm, regular use of corticosteroids, use of MAO inhibitor, anaesthetic, anti-depressant or anti-psychotic drugs within the week prior to the start of the study, use of psychotropic medication or recreational drugs/alcohol 24 hours before each test day, and first degree family members with schizophrenia, bipolar disorder, ventricular arrhythmia or sudden death. Inclusion age range was 18-45 years old.

Appendix 2 - Baseline measures and mood ratings

Prior to capsule intake, subjects completed a Dutch reading test (NLV, Schmand et al., 1991) as a proxy of verbal intelligence on day 1, and the Listening Span Test (Daneman and Carpenter, 1980; Salthouse and Babcock, 1991) on day 2. Subsequently subjects completed the Digit Span Test (forward and backward; Wechsler, 2008) and the training phase of a Pavlovian-Instrumental Transfer task (PIT, Geurts et al., 2013; Huys et al., 2011) of which data will be reported elsewhere. Between test days, subjects completed a number of self-report questionnaires. The group that received MPH on day 1 did not differ significantly on any of the baseline measures from the group that received placebo on day 1 ($p<0.05$). See Appendix 2—table 1 for an overview of the neuropsychological test scores and self-report questionnaires.

Mood ratings, heart rate and blood pressure were monitored for safety reasons three times during each test day, (i) before capsule intake, (ii) upon start task battery, and (iii) upon completion of the task battery. The mood ratings consisted of the Positive and Negative Affect Scale (PANAS, Watson et al., 1988) and the Bond and Lader Visual Analogues Scales (calmness, contentedness, alertness; Bond and Lader, 1974), as well as a medical Visual Analogues Scale. We assessed whether MPH affected mood and medical symptoms. For this control analysis we performed a repeated measures MANOVA using Pillai's trace with the within subject factors Time (baseline/start testing/end testing) and Drug (MPH/placebo), and dependent variables Positive Affect, Negative Affect, Calmness, Contentedness, Alertness, and Medical Symptoms. Significant effects were further explored with Bonferonni corrected repeated measures ANOVA, where alpha = 0.05/6 ≈ 0.008. Greenhouse-Geisser correction was applied when the assumption of sphericity was not met.

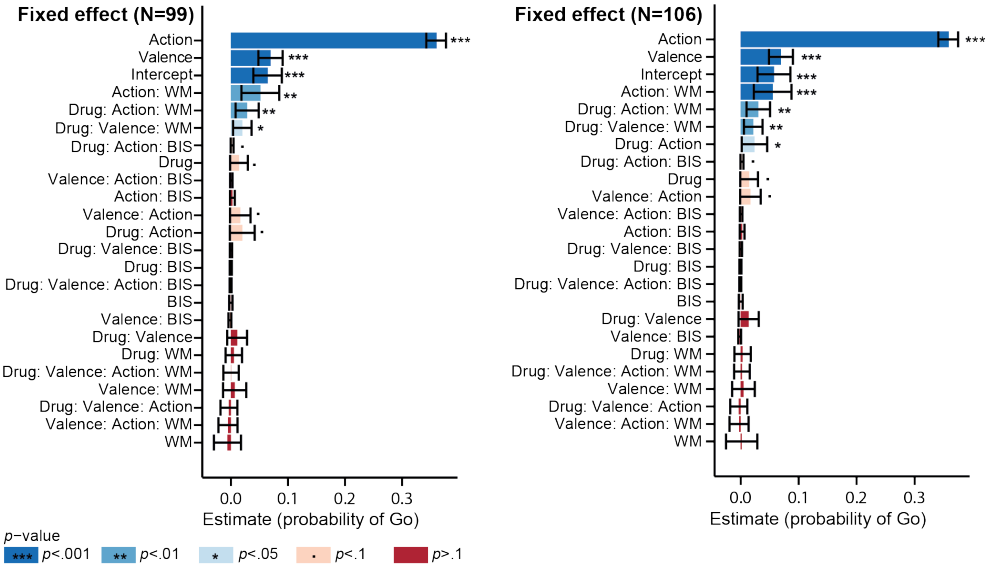
MPH affected these self-report ratings (Time x Drug: $V=0.38$, $F_{12,90}=4.7$, $p<0.001$), in the absence of baseline differences between the MPH and placebo groups ($V=0.07$, $F_{6,96}=1.1$, $p=0.359$). After capsule intake MPH increased Positive Affect ($F_{1,101}=17.5$, $p<0.001$), Alertness ($F_{1,101}=15.2$, $p<0.001$), and Medical Symptoms ($F_{1,101}=11.1$, $p=0.001$), and decreased Calmness ($F_{1,101}=8.6$, $p=0.004$), relative to placebo. We confirmed that the effects of MPH on the self-report ratings did not further interact with Listening Span and Impulsivity ($p>0.05$). Thus, the MPH-induced changes in mood and medical symptoms were orthogonal to the Listening Span dependent MPH effects we observed in the task.

		Group 1 (Placebo Day 1)	Group 2 (MPH Day 1)	
Neuropsychological tests	Listening span	5.0 (0.9)	4.6 (1.2)	*p*=0.16
	NLV	94.4 (7.6)	92.6 (7.6)	*p*=0.23
	Digit span – forward	17.2 (3.7)	16.2 (3.6)	*p*=0.16
	Digit Span - backward	14.7 (3.4)	13.9 (2.7)	*p*=0.22
Self-report questionnaires	Impulsivity (BIS-11)	63.5 (8.9)	60.2 (7.9)	*p*=0.052*
	Behavioural inhibition (BIS)	16.4 (3.7)	16.3 (3.5)	*p*=0.90
	Behavioural activation (BAS)	22.8 (3.9)	23.9 (4.0)	*p*=0.17
	Need for cognition (NCS)	64.5 (10.5)	62.2 (10.5)	*p*=0.26
	Social support (MSPSS)	71.1 (10.1)	69.3 (9.6)	*p*=0.35
	Social status (BSMSS)	49.8 (12.1)	45.9 (12.7)	*p*=0.11
	Social dominance (SADQ)	4.1 (0.9)	4.1 (0.8)	*p*=0.82
	Aggressive dominance (SADQ)	2.6 (0.6)	2.6 (0.6)	*p*=0.69
	Depressive symptoms (BDI-II)	3.5 (3.7)	3.6 (3.9)	*p*=0.97
	Anxiety symptoms (STAI)	32.4 (6.6)	32.4 (7.2)	*p*=1.0

Appendix 2 - table 1. **Mean(SD) scores for neuropsychological tests and self-report questionnaires for the group that received placebo and MPH on day 1.**
Significance levels for the between group differences are reported. Self-report questionnaires include the Barratt Impulsiveness Scale (BIS-11; Patton et al., 1995), the Behavioural Inhibition Scale/Behavioural Activation Scale (BISBAS; Carver and White, 1994), Need for Cognition Scale (NCS, Cacioppo et al., 1984), Multidimensional Scale of Perceived Social Support (MSPSS, Zimet et al., 1988), Barratt Simplified Measure of Social Status (BSMSS, Barratt, 2006), Sociable and Aggressive Dominance Questionnaire (SADQ, Kalma et al., 1993), Beck Depression Inventory II (BDI-II; Beck et al., 1996), Spielberger Trait Anxiety Inventory (STAI; Spielberger et al., 1983). *One subject had an outlying score on the BIS-11. Without outlier: *p*=0.09.

Appendix 3 - General performance and individual differences in drug effects on task performance

In the main manuscript, we report the results of 99 (out of 106) subjects. Four subjects did not complete both test days, two subjects dissolved the capsules before swallowing, and one subject did not sufficiently engage in the task (only 13/2% Go responses on day 1/2). We then repeated the analyses with these subjects included to confirm that this did not alter the conclusions (Appendix 3 - Figure 1).



Appendix 3 - figure 1. **Logistic mixed model estimates of the probability of Go responses to verify that exclusion of a subset of subjects (7) did not affect our inference.**
Left: N = 99; Right: N = 106. Fixed effect estimates and 95% confidence interval (CI) are plotted on probability scale. Effects are sorted by lower bound of the CI. The results including all 106 subjects replicate the findings when discarding the subset of subjects (four subjects dropped out after the first test day, two subjects dissolved the capsules before swallowing, one subject did not sufficiently engage in the task).

MPH increased the proportion of Go responses to cues requiring a Go response depending on working memory span (Required Action x Drug x Listening Span: X^2_1=7.5, p=.006). Under MPH, higher span subjects made more Go responses to Go than NoGo cues (MPH: X^2_1=18.3, p<.001), while this was not the case under placebo (Placebo: X^2_1=1.2, p=.264). This effect of MPH was not significant across the group (independent of span) either (Required Action x Drug: X^2_1=3.2, p=.073). Thus, independent of the cue valence, MPH altered Go/NoGo responding as a function of the optimal action. Again, this effect of MPH covaried with working memory span, and not trait impulsivity (Appendix 3 - Figure 1). High span subjects made more (fewer) Go responses to cues requiring Go (NoGo) responses under MPH relative to placebo. Low span subjects showed the opposite pattern. These results could be interpreted as a cognitive enhancing effect of MPH in high span subjects, but not in low span subjects. This MPH-induced response accuracy is orthogonal to our effect of interest, and could thus not be attributed to an altered Pavlovian response bias or instrumental learning bias. Although this MPH effect on response accuracy is interesting in itself, it was not the focus of the current study, and therefore serves primarily as an invitation for future studies to assess the cognitive enhancing effects of MPH on instrumental responding.

Appendix 4 - Reaction times

For completeness, we analysed reaction times as a measure of behavioural vigour. First, we confirmed that the expected task effects are present. Second, we assessed whether the MPH effects on Go responding were accompanied by effects on RT, potentially indicative of a speed-accuracy trade-off. RT data were log(ln)-transformed to improve normality and analysed with linear mixed-level models using the lme4 package in R (Bates et al., 2014; R Development Core Team, 2015). We assessed RTs of all Go responses, irrespective of the accuracy of the Go responses, in a model including the within subject factors Drug (MPH vs. placebo), Valence (Win vs. Avoid cue), and Required Action (Go vs. NoGo), and the between subject factor Listening Span.

Regarding the expected task effects, subjects were faster when they made Go responses to Go vs. NoGo cues (Required Action: X^2_1=296.2, p<0.001), indicative of learning (i.e. faster to correct than incorrect responses). We also observed effects of the motivational bias in reaction times, where cue valence influenced RTs (Valence: X^2_1=89.5, p<0.001), such that RTs were shorter to Win vs. Avoid cues. This effect of cue valence was stronger for NoGo compared to Go cues (Required Action x Valence: X^2_1=11.5, p<0.001), though both were highly significant (Go: X^2_1=53.7, p<0.001; NoGo: X^2_1=66.6, p<0.001).

Regarding the MPH effects on RT, there was no effect of MPH on the motivational valence effect on RT (Valence x Drug: X^2_1=0.8, p=0.37), in line with the absence of any MPH main effect on Go responding. In contrast to Go responding, there were no Listening Span-dependent effects of MPH on RTs (all p>0.7). The absence of MPH effects on RTs suggests that the MPH effects reported in the main manuscript are not due to an effect on speed-accuracy trade-off. Perhaps of interest, but beyond the scope of this article, is that we did observe span-dependent effects independent of drug treatment. Higher span subjects sped up more for Win relative to Avoid cues (Valence x Listening Span: X^2_1=4.2, p=0.041), and for Go relative to NoGo cues (Required Action x Listening Span: X^2_1=5.2, p=0.023). No other effects were significant (p>0.05).

Appendix 5 - Computational modelling

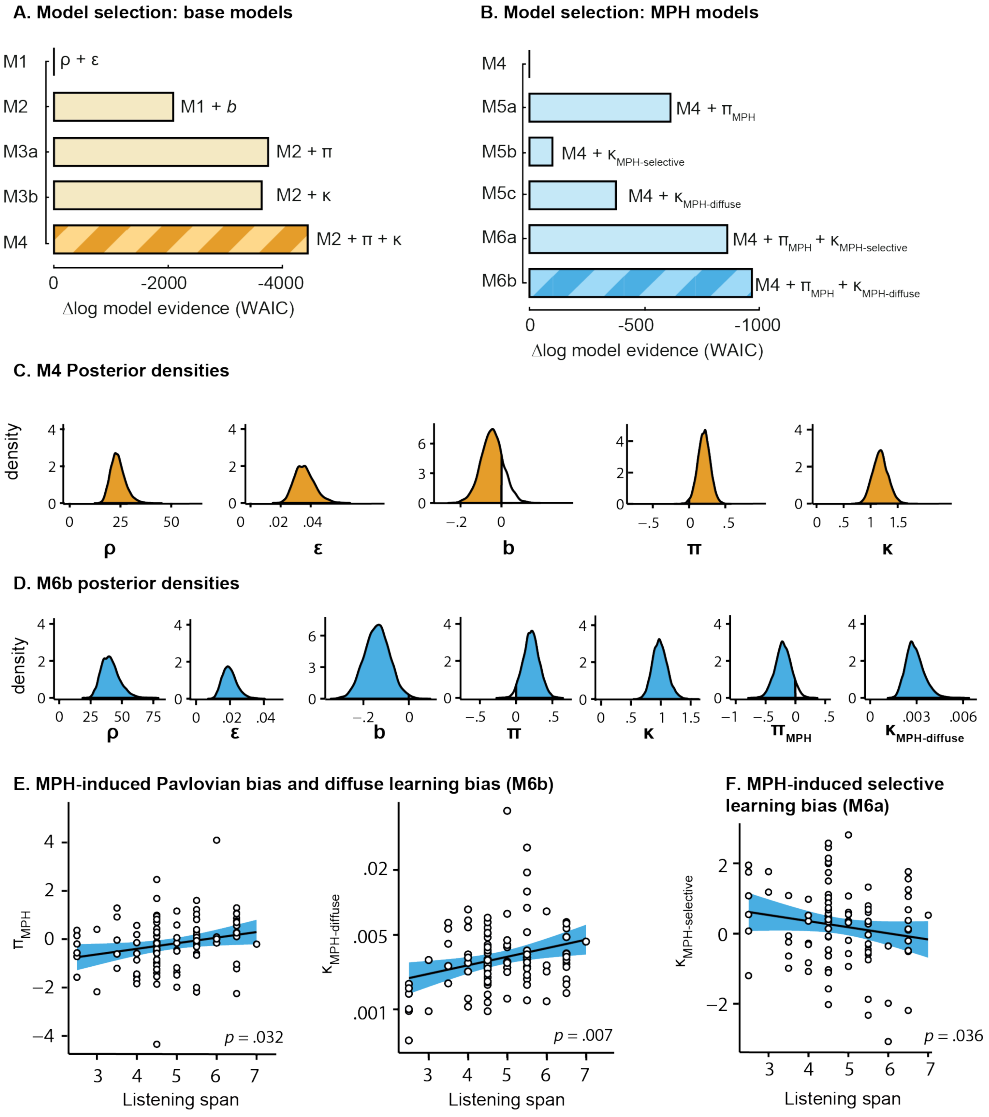
In the main article, we report five base models (M1, M2, M3a, M3b, M4) to disentangle the role of Pavlovian and instrumental learning mechanisms in driving motivational biasing of action. The winning base model was then extended in three competing models (M5a-c) and a composite model (M6) to assess the effects of MPH on these mechanisms. Not all models reached convergence when including all subjects of the behavioural analysis (N=99). For 5 subjects, $\hat{R}$ exceeded 1.1 in one or more of the models M1/M2/M5a/M6. Therefore, we repeated model estimation while excluding the 5 subjects for whom initially $\hat{R}$ exceeded 1.1 in any one of the models, resulting in model convergence for all models (see main article). In Appendix 5 - Figure 1A-E we report the model comparison results and parameter inference for the models including all subjects, to demonstrate our conclusions do not depend on the exclusion of these 5 non-converging subjects.

Note that the go bias estimates of the winning base model M4 did not significantly deviate from 0 across the group (83 % of posterior distribution < 0, cut-off is usually considered 90%). The fact that inclusion of this go bias parameter did improve model evidence suggests large individual variance. In other words, inclusion of this parameter was important for explaining the data, but the direction of its effect was variable across subjects. It is noteworthy that the go bias estimates are on average negative (even if not significantly different from 0), in contrast to previous studies (Cavanagh et al., 2013; Guitart-Masip et al., 2012). This discrepancy likely is the result of incorporation of the additional Go response in the current paradigm, such that chance level of Go responses is 67%, rather than 50%, and so a positive bias estimate corresponds to a greater than 2/3 proportion of Go responses overall.

Furthermore, in the extended MPH models, both π_{MPH} (M5a) and $\kappa_{MPH\text{-}diffuse}$ (M5c) greatly increased model evidence, in contrast to addition of $\kappa_{MPH\text{-}selective}$ (M5b), which only marginally increased model evidence. Therefore, to assess whether both Pavlovian (π_{MPH}) and instrumental learning ($\kappa_{MPH\text{-}diffuse}$) effects explained Listening Span-dependent MPH variance independently (i.e. whether there was evidence for both of these effects), we constructed a composite model (M6 in main text; M6b in Appendix 5 - Figure 1) containing both parameters. For completeness, here we also report the composite model containing both the π_{MPH} and $\kappa_{MPH\text{-}selective}$ parameters (M6a). As expected, model selection favours the composite model with a spread of credit assignment ($\kappa_{MPH\text{-}diffuse}$, M6b; $WAIC_{N=94}$=66069) over the model that includes a strengthening of the selective instrumental learning bias ($\kappa_{MPH\text{-}selective}$, M6a; $WAIC_{N=94}$=66153). Furthermore, in this model $\kappa_{MPH\text{-}selective}$ relates negatively to Listening Span (R=.22, *p*=.036; Appendix 5 - Figure 1F), now that this model accounts for the MPH-induced Pavlovian bias variance. This negative correlation cannot explain the positive relation between the relation between working memory span and the effects of MPH on motivational bias, and as such further corroborates our conclusion that MPH does not simply increase the strength of the instrumental learning bias as a function of listening span.

In the main article, we report the subject-level parameter estimates in model space (Figure 3,5). Here we additionally report the untransformed parameter estimates (Appendix 5 - Table 1: subject-level, Appendix 5 - Table 2: top-level) and the confidence of top-level parameters deviating from 0 for each model (Appendix 5 - Table 3). In Appendix 5 - Figure 2,3 we display the one-step-ahead predictions for the both the winning and non-winning base and MPH models.

We refer to the *Decision Letter* and *Author Response* available at elifesciences.org/articles/22169#decision-letter for a discussion on the potential confound of asymmetric reward/punishment sensitivities, where we show control analyses that speak against this potential confound.



Appendix 5—figure 1. **Model selection and parameter inference for base models and extended MPH models including all subjects (N = 99).**

(**a-b**) Model selection favours M4 of the base models and M6b of the extended MPH models as reported in the main article. Note that M6b in this figure corresponds to M6 in the main manuscript. (**c-d**) Posterior densities of top-level parameters of the winning base and MPH model, in model space (i.e. transformed). Only κ is presented untransformed (i.e. in sample space), as it is added to ε_0 prior to transformation. (**e**) As reported in the main article, π_{MPH} and $\kappa_{MPH\text{-}diffuse}$ of M6b positively correlate with Listening Span. (**f**) In the composite model M6a, $\kappa_{MPH\text{-}selective}$ correlates negatively with Listening Span (N=94; Rho=–0.22, p=0.036), which further supports that this parameter cannot capture the positive relation between listening span and the effect of MPH on motivational bias. Note that we report the correlation here for the 94 subjects for whom the parameters were reliably estimated, i.e. model convergence was reached.

	Base models					Extended MPH models				
	M1	M2	M3a	M3b	M4	M5a	M5b	M5c	M6a	M6b
ρ	3.4 (1.2)	3.4 (1.2)	3.3 (1.2)	3.2 (1.1)	3.2 (1.1)	3.2 (1.1)	3.2 (1.1)	3.7 (1.3)	3.2 (1.1)	3.7 (1.3)
ε	−3.7 (1.6)	−3.7 (1.5)	−3.6 (1.5)	−3.4 (1.5)	−3.5 (1.4)	−3.4 (1.4)	−3.5 (1.4)	−4.0 (1.7)	−3.4 (1.3)	−4.0 (1.7)
b		−0.2 (0.5)	−0.2 (0.5)	−0.04 (0.6)	−0.05 (0.5)	−0.06 (0.5)	−0.07 (0.5)	−0.13 (0.5)	−0.08 (0.5)	−0.15 (0.5)
π			0.5 (0.8)		0.15 (0.7)	0.13 (1.0)	0.18 (0.7)	0.04 (0.7)	0.17 (1.1)	0.16 (0.9)
κ				1.7 (1.3)	1.2 (0.8)	1.1 (0.7)	1.1 (0.8)	1.2 (0.8)	1.1 (1.0)	1.0 (0.7)
π_{MPH}						0.08 (1.2)			0.09 (1.5)	−0.19 (1.2)
$\kappa_{MPH\text{-}selective}$							0.09 (0.8)		0.22 (1.1)	
$\kappa_{MPH\text{-}diffuse}$								−5.9 (0.7)		−5.7 (0.5)

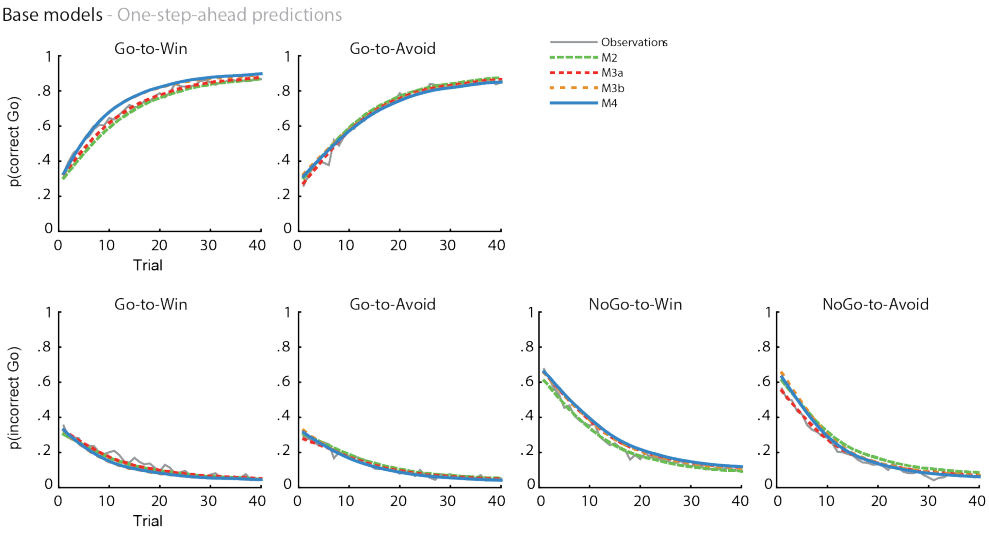
Appendix 5—table 1. **Untransformed subject-level parameter means (SD).**

	Base models					Extended MPH models				
	M1	M2	M3a	M3b	M4	M5a	M5b	M5c	M6a	M6b
ρ	3.4 (0.2)	3.4 (0.2)	3.3 (0.2)	3.2 (0.16)	3.2 (0.15)	3.2 (0.15)	3.2 (0.15)	3.7 (0.18)	3.2 (0.15)	3.7 (0.17)
ε	−3.7 (0.2)	−3.7 (0.2)	−3.5 (0.2)	−3.4 (0.19)	−3.4 (0.19)	−3.4 (0.19)	−3.4 (0.19)	−4.0 (0.22)	−3.3 (0.18)	−3.9 (0.22)
b		−0.2 (0.1)	−0.2 (0.1)	−0.04 (0.07)	−0.05 (0.06)	−0.06 (0.06)	−0.07 (0.06)	−0.13 (0.06)	−0.08 (0.06)	−0.15 (0.06)
π			0.5 (0.1)		0.15 (0.09)	0.13 (0.12)	0.18 (0.09)	0.04 (0.09)	0.17 (0.13)	0.16 (0.11)
κ				1.65 (0.21)	1.2(0.15)	1.1 (0.14)	1.1 (0.17)	1.16 (0.15)	1.09 (0.18)	1.00 (0.13)
π_{MPH}						0.08 (0.14)			0.09 (0.17)	−0.19 (0.14)
$\kappa_{MPH\text{-}selective}$							0.09 (0.19)		0.22 (0.25)	
$\kappa_{MPH\text{-}diffuse}$								−5.9 (0.24)		−5.7 (0.21)

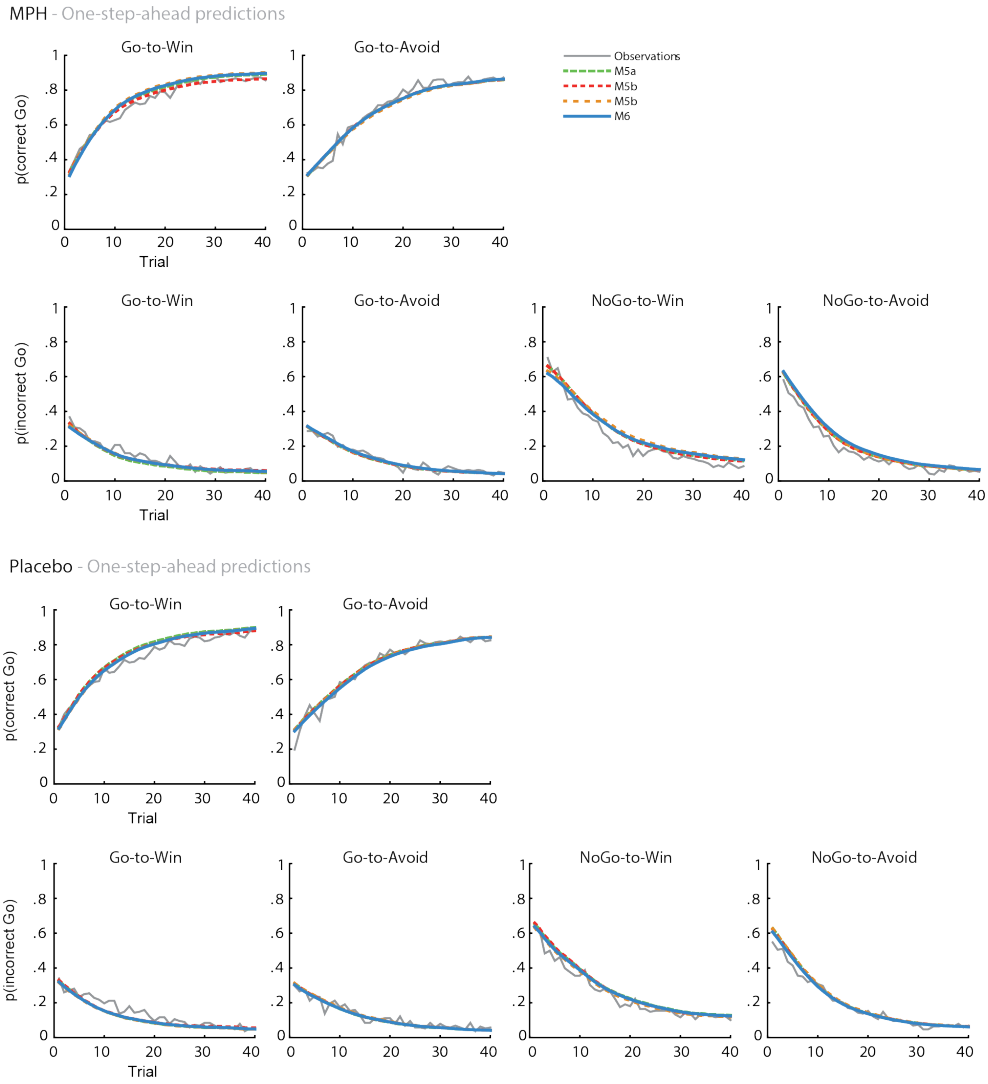
Appendix 5—table 2. **Untransformed top-level parameter means (SD).**

	Base models					Extended MPH models				
	M1	**M2**	**M3a**	**M3b**	**M4**	**M5a**	**M5b**	**M5c**	**M6a**	**M6b**
ρ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ε	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
b		0.00	0.00	0.29	0.17	0.13	0.12	0.01	0.09	0.01
π			1.00		0.96	0.87	0.98	0.70	0.91	0.91
κ				1.00	1.00	1.00	1.00	1.00	1.00	1.00
π_{MPH}						0.72			0.71	0.09
$\kappa_{MPH\text{-}selective}$							0.67		0.81	
$\kappa_{MPH\text{-}diffuse}$								0.00		0.00

Appendix 5—table 3. **Confidence/probability that top-level parameter is larger than 0.**



Appendix 5—figure 2. **Average one-step-ahead predictions for the base models M2-4 overlaid on the observations in grey.**
The one-step-ahead predictions indicate the action probabilities as predicted by the model, using each subject's actual choices and outcomes.



Appendix 5—figure 3. **Average one-step-ahead predictions for the extended MPH models M5-6 overlaid on the observations in grey.**

The one-step-ahead predictions generate the action probability of each choice, based on the history of the subject's actual choices and outcomes preceding the choice. The predictions are separately plotted for MPH (top) and placebo (bottom). We observed no main effect of MPH on the motivational bias (i.e. more Go to Win cues relative to Avoid cues). Accordingly, all models make highly similar predictions under MPH and placebo across the group.

References

Arnsten, A.F., Dudley, A.G., 2005. Methylphenidate improves prefrontal cortical cognitive function through alpha2 adrenoceptor and dopamine D1 receptor actions: Relevance to therapeutic effects in Attention Deficit Hyperactivity Disorder. Behav. Brain Funct. 1, 2. doi:10.1186/1744-9081-1-2

Barr, D.J., 2013. Random effects structure for testing interactions in linear mixed-effects models. Front. Psychol. 4, 328. doi:10.3389/fpsyg.2013.00328

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278. doi:10.1016/j.jml.2012.11.001

Barratt, W., 2006. The Barratt Simplified Measure of Social Status (BSMSS): Measuring SES. Terre Haute Dep. Educ. Leadership, Adm. Found. Indiana State Univ.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4. R Packag. version. doi:citeulike-article-id:7112638

Beck, A.T., Steer, R.A., Ball, R., Ranieri, W., 1996. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. J. Pers. Assess. 67, 588–597. doi:10.1207/s15327752jpa6703_13

Beeler, J.A., Frank, M.J., McDaid, J., Alexander, E., Turkson, S., Sol Bernandez, M., McGehee, D.S., Zhuang, X., 2012. A Role for Dopamine-Mediated Learning in the Pathophysiology and Treatment of Parkinson's Disease. Cell Rep. 2, 1747–1761. doi:10.1016/j.celrep.2012.11.014

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Duezel, E., Dolan, R., Dayan, P., 2013. Dopamine Modulates Reward-Related Vigor. Neuropsychopharmacology 38, 1495–1503. doi:10.1038/npp.2013.48

Berridge, K.C., 2007. The debate over dopamine's role in reward: The case for incentive salience. Psychopharmacology (Berl). doi:10.1007/s00213-006-0578-x

Bond, A., Lader, M., 1974. The use of analogue scales in rating subjective feelings. Br J Med Physiol. 47, 211–8. doi:10.1111/j.2044-8341.1974.tb02285.x

Bromberg-Martin, E.S., Matsumoto, M., Hikosaka, O., 2010. Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. Neuron. doi:10.1016/j.neuron.2010.11.022

Buckholtz, J.W., Treadway, M.T., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., Baldwin, R.M., Schwartzman, A.N., Shelby, E.S., Smith, C.E., Kessler, R.M., Zald, D.H., 2010. Dopaminergic Network Differences in Human Impulsivity. Science (80-.). 329, 532. doi:10.1126/science.1185778

Cacioppo, J.T., Petty, R.E., Kao, C.F., 1984. The efficient assessment of need for cognition.: EBSCOhost. J. Pers. Assess. 48, 306–307. doi:10.1207/s15327752jpa4803_13

Carver, C.S., White, T.L., 1994. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. J. Pers. Soc. Psychol. 67, 319–333. doi:10.1037/0022-3514.67.2.319

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., Frank, M.J., 2013. Frontal Theta Overrides Pavlovian Learning Biases. J. Neurosci. 33, 8541–8548. doi:10.1523/JNEUROSCI.5754-12.2013

Clatworthy, P.L., Lewis, S.J.G., Brichard, L., Hong, Y.T., Izquierdo, D., Clark, L., Cools, R., Aigbirhio, F.I., Baron, J.-C., Fryer, T.D., Robbins, T.W., 2009. Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. J. Neurosci. 29, 4690–4696. doi:10.1523/JNEUROSCI.3266-08.2009

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482, 85–88. doi:10.1038/nature10754

Collins, A.G.E., Frank, M.J., 2015. Surprise! Dopamine signals mix action, value and error. Nat. Neurosci. 19, 3–5. doi:10.1038/nn.4207

Collins, A.G.E., Frank, M.J., 2014. Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. 121, 337–66. doi:10.1037/a0037015

Cools, R., D'Esposito, M., 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol. Psychiatry. doi:10.1016/j.biopsych.2011.03.028

Cools, R., Frank, M.J., Gibbs, S.E., Miyakawa, A., Jagust, W., D'Esposito, M., 2009. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J. Neurosci. 29, 1538–43. doi:10.1523/JNEUROSCI.4467-08.2009

Cools, R., Gibbs, S.E., Miyakawa, a, Jagust, W., D'Esposito, M., 2008. Working memory capacity predicts dopamine synthesis capacity in the human striatum. J.Neurosci. 28, 1208–1212. doi:10.1523/JNEUROSCI.4475-07.2008

Cools, R., Sheridan, M., Jacobs, E., D'Esposito, M., 2007. Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. J. Neurosci. 27, 5506–5514. doi:10.1523/JNEUROSCI.0601-07.2007

Dagher, A., Robbins, T.W., 2009. Personality, Addiction, Dopamine: Insights from Parkinson's Disease. Neuron. doi:10.1016/j.neuron.2009.01.031

Daneman, M., Carpenter, P. a., 1980. Individual differences in working memory and reading. J. Verbal Learning Verbal Behav. 19, 450–466. doi:10.1016/S0022-5371(80)90312-6

Danjo, T., Yoshimi, K., Funabiki, K., Yawata, S., Nakanishi, S., 2014. Aversive behavior induced by optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in the nucleus accumbens. Proc. Natl. Acad. Sci. 111, 6455–6460. doi:10.1073/pnas.1404323111

Davis, H., Wright, J., 1979. procedural and parametric variability in studies of conditioned suppression. Bull. Psychon. Soc. 14, 149–150.

Day, J.J., Roitman, M.F., Wightman, R.M., Carelli, R.M., 2007. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat. Neurosci. 10, 1020–1028. doi:10.1038/nn1923

Dayan, P., Niv, Y., Seymour, B., D. Daw, N., 2006. The misbehavior of value and the discipline of the will. Neural Networks 19, 1153–1160. doi:10.1016/j.neunet.2006.03.002

DeLong, M.R., Wichmann, T., 2007. Circuits and circuit disorders of the basal ganglia. Arch. Neurol. 64, 20–24. doi:10.1001/archneur.64.1.20

Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18. doi:10.3758/BF03199951

Dickinson, A., Smith, J., Mirenowicz, J., 2000. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav. Neurosci. 114, 468–483. doi:10.1037/0735-7044.114.3.468

Dreyer, J.K., Hounsgaard, J., 2012. A mathematical model of dopamine autoreceptors and uptake inhibitors and their influence on tonic and phasic dopamine signaling. J. Neurophysiol. 171–182. doi:10.1152/jn.00502.2012

Estes, W.K., 1943. Discriminative conditioning. I. A discriminative property of conditioned anticipation. J. Exp. Psychol. 32, 150–155. doi:10.1037/h0058316

Estes, W.K., Skinner, B.F., 1941. Some quantitative properties of anxiety. J. Exp. Psychol. 29, 390–400. doi:10.1037/h0062283

Faraone, S. V., Buitelaar, J., 2010. Comparing the efficacy of stimulants for ADHD in children and adolescents using meta-analysis. Eur. Child Adolesc. Psychiatry 19, 353–364. doi:10.1007/s00787-009-0054-3

Faure, A., Reynolds, S.M., Richard, J.M., Berridge, K.C., 2008. Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. J. Neurosci. 28, 7184–7192. doi:28/28/7184 [pii] 10.1523/JNEUROSCI.4961-07.2008

Frank, M.J., 2006. Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. Neural Networks 19, 1120–1136. doi:10.1016/j.neunet.2006.03.006

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. 17, 51–72. doi:10.1162/0898929052880093

Frank, M.J., Fossella, J. a, 2011. Neurogenetics and pharmacology of learning, motivation, and cognition. Neuropsychopharmacology 36, 133–152. doi:10.1038/npp.2010.96

Frank, M.J., O'Reilly, R.C., 2006. A Mechanistic Account of Striatal Dopamine Function in Human Cognition: Psychopharmacological Studies With Cabergoline and Haloperidol. Behav. Neurosci. 120, 497–517. doi:10.1037/0735-7044.120.3.497

Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science (80-.). 306, 1940–1943. doi:10.1126/science.1102941

Gelman, A., 2006. Prior distribution for variance parameters in hierarchical models. Bayesian Anal. 1, 515–533. doi:10.1214/06-BA117A

Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016. doi:10.1007/s11222-013-9416-2

Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. doi:10.1214/ss/1177011136

Geurts, D.E.M., Huys, Q.J.M., Den Ouden, H.E.M., Cools, R., 2013. Aversive Pavlovian control of instrumental behavior in humans. J. Cogn. Neurosci. 25, 1428–1441.

Gibbs, S.E.B., D'Esposito, M., 2005. Individual capacity differences predict working memory performance and prefrontal activity following dopamine receptor stimulation. Cogn. Affect. Behav. Neurosci. 5, 212–221. doi:10.3758/CABN.5.2.212

Guitart-Masip, M., Duzel, E., Dolan, R., Dayan, P., 2014a. Action versus valence in decision making. Trends Cogn. Sci. doi:10.1016/j.tics.2014.01.003

Guitart-Masip, M., Economides, M., Huys, Q.J.M., Frank, M.J., Chowdhury, R., Duzel, E., Dayan, P., Dolan, R.J., 2014b. Differential, but not opponent, effects of l-DOPA and citalopram on action learning with reward and punishment. Psychopharmacology (Berl). 231, 955–966. doi:10.1007/s00213-013-3313-4

Guitart-Masip, M., Huys, Q.J.M., Fuentemilla, L., Dayan, P., Duzel, E., Dolan, R.J., 2012. Go and no-go learning in reward and punishment: Interactions between affect and effect. Neuroimage 62, 154–166. doi:10.1016/j.neuroimage.2012.04.024

Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy, R.T., Aragona, B.J., Berke, J.D., 2015. Mesolimbic dopamine signals the value of work. Nat. Neurosci. 19, 117–126. doi:10.1038/nn.4173

Hebart, M.N., Gläscher, J., 2015. Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. Psychopharmacology (Berl). 232, 437–451. doi:10.1007/s00213-014-3682-3

Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A., Galarraga, E., 1997. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. J Neurosci 17, 3334–3342.

Hernandez-Lopez, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., Surmeier, D.J., 2000. D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLC[beta]1-IP3-calcineurin-signaling cascade. J. Neurosci. 20, 8987–8995. doi:20/24/8987 [pii]

Hikida, T., Kimura, K., Wada, N., Funabiki, K., Nakanishi Shigetada, S., 2010. Distinct Roles of Synaptic Transmission in Direct and Indirect Striatal Pathways to Reward and Aversive Behavior. Neuron 66, 896–907. doi:10.1016/j.neuron.2010.05.011

Huys, Q.J.M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R.J., Dayan, P., 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. PLoS Comput. Biol. 7. doi:10.1371/journal.pcbi.1002028

Kalma, A.P., Visser, L., Peeters, A., 1993. Sociable and aggressive dominance: Personality differences in leadership style? Leadersh. Q. 4, 45–64. doi:10.1016/1048-9843(93)90003-C

Kass, R., Raftery, A., 1995. Bayes Factors. J. Am. Stat. Assoc. 90, 773–795. doi:10.1080/01621459.1995.10476572

Kim, J.-H., Son, Y.-D., Kim, H.-K., Lee, S.-Y., Kim, Y.-B., Cho, Z.-H., 2014. Dopamine D(2/3) receptor availability and human cognitive impulsivity: a high-resolution positron emission tomography imaging study with [^{11}C] raclopride. Acta Neuropsychiatr. 26, 35–42. doi:10.1017/neu.2013.29

Kimberg, D.Y., D'Esposito, M., Farah, M.J., 1997. Effects of bromocriptine on human subjects depend on working memory capacity. Neuroreport 8, 3581–3585. doi:10.1097/00001756-199711100-00032

Kimko, H.C., Cross, J.T., Abernethy, D.R., 1999. Pharmacokinetics and clinical effectiveness of methylphenidate. Clin. Pharmacokinet. 37, 457–470. doi:10.2165/00003088-199937060-00002

Kravitz, A. V, Tye, L.D., Kreitzer, A.C., 2012. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nat. Neurosci. 15, 816–818. doi:10.1038/nn.3100

Landau, S.M., Lal, R., O'Neil, J.P., Baker, S., Jagust, W.J., 2009. Striatal dopamine and working memory. Cereb. Cortex 19, 445–54. doi:10.1093/cercor/bhn095

Lee, B., London, E.D., Poldrack, R. a, Farahi, J., Nacca, A., Monterosso, J.R., Mumford, J. a, Bokarius, A. V, Dahlbom, M., Mukherjee, J., Bilder, R.M., Brody, A.L., Mandelkern, M. a, 2009. Striatal dopamine d2/d3 receptor availability is reduced in methamphetamine dependence and is linked to impulsivity. J. Neurosci. 29, 14734–40. doi:10.1523/JNEUROSCI.3765-09.2009

Lex, A., Hauber, W., 2008. Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. Learn. Mem. 15, 483–491. doi:10.1101/lm.978708

Lloyd, K., Dayan, P., 2016. Safety out of control: dopamine and defence. Behav. Brain Funct. 12, 15. doi:10.1186/s12993-016-0099-7

Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459, 837–841. doi:10.1038/nature08028
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092. doi:http://dx.doi.org/10.1063/1.1699114
Mink, J.W., Thach, W.T., 1991. Basal ganglia motor control. III. Pallidal ablation: normal reaction time, muscle cocontraction, and slow movement. J. Neurophysiol. 65, 330–351.
Montague, P.R., Hyman, S.E., Cohen, J.D., 2004. Computational roles for dopamine in behavioural control. Nature 431, 760–767. doi:10.1038/nature03015
Nassar, M.R., Frank, M.J., 2016. Taming the beast: Extracting generalizable knowledge from computational models of cognition. Curr. Opin. Behav. Sci. doi:10.1016/j.cobeha.2016.04.003
Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: Opportunity costs and the control of response vigor. Psychopharmacology (Berl). 191, 507–520. doi:10.1007/s00213-006-0502-4
Patton, J.H., Stanford, M.S., Barratt, E.S., 1995. Factor structure of the Barratt impulsiveness scale. J. Clin. Psychol. 51, 768–774. doi:10.1002/1097-4679(199511)51:6<768
Pederson, M.L., Frank, M.J., Biele, G., 2016. The drift diffusion model as the choice rule in reinforcement learning. Psychon. Bull. Rev.
Prince, J., 2008. Catecholamine dysfunction in attention-deficit/hyperactivity disorder: an update. J. Clin. Psychopharmacol. 28, S39–S45. doi:10.1097/JCP.0b013e318174f92a
R Developement Core Team, 2015. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. 1, 409. doi:10.1007/978-3-540-74686-7
Reeves, S.J., Polling, C., Stokes, P.R.A., Lappin, J.M., Shotbolt, P.P., Mehta, M.A., Howes, O.D., Egerton, A., 2012. Limbic striatal dopamine D2/3 receptor availability is associated with non-planning impulsivity in healthy adults after exclusion of potential dissimulators. Psychiatry Res. - Neuroimaging 202, 60–64. doi:10.1016/j.pscychresns.2011.09.011
Rescorla, R. a, Wagner, a R., 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Class. Cond. II Curr. Res. Theory 21, 64–99. doi:10.1101/gr.110528.110
Robbins, T.W., Everitt, B.J., 2007. A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). Psychopharmacology (Berl). doi:10.1007/s00213-006-0528-7
Robbins, T.W., Everitt, B.J., 1996. Neurobehavioural mechanisms of reward and motivation. Curr. Opin. Neurobiol. doi:10.1016/S0959-4388(96)80077-8
Salthouse, T.A., Babcock, R.L., 1991. Decomposing adult age differences in working memory. Dev. Psychol. 27, 763–776. doi:10.1037/0012-1649.27.5.763
Schmand, B., Bakker, D., Saan, R., Louman, J., 1991. [The Dutch Reading Test for Adults: a measure of premorbid intelligence level]. Tijdschr. Gerontol. Geriatr. 22, 15–19.
Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science (80-.). 275, 1593–1599. doi:10.1126/science.275.5306.1593
Schultz, W., Tremblay, L., Hollerman, J.R., 1998. Reward prediction in primate basal ganglia and frontal cortex, in: Neuropharmacology. pp. 421–429. doi:10.1016/S0028-3908(98)00071-9
Shen, W., Flajolet, M., Greengard, P., Surmeier, D.J., 2008. Dichotomous dopaminergic control of striatal synaptic plasticity. Science 321, 848–51. doi:10.1126/science.1160575
Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K.P., Dayan, P., Dolan, R.J., 2012. Dopamine and performance in a reinforcement learning task: Evidence from Parkinson's disease. Brain 135, 1871–1883. doi:10.1093/brain/aws083
Skolnick, P., 2005. Dopamine and depression, in: Dopamine and Glutamate in Psychiatric Disorders. pp. 199–214. doi:10.1007/978-1-59259-852-6_9
Smittenaar, P., Chase, H.W., Aarts, E., Nusselein, B., Bloem, B.R., Cools, R., 2012. Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection - learning or performance? Eur. J. Neurosci. 35, 1144–1151. doi:10.1111/j.1460-9568.2012.08043.x
Spencer, R.C., Devilbiss, D.M., Berridge, C.W., 2015. The cognition-enhancing effects of psychostimulants involve direct action in the prefrontal cortex. Biol. Psychiatry. doi:10.1016/j.biopsych.2014.09.013
Spielberger, C.D., Gorsuch, R.L., Lushene, P.R., Vagg, P.R., Jacobs, A.G., 1983. Manual for the State-Trait Anxiety Inventory (Form Y), Manual for the statetrait anxiety inventory STAI.
Stan Development Team, 2016. RStan: the R interface to Stan, Version 2.9.0. [WWW Document]. URL http://mc-stan.org

Steingroever, H., Wagenmakers, E.-J., 2014. Performance and awareness in the Iowa Gambling Task. Behav. Brain Sci. 37, 41–2. doi:10.1017/S0140525X13000861

Syed, E.C.J., Grima, L.L., Magill, P.J., Bogacz, R., Brown, P., Walton, M.E., 2015. Action initiation shapes mesolimbic dopamine encoding of future rewards. Nat. Neurosci. 19, 1–6. doi:10.1038/nn.4187

Taylor, J.R., Robbins, T.W., 1986. 6-Hydroxydopamine lesions of the nucleus accumbens, but not of the caudate nucleus, attenuate enhanced responding with reward-related stimuli produced by intra-accumbens d-amphetamine. Psychopharmacology (Berl). 90, 390–397. doi:10.1007/BF00179197

Taylor, J.R., Robbins, T.W., 1984. Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. Psychopharmacology (Berl). 84, 405–412. doi:10.1007/BF00555222

Thorndike, E.L., 1933. A proof of the law of effect. Science (80-.). 77, 173–175.

Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. Science 307, 1642–5. doi:10.1126/science.1105370

van der Schaaf, M.E., Fallon, S.J., Ter Huurne, N., Buitelaar, J., Cools, R., 2013. Working memory capacity predicts effects of methylphenidate on reversal learning. Neuropsychopharmacology 38, 2011–8. doi:10.1038/npp.2013.100

Volkow, N.D., Wang, G.J., Fowler, J.S., Logan, J., Franceschi, D., Maynard, L., Ding, Y.S., Gatley, S.J., Gifford, A., Zhu, W., Swanson, J.M., 2002. Relationship between blockade of dopamine transporters by oral methylphenidate and the increases in extracellular dopamine: Therapeutic implications. Synapse 43, 181–187. doi:10.1002/syn.10038

Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. J. Mach. Learn. Res. 11, 3571–3594.

Watson, D., Clark, L.A., Tellegen, A., 1988. Development and Validation of Brief Measures of Positive and Negative Affect. The PANAS Scales 54, 1063-1070.

Wechsler, D., 2008. Wechsler adult intelligence scale - Fourth Edition (WAIS-IV). San Antonio 1–3.

Wigal, S.B., Wigal, T., Schuck, S., Brams, M., Williamson, D., Armstrong, R.B., Starr, H.L., 2011. Academic, behavioral, and cognitive effects of OROS® methylphenidate on older children with attention-deficit/hyperactivity disorder. J. Child Adolesc. Psychopharmacol. 21, 121–31. doi:10.1089/cap.2010.0047

Wise, R.A., 2004. Dopamine, learning and motivation. Nat. Rev. Neurosci. 5, 483–494. doi:10.1038/nrn1406

Wyvell, C.L., Berridge, K.C., 2000. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. J. Neurosci. 20, 8122–8130. doi:http://www.jneurosci.org/content/21/19/7831

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G.C.R., Urakubo, H., Ishii, S., Kasai, H., 2014. A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science (80-.). 345, 1616–1620. doi:10.1126/science.1255514

Zalocusky, K.A., Ramakrishnan, C., Lerner, T.N., Davidson, T.J., Knutson, B., Deisseroth, K., 2016. Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. Nature 531, 642–6. doi:10.1038/nature17400

Zimet, G.D., Dahlem, N.W., Zimet, S.G., Farley, G.K., 1988. The Multidimensional Scale of Perceived Social Support. J. Pers. Assess. 52, 30–41. doi:10.1207/s15327752jpa5201_2

Chapter 4

Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action

It ain't what you don't know that gets you into trouble.
It's what you know for sure that just ain't so.
- Mark Twain

Published as:
Swart JC, Frank MJ, Määttä JI, Jensen O, Cools R, & den Ouden HEM (2018). Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. PLOS Biology, 16(10):e2005979.

Abstract

Motivation exerts control over behaviour by eliciting Pavlovian responses, which can either match or conflict with instrumental action. We can overcome maladaptive motivational influences, putatively through frontal cognitive control. However, the neurocomputational mechanisms subserving this control are unclear; does control entail upregulating instrumental systems, downregulating Pavlovian systems, or both? We combined EEG recordings with a motivational Go/NoGo learning task (N=34), where multiple Go options enabled us to disentangle selective action learning from non-selective Pavlovian responses. Midfrontal theta-band (4-8Hz) activity covaried with the level of Pavlovian conflict, and was associated with reduced Pavlovian biases, rather than reduced instrumental learning biases. Motor and lateral prefrontal regions synchronized to the midfrontal cortex, and these network dynamics predicted the reduction of Pavlovian biases over and above local, midfrontal theta activity. This work links midfrontal processing to detecting Pavlovian conflict, and highlights the importance of network processing in reducing the impact of maladaptive, Pavlovian biases.

PLOS Biology Author Summary

The anticipation of reward and punishment are key drivers of behavior: we tend to take action for rewards, while holding back in the face of punishment. This motivational bias might have an overall evolutionary advantage, but can also work against us in specific situations. Here we first asked whether this motivational bias relies on innate, automatic action tendencies, or whether this bias might actually itself be learned. Secondly, we studied which brain processes reduce the impact of these motivational drives when they become dysfunctional. By comparing the actions of human participants to the predictions of several mathematical models, we showed that the motivational bias in action relies partly on automatic tendencies and partly on asymmetric learning from experience. We then observed that activity over the midfrontal cortex specifically increased as a function of how dysfunctional the automatic tendencies were. Additionally, this midfrontal cortex activity was functionally connected to the motor and lateral frontal cortex, which play a role in activating / inhibiting behavior. By incorporating this connectivity into the mathematical models, we showed that this midfrontal connectivity predicted reduced impact of dysfunctional automatic tendencies on behavior. We propose that the midfrontal cortex detects dysfunctional action tendencies, and implements cognitive control by signaling across the network.

Introduction

Potential rewards and losses are key drivers of our actions, with consequences ranging from highly desirable (e.g., raise or promotion) to unwanted (e.g., bankruptcy) and even inconceivable (the collapse of the worldwide financial market). The valence of these outcomes is particularly well known to bias our actions: whereas anticipated rewards promote taking action, anticipated losses promote holding back from taking action (Dickinson and Balleine, 1994; Guitart-Masip et al., 2014a; Huys et al., 2011). These motivational biases are often beneficial (e.g., working harder to gain a promotion, and stop spending money to avoid bankruptcy), but they can also conflict with instrumental requirements imposed by the environment (Hershberger, 1986) (e.g., the need to stop side-tracking to obtain the promotion more effectively, and work harder at job applications to avoid the bankruptcy). Fortunately, we are not enslaved to our motivational drives; we can often overcome our motivational biases and adapt to the environmental requirements, putatively by recruiting the midfrontal cortex (Cavanagh et al., 2013; Cavanagh and Frank, 2014) as a hub in frontal control networks (Cohen, 2011). Here, we set out to uncover the neurocomputational mechanisms by which the midfrontal cortex reduces motivational control over our actions.

The well-established motivational biases in action (Davis and Wright, 1979; Duffy, 1962; Estes, 1943; Estes and Skinner, 1941; Geurts et al., 2013; Huys et al., 2011; Swart et al., 2017) arise at least partly from Pavlovian mechanisms (Estes, 1943; Estes and Skinner, 1941), such that appetitive conditioned cues globally promote behavioural activation, and aversive conditioned cues globally inhibit behavioural activation, independently of instrumental requirements. In other words, cue valence elicits non-selective Pavlovian (in)action, rather than enhancing selective instrumental responses. These Pavlovian response tendencies can be helpful in reducing computational load by shaping our actions in a hardwired manner (Dayan et al., 2006). Consequently, the Pavlovian response tendencies become disadvantageous when they are incongruent with instrumental requirements, requiring us to rely more on the relatively flexible, yet slower, instrumental system. We have recently demonstrated, however, that the instrumental system is susceptible to motivational valence biases as well, which are manifest as a result of biases in learning rather than directly on choice (Swart et al., 2017). More specifically, reward outcomes are more effective in reinforcing specific Go actions and less effective in reinforcing NoGo responses. Conversely, punishment outcomes are more effective in inducing avoidance of specific Go responses that preceded them and less effective following NoGo responses. These biases are consistent with an emerging understanding of the dopaminergic mechanisms of corticostriatal plasticity (Collins and Frank, 2014). Moreover, this instrumental learning bias explains a substantial portion of the behaviour otherwise attributed to a Pavlovian system. Taken together, Pavlovian and instrumental learning biases complementarily contribute to the well-known motivational bias in action.

The motivational biasing of action is often adaptive, but becomes maladaptive when

these biases are incongruent with the environmental requirements. The midfrontal cortex is the key region that has been linked to reducing motivational biases when these biases become maladaptive (Cavanagh et al., 2013; Cavanagh and Frank, 2014), which has been assumed so far to reflect a modulation of the Pavlovian response biases. However, previous studies did not disentangle Pavlovian and instrumental learning biases, and thus the reduced motivational biases might just as well reflect a modulation of the instrumental learning biases. Moreover, the midfrontal cortex has extensively been linked to reinforcement learning signals, both at time of response and feedback (Hauser et al., 2014; Holroyd and Coles, 2002; Marco-Pallares et al., 2008; van de Vijver et al., 2011; Van de Vijver et al., 2014), and has been proposed to modulate learning in downstream target areas (Cohen et al., 2011). Therefore, it is pertinent to assess whether the midfrontal signals are related to modulations of the Pavlovian or instrumental system.

Here, we propose that the midfrontal cortex is specifically implicated in reducing motivational biases by detecting and signalling conflict between the Pavlovian and instrumental systems, relying on similar neural mechanisms as evident in classic response conflict. Classically, the midfrontal cortex has been linked to detection of response conflict (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2011; Pastötter et al., 2013; van Driel et al., 2015), signalling the need for cognitive control to elevate the decision threshold and prevent impulsive responses (Aron et al., 2016; Cavanagh et al., 2011; Cavanagh and Frank, 2014; Frank et al., 2015; Herz et al., 2016; Kelley et al., 2018; Zavala et al., 2014). In classic response conflict tasks (Simon and Rudell, 1967), task-irrelevant features trigger prepotent responses that can conflict with the required response as signalled by the task-relevant features. During these conflict trials, oscillatory activity in the theta frequency range (4-8Hz) increases over the midfrontal cortex (Cohen and Ridderinkhof, 2013; van Driel et al., 2015), putatively reflecting the detection of conflict, and this activity is in turn predictive of behavioural performance (Cohen and Cavanagh, 2011). Moreover, successful resolution of response conflict is accompanied by increased functional connectivity between the midfrontal cortex and task-related regions (most notably, the dorsolateral prefrontal cortex and motor cortex) (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013), thought to reflect signalling of the increased need for control in order to elevate the decision threshold accordingly (Aron et al., 2016; Cavanagh et al., 2011; Frank et al., 2015). We hypothesized that the midfrontal cortex similarly i) detects conflict between the Pavlovian and instrumental systems, rather than modulating Pavlovian and instrumental learning biases in general, and ii) signals the Pavlovian conflict to the dorsolateral prefrontal cortex and motor cortex in order to facilitate instrumental behaviour by preventing impulsive, Pavlovian responses, yet would not be predictive of the specific required response per se.

In the current study, we employ a motivational Go/NoGo learning task with multiple Go response options (Swart et al., 2017) and concurrent EEG surface recordings to firstly

disentangle non-selective Pavlovian activation from selective instrumental responses, and test whether midfrontal theta activity covaries with the level of Pavlovian conflict, in line with detection of the Pavlovian conflict. Second, we assess whether the midfrontal theta responses are associated with reduced Pavlovian response biases, reduced instrumental learning biases, or both. Finally, we test whether synchronization of the dorsolateral prefrontal cortex and motor cortex predicts the reduction of the motivational biases over and above the local, midfrontal theta activity, in line with conflict resolution being instantiated by signalling to downstream targets (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2012; van Driel et al., 2015, 2012).

Results

Task performance: Cue and outcome valence complementarily bias motivated action

In the motivational Go/NoGo learning task, 34 healthy subjects needed to learn the correct responses (Go-left/Go-right/NoGo) by trial-and-error in order to gain rewards (Win cues) and avoid losses (Avoid cues), see Figure 1. The inclusion of multiple Go response options, enables disentangling the impact of reward / punishment value on global, Pavlovian behavioural activation from that on selective, instrumental action selection and learning (see computational modelling below).

Over trials, subjects increased Go responding to the cues requiring Go vs. NoGo responses (X^2_1=135.7, p<.001; Figure 1), indicative of task learning. Crucially, motivational valence strongly biased Go responding; subjects made more Go responses to Win than Avoid cues (X^2_1=16.3, p<.001), independent of the Go/NoGo requirements (X^2_1=.6, p=.427). This increase in Go responses was by definition incorrect for the NoGo cues, and driven by both correct and incorrect Go responses for the Go cues (correct Go: X^2_1=10.8, p=.001; incorrect Go: X^2_1=5.1, p=.024). Moreover, the accuracy of Go responses did not significantly differ for Go-to-Win and Go-to-Avoid cues (X^2_1=1.6, p=.200). Thus, we observed a clear effect of motivational valence on Go responding, which was driven by both correct and incorrect Go responses.

Previously, we showed that this motivational biasing of Go/NoGo responding could partly be attributed to generalized, Pavlovian response biasing, and partly to biased instrumental learning of selective actions (Swart et al., 2017). That is, reward cues directly promoted non-selective behavioural activation, whereas reward outcomes preferentially facilitated instrumental learning for selective Go responses compared to NoGo responses. Similar Pavlovian and instrumental learning biases were observed for punishment and inaction. In the following, we disentangle the contribution of these cue-based Pavlovian response biases and outcome-based instrumental learning biases to the observed motivational bias in Go responding, before turning to how they are impacted by frontal EEG signals.

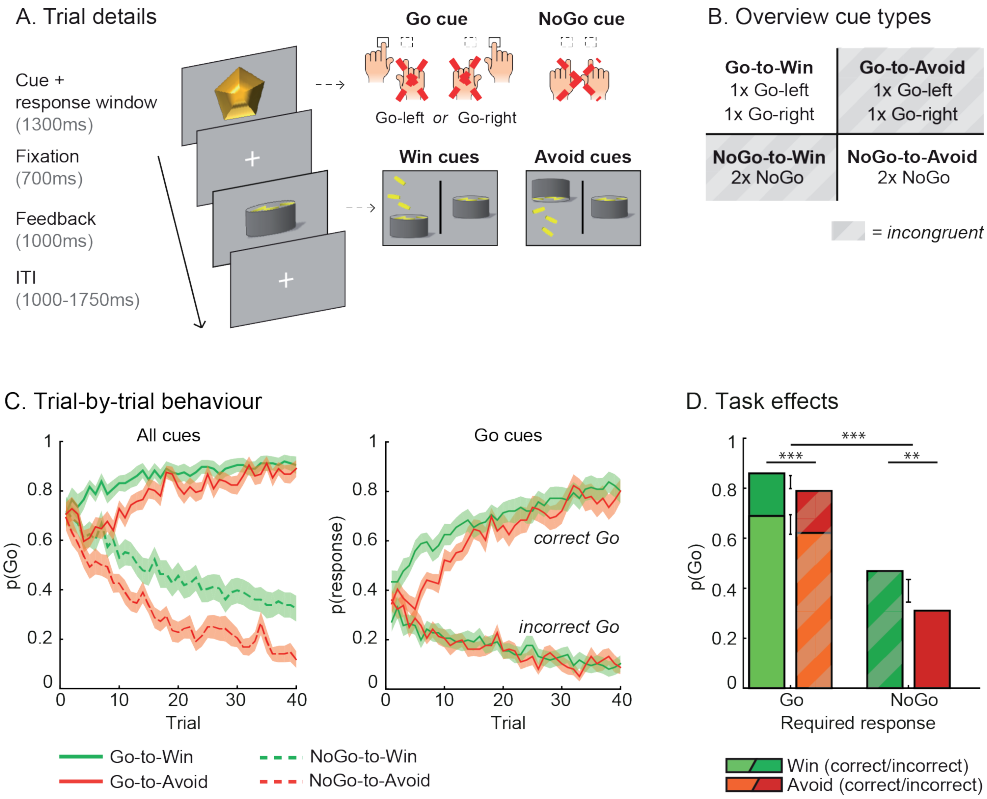


Figure 1. **Motivational Go/NoGo learning task and performance.**
(**a**) On each trial, a cue appears on screen and response-dependent feedback follows. By trial-and-error, subjects should learn to press left or right (Go cues) and withhold responding (NoGo cues) during cue presentation. Feedback is probabilistic; correct responses are followed by rewards for Win cues and neutral outcomes for Avoid cues 80% of the time, and by neutral outcomes for Win cues and punishments for Avoid cues otherwise. For incorrect responses these probabilities are reversed. Rewards and punishments are visualized by money falling in and out of a basket respectively. Image adapted from (Swart et al., 2017). (**b**) Each cue has only one correct response; Go-left, Go-right, or NoGo. In total there are 8 different cues over which cue valence (Win vs. Avoid) is orthogonal to the required action (Go vs. NoGo). The motivationally incongruent cues, where the Pavlovian response tendencies are opposite to the instrumental requirements, are marked in grey. (**c**) Average trial-by-trial behaviour (shaded areas indicate the standard error of the mean), collapsed over cues of the same category. Subjects (N=34) decrease incorrect Go responses (left: NoGo cues; right: Go cues), and increase correct Go responses over trials (right), illustrative of task learning. Note that only one of the Go responses (Go-left/Go-right) was considered correct for the Go cues, whereas all Go responses were incorrect for the NoGo cues. Notably, the initial dip in Go responses for the Go-to-Avoid cues suggests that cue valence affects Go responding once the cue valence is known, i.e. once punishment has been experienced. (**d**) Overall, subjects make significantly more Go responses for Go than NoGo cues. Orthogonal to the action requirements, subjects make more correct and incorrect Go responses for the Win than Avoid cues, which we refer to as the motivational bias. Error bars represent the standard error of the difference. ***p<.001.

We set out to disentangle the influence of Pavlovian response biases and biased instrumental learning in a computational modelling framework (Swart et al., 2017). First, we fitted a simple reinforcement-learning model (M1) to the subjects' choices, and estimated model evidence using the Watanabe-Akaike Information Criterion (WAIC), which provides a metric of model goodness, by assessing fit to the data while penalizing for model complexity. This simple model M1 included a learning rate (ε_0) and feedback sensitivity parameter (ρ; $WAIC_{M1}$=35368; R^2=33.2%). Stepwise addition of the Go bias parameter (*b;* M2), modelling a non-selective tendency towards Go responses, improved model fit ($WAIC_{M2}$=34769; R^2=34.7%). Addition of the Pavlovian response bias parameter (π; M3a; $WAIC_{M3a}$=33703; R^2=37.3%), and instrumental learning bias parameter (κ; M3b; $WAIC_{M3b}$=34261; R^2=36.0%) further improved WAIC, indicating that Go vs. NoGo responding was differentially influenced by cue and outcome valences beyond that explained by simple motor biases or instrumental learning alone. The model that best explained the task performance included both the Pavlovian response bias and instrumental learning bias parameter (M3c; $WAIC_{M3c}$=33574; R^2=37.6%; Figure 2). In other words, cue and outcome valence biased behavioural activation in a complementary fashion through instantaneous, global activation and experience-dependent selective learning, respectively. See Box 1 for an overview of the model equations.

In the winning model M3c, the Pavlovian bias estimates were positive at the group-level (95.5% group-level samples>0), indicating that Win cues globally promoted Go responding, while Avoid cues globally suppressed Go responding. The instrumental learning bias estimates were positive as well (100% group-level samples>0), indicating that reward enhanced the selective learning of Go responses (evidenced by higher learning rates for rewarded Go responses; $\varepsilon_0+\kappa$). Thus, subjects were more likely to repeat rewarded Go responses relative to rewarded NoGo responses. Conversely, punishment hampered the unlearning of NoGo responses (evidenced by lower learning rates for punished NoGo responses; $\varepsilon_0-\kappa$). In other words, subjects were more likely to repeat punished NoGo responses relative to repeating punished Go responses. As reported previously, the Pavlovian bias parameter decreased when including the instrumental learning bias parameter (median [25-75 percentile] π for M3a: .6 [-.2 1.2]; for M3c: .4 [-.5 .9]), suggesting that variance corresponding to the instrumental learning bias would otherwise be attributed to the Pavlovian bias parameter. The impact of the Pavlovian response bias and instrumental learning bias on behaviour has distinct dynamics however(Swart et al., 2017); whereas the impact of the instrumental learning bias is experience-dependent and develops over time, the impact of the Pavlovian bias reduces as instrumental action values are learned. Altogether, we observed a clear influence of both cue and outcome valence in biasing instrumental responding, consistent with our prior report (Swart et al., 2017).

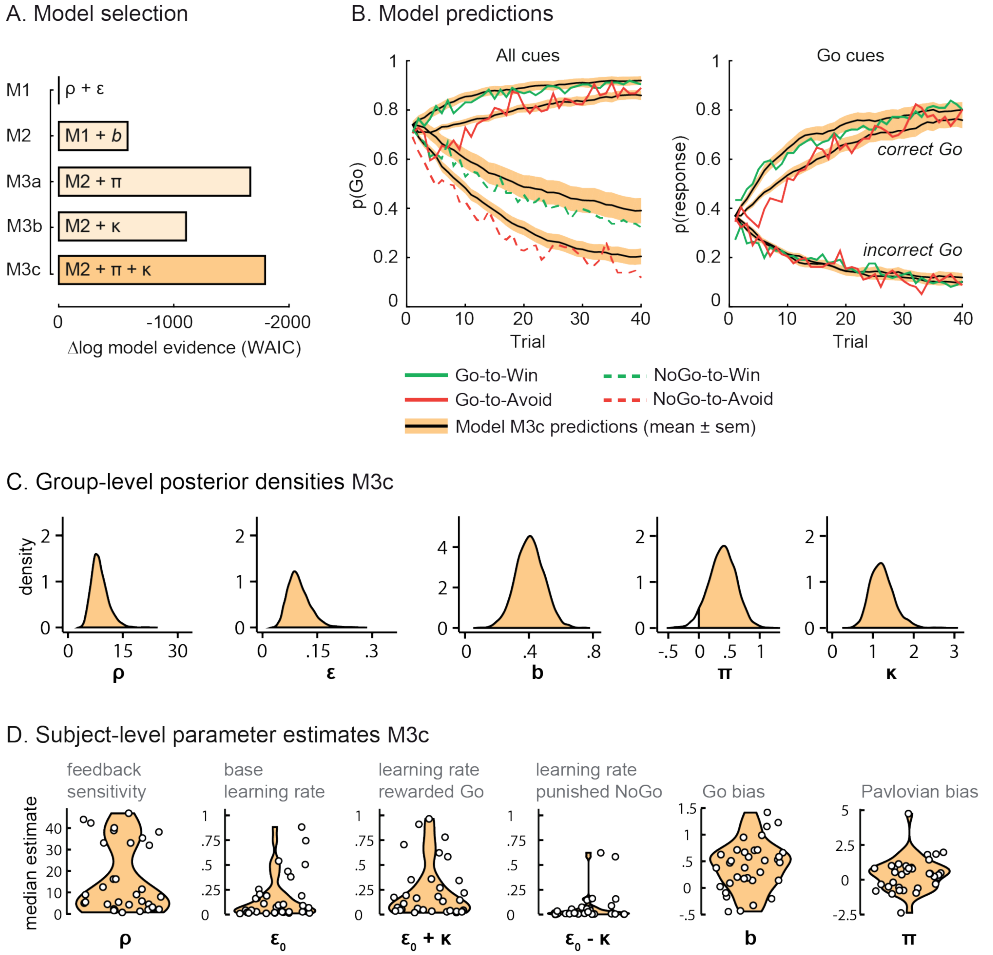


Figure 2. **Computational modelling of Pavlovian response bias and instrumental learning bias.** (**a**) Model evidence, relative to simplest model M1, favours M3c (marked by darkest colour). The simplest model M1 contains a feedback sensitivity (ρ) and learning rate (ε) parameter. Stepwise addition of the Go bias (*b*), Pavlovian bias (π), and instrumental learning bias (κ) parameter improves model fit. (**b**) Absolute post-hoc model fit. The model predictions of winning model M3c (black lines; shaded areas indicate standard error of the mean) capture the key features of the data (coloured lines); responses are learned (more Go responding for Go cues vs. NoGo cues) and a motivational bias (more Go responding for Win vs. Avoid cues). (**c**) Posterior densities of the winning model M3c. The group-level estimates are positive for the Go bias (100% samples>0), the Pavlovian bias (95.5% samples>0), and the instrumental learning bias (100% samples>0). (**d**) Subject-level parameter estimates of the winning model M3c. The feedback sensitivity and learning rate were strongly anti-correlated over subjects (R=-.77, p<.001), such that the impact of high feedback sensitivity was limited by a low learning rate. Subjects have increased learning rates for rewarded Go responses (ε_0+κ) and decreased learning rates for punished NoGo responses (ε_0-κ), which we refer to as the instrumental learning bias. Note that κ is added to ε_0 prior to [0 1]-constraining, see S2 Text.

$p(a_t\|s_t) = \left[\frac{\exp(w(a_t,s_t))}{\sum_{a'}\exp(w(a',s_t))}\right]$	Eq. 1	M1-3
$Q_t(a_t,s_t) = Q_{t-1}(a_t,s_t) + \varepsilon\left(\rho r_t - Q_{t-1}(a_t,s_t)\right)$	Eq. 2	M1-3
$w(a_t,s_t) = \begin{cases} q(a_t,s_t) + \pi V(s) + b & if\ a = Go \\ q(a_t,s_t) & else \end{cases}$	Eq. 3	M2(*b* only), M3a,c
$\varepsilon = \begin{cases} \varepsilon_0 + \kappa & if\ r_t = 1\ \&\ a = go \\ \varepsilon_0 - \kappa & if\ r_t = -1\ \&\ a = nogo \\ \varepsilon_0 & else \end{cases}$	Eq. 4	M3b-c

Box 1. **Overview of the behavioural computational models.**
In all models, the probability of each response (*a*) is estimated based on computed action weights (*w*) using a softmax function (Eq. 1; Where *t* indicates the trial and *s* the state, i.e. the cue). The action weights are determined by the learned instrumental Q values (Eq. 2, where ε is the learning rate and ρ the feedback sensitivity) and by a non-selective Go bias (*b*) and Pavlovian response bias (π) from respectively M2 and M3a onwards. Static Pavlovian values (V) were modelled once the first reward/punishment outcome has been experienced (also see S2 Text). In M3b the instrumental learning bias (κ) is introduced (Eq. 4). M3c contained both the π and κ parameters. For an elaborate description of the computational models and parameter constraints, see S2 Text.

Midfrontal oscillatory theta power reflects Pavlovian conflict

After establishing the impact of both cue-driven (Pavlovian) and outcome-driven (instrumental learning) motivational biases, we set out to replicate the finding that midfrontal oscillatory activity in the theta frequency range (4-8Hz) relates to reducing the motivational biasing of action (Cavanagh et al., 2013). Moreover, the current paradigm allows us to address specifically whether the midfrontal cortex does so by modulating the Pavlovian response bias, the instrumental learning bias, or both.

We reasoned that if midfrontal oscillatory theta activity relates to modulating the Pavlovian response bias (Cavanagh et al., 2013), midfrontal theta power would increase for motivationally incongruent cues (Go-to-Avoid, NoGo-to-Win), compared to congruent cues (Go-to-Win, NoGo-to-Avoid). As midfrontal theta has been linked to a conflict-induced adjustment of the decision threshold, preventing impulsive responses (Cavanagh et al., 2011; Cavanagh and Frank, 2014; Frank et al., 2015; Herz et al., 2016; Zavala et al., 2014), we further hypothesized that increases in midfrontal theta would be particularly evident when motivational conflict is correctly resolved. Put simply, we reasoned that a control-related signal should prevail when control is successfully implemented. To this end, we specifically assessed correct trials in line with classic cognitive control studies (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2011; Pastötter et al., 2013; van Driel et al., 2015) and report the analysis of incorrect trials in S7 Text. Time-wise permutation testing over our *a priori* midfrontal channels (see Methods), indeed revealed

a cue-locked time-window during which midfrontal theta power was significantly higher for motivationally incongruent relative to congruent cues (450ms - 650ms; p=.024; Figure 3). Importantly, the significant time-window was prior to the average response time (mean=753ms; range=594-1077ms), as might be expected for a control-related signal. Corroborating these findings, response-locked permutation testing also indicated a pre-response time-window (-826 to -150ms; p=.002; Figure 3) during which midfrontal theta power increased for Go-to-Avoid (incongruent) trials relative to Go-to-Win (congruent) trials. Moreover, these temporally specific cue-locked and response-locked enhancements of midfrontal theta for incongruent trials were not accompanied by any converse time points in which theta power showed the reverse effect (i.e., increases for congruent relative to incongruent). Altogether, we observed a clear midfrontal theta response to motivational conflict, where theta power increased prior to correctly responding, when cue valence was incongruent with the instrumental Go/NoGo requirements.

Recent theories suggest that the midfrontal cortex might be responsible for detection of conflict (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2011; Pastötter et al., 2013; van Driel et al., 2015) and signalling the need for control to downstream targets (other cortical and subcortical sites) (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2012; van Driel et al., 2015, 2012). Accordingly, we reasoned that if midfrontal theta power relates to the detection of Pavlovian-instrumental conflict, the midfrontal theta signal would covary with the level of conflict. To this end, we assessed whether there was evidence for a trial-by-trial relation between the cue-locked midfrontal theta power and the level of Pavlovian-instrumental conflict. Here, we quantified Pavlovian-instrumental conflict by the extent to which instrumental values for NoGo responses were higher than those for Go responses on Win trials, and opposite for Avoid trials. Across subjects, this correlation was significantly positive (M3c: t_{29}=2.9, p=.007; see Figure 3), supporting the trial-by-trial relation of midfrontal theta power and the level of motivational conflict. In sum, midfrontal theta power covaried with the level of motivational conflict, putatively reflecting the detection of conflict.

To assess whether the conflict-related midfrontal theta signal was indeed associated with reducing the motivational biasing of action, we extended our computational modelling approach. We employed the models developed by Cavanagh and colleagues (2013) to test the following differential mechanisms by which midfrontal theta might modulate the motivational biasing of action; We tested whether midfrontal theta power altered behaviour by modulating the Pavlovian response tendencies (M4a), the instrumental contribution (M4b), and/or the balance between the Pavlovian and instrumental contribution (M4c). Note that the current design is particularly well-suited to differentiate between these alternative mechanisms (M4a-c), as the multiple Go options enable us to disentangle whether frontal theta facilitates selection of particular instrumental actions or

rather reduces global Pavlovian (in)activation. Finally, we tested the alternative possibility that cue-related midfrontal theta power instead modulates the instrumental learning bias. That is, midfrontal theta power at the time of the decision might affect instrumental learning that takes place thereafter. We assessed this novel hypothesis in model M4d. See Box 2 for an overview of the equations for the EEG model extension.

$$w(Go'_t, s_t) = \begin{cases} Q(Go'_t, s_t) + (\pi + \beta * \theta_t) * V(s) + b & \textit{if conflict} \\ Q(Go'_t, s_t) + \pi V(s) + b & \textit{else} \end{cases} \quad \text{Eq. 1 (M4a)}$$

$$w(Go'_t, s_t) = \begin{cases} (1 - \beta * \theta_t) * Q(Go'_t, s_t) + \pi V(s) + b & \textit{if conflict} \\ Q(Go'_t, s_t) + \pi V(s) + b & \textit{else} \end{cases}$$

$$w(NoGo_t, s_t) = \begin{cases} (1 - \beta * \theta_t) * Q(NoGo_t, s_t) & \textit{if conflict} \\ Q(NoGo_t, s_t) & \textit{else} \end{cases} \quad \text{Eq. 2 (M4b)}$$

$$w(Go'_t, s_t) = \begin{cases} (1 - (\tau + \beta * \theta_t)) * Q(Go'_t, s_t) + (\tau + \beta * \theta_t) * V(s) + b & \textit{if conflict} \\ (1 - \tau) * Q(Go'_t, s_t) + \tau * V(s) + b & \textit{else} \end{cases}$$

$$w(NoGo_t, s_t) = \begin{cases} (1 - (\tau + \beta * \theta_t)) * Q(NoGo_t, s_t) & \textit{if conflict} \\ (1 - \tau) * Q(NoGo_t, s_t) & \textit{else} \end{cases} \quad \text{Eq. 3 (M4c)}$$

$$\varepsilon_{rewarded\ Go} = \begin{cases} (1 - \beta * \theta_t) * (\varepsilon_0 + \kappa) + (\beta * \theta_t) * \varepsilon_0 & \textit{if conflict} \\ \varepsilon_0 + \kappa & \textit{else} \end{cases}$$

$$\varepsilon_{punished\ NoGo} = \begin{cases} (1 - \beta * \theta_t) * (\varepsilon_0 - \kappa) + (\beta * \theta_t) * \varepsilon_0 & \textit{if conflict} \\ \varepsilon_0 - \kappa & \textit{else} \end{cases} \quad \text{Eq. 4 (M4d)}$$

Box 2. **Behavioural computational models extended with trial-by-trial EEG data.**
All EEG models were an extension of the behavioural model M3c, see Box 1. The β parameter scaled the impact of trial-by-trial theta power (θ_t) on the Pavlovian bias (Eq. 1; M4a), the instrumental contribution (Eq. 2; M4b), the balance between the Pavlovian and instrumental contribution (Eq. 3; M4c), or the instrumental learning bias (Eq. 4; M4d). In model M4c the Pavlovian bias parameter was replaced by τ, scaling the relative contribution of the Pavlovian and instrumental values. For model M5a-b, the trial-by-trial theta power estimates in Eq. 1 were replaced by trial-by-trial estimates of midfrontal-lateral prefrontal phase synchrony (M5a) and midfrontal-contralateral motor phase synchrony (M5b). For an elaborate description of the computational models and parameter constraints, see S2 Text.

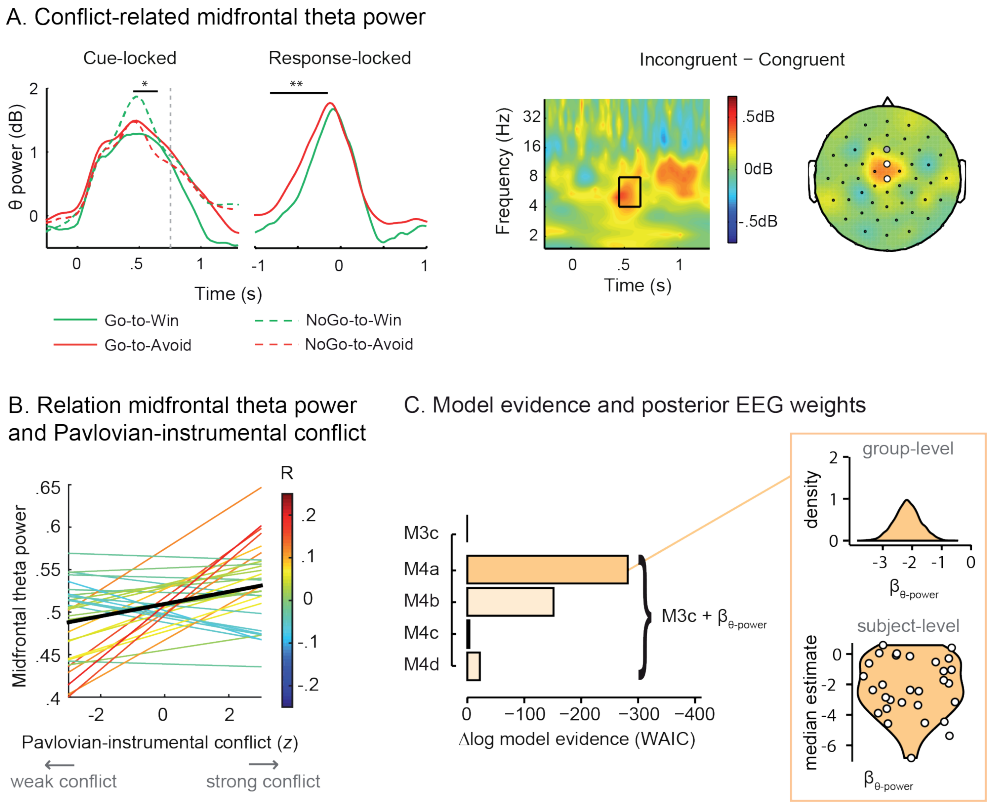


Figure 3. **Cue-related midfrontal theta power.**
(**a**) Left: Time-wise permutation testing reveals one cue-locked (450ms to 650ms) and response-locked (-826ms to -150ms) time-window during which midfrontal theta power (4-8Hz) increased for motivationally incongruent trials (NoGo-to-Win, Go-to-Avoid) relative to congruent trials (Go-to-Win, NoGo-to-Avoid). *p<.05; **p<.01. Right: Cue-locked contrast of motivational conflict. Time-frequency plot for the midfrontal channels, with the resulting time-window indicated by the box. Topoplot for theta power in the resulting time-window with the significant (non-significant) *a priori* midfrontal channels indicated by white (grey) discs (post-hoc Bonferroni corrected alpha=.017). The *a priori* channels showing a significant cue-locked congruency effect were used in the computational models and served as seeds for the connectivity analyses. (**b**) Within-subject regression lines (coloured) of trial-by-trial midfrontal theta power and Pavlovian-instrumental conflict according to model M3c. Across subjects (black line), there is a positive association between the level of conflict and midfrontal theta power (p=.007), putatively reflecting the detection of the conflict. Midfrontal theta power did not significantly covary with trial number (p=.94), rendering a confound of trials on task unlikely. (**c**) Model evidence relative to winning behavioral model M3c. Addition of the $\beta_{\theta\text{-power}}$ parameter improves model fit most when scaling the impact of theta power on the Pavlovian bias (M4a), rather than scaling the impact on the instrumental contribution (M4b), the balance between the Pavlovian and instrumental contribution (M4c), or the instrumental learning bias (M4d). The negative $\beta_{\theta\text{-power}}$ parameter estimates indicate that the conflict-related increases in midfrontal theta power were associated with reduction of the Pavlovian response bias.

For model comparison, we re-fitted the winning base model M3c over the behavioural data of the subjects included in the EEG analyses (N=30; EEG data of 4 subjects contained excessive

noise), as model evidence can only be directly compared when estimated over identical data. When allowing midfrontal theta power to scale the impact of the Pavlovian bias, model evidence greatly increased (M4a; $WAIC_{M4a}$=28503; R^2=40.5%), relative to the re-fitted base model M3c ($WAIC_{M3c,N=30}$=28785; R^2=39.8%; Figure 3), indicating that trial-wise midfrontal theta values are informative about the modulation of Pavlovian biases over and above what could be inferred on average across trials. There was less evidence for midfrontal theta power scaling the impact of the instrumental action values (M4b; $WAIC_{M4b}$=28634; R^2=40.1%) or the balance between the Pavlovian and instrumental contribution (M4c; $WAIC_{M4c}$=28781; R^2=39.5%), although both models did better than base model M3c. Although the improvements in the R^2 values might appear small, these improvements reflect the improvement *for every subject on every trial*. Additionally, the impact of the additional parameters in the EEG models is restricted as i) EEG data was rejected for trials with EEG artefacts and noise (~10% of trials), and ii) the EEG parameters are only modelled half the trials (i.e., incongruent cues). Model evidence hardly increased relative to M3c when midfrontal theta power was allowed to modulate the instrumental learning bias (M4d; $WAIC_{M4d}$=28763; R^2=39.8%). Thus, computational modelling indeed implicated the conflict-related midfrontal theta signals in modulation of the motivational biasing of action, where the theta signals particularly scaled the Pavlovian response bias (M4a), rather than the instrumental learning bias (M4d).

In the winning model M4a, the impact of midfrontal theta power on the Pavlovian bias, as assessed by $\beta_{\theta\text{-power}}$, has negative group-level estimates (100% samples<0). These negative estimates indicate that midfrontal theta power relates to *weaker* impact of the Pavlovian response tendencies, such that higher levels of theta power are associated with better ability to resolve the Pavlovian conflict. Crucially, in model M4a midfrontal theta power reduces the Pavlovian response tendencies only under conflict between the Pavlovian response tendencies and the instrumental action values. Modelling an effect of midfrontal theta power on all trials (i.e., including motivationally congruent trials), greatly reduces model fit relative to M4a (ΔWAIC=+266), consistent with previous reports showing that midfrontal theta power relates to the decision threshold specifically on conflict trials (Cavanagh et al., 2011; Frank et al., 2015; Herz et al., 2016; Kelley et al., 2018; Zavala et al., 2014). Altogether, these results suggest that cue-related midfrontal theta power is associated with resolving Pavlovian-instrumental conflict, rather than promoting unbiased behaviour in general.

Although the cue-related midfrontal theta power was not linked to a modulation of the instrumental learning bias (M4d), theta power at the time of outcome, when reinforcement learning would occur, could in principle modulate the learning bias nonetheless. To this end, we next assessed whether feedback-related midfrontal theta power increased differentially following actions and outcomes that are consistent with instrumental learning biases. First, we determined the optimal time-window for feedback-related theta power by contrasting non-preferred outcomes (i.e. neutral outcomes for Win cues and punishment for Avoid cues) and preferred outcomes (i.e. reward for Win cues and neutral outcomes for Avoid cues). Time-wise permutation testing using the *a priori* midfrontal channels revealed one time-

window during which midfrontal theta power was significantly higher for non-preferred than preferred outcomes (150 - 476ms; p=.002; Figure 4). We then assessed whether theta power in this time-window increased for the conditions where we observed biased learning (i.e. enhanced Go learning after reward and decreased NoGo unlearning after punishment, see Task performance). We did not observe a significant change in midfrontal theta power for the biased learning conditions (rewarded Go: $\varepsilon_0+\kappa$, punished NoGo: $\varepsilon_0 - \kappa$) relative to their unbiased counterparts (rewarded NoGo, punished Go; $F_{1,28}$=1.1, p=.312, BF_{01}=4.9). Alternatively, feedback-related midfrontal theta power might reflect the amount of instrumental learning that takes place. However, midfrontal theta power did not significantly differ for the conditions where we observed relatively stronger (rewarded Go, punished Go) vs. weaker (rewarded NoGo, punished NoGo) instrumental learning as observed at the behavioural level ($F_{1,28}$=3.0, p=.095, BF_{01}=3.1). Altogether, we observed well-established feedback-related modulations of midfrontal theta power (Cohen et al., 2011), yet this feedback-related theta power was not linked to biased instrumental learning of Go/NoGo responses. For full analysis of the feedback-related theta power, including the neutral outcomes, we refer to S5 Text.

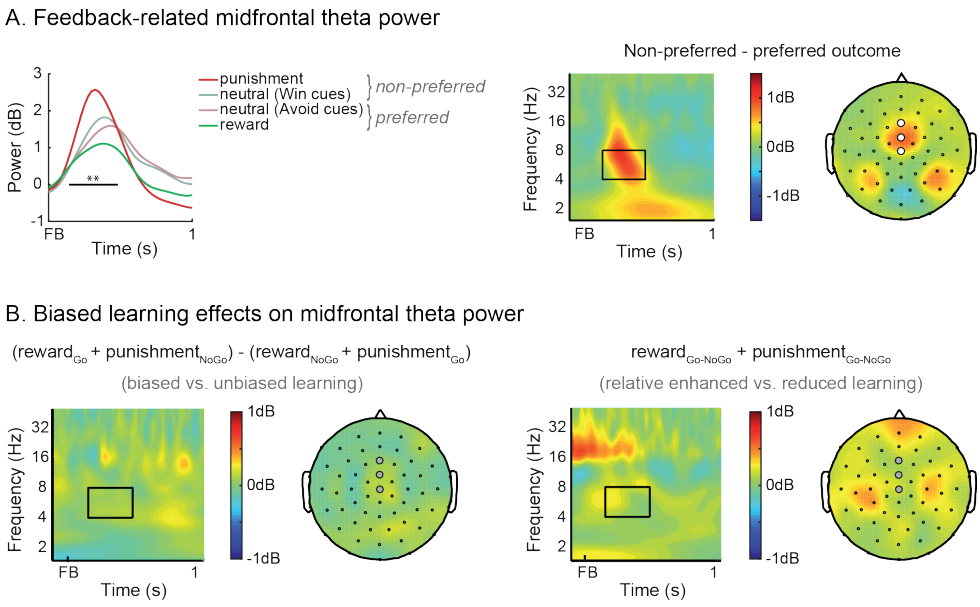


Figure 4. **Feedback-related midfrontal theta power.**
(**a**) Left: Time-wise permutation testing reveals one feedback-locked (150ms to 467ms) time-window during which midfrontal theta power increased for non-preferred relative to preferred outcomes. **p<.01. During this time-window, feedback-related theta power additionally increased more for Avoid than Win cues (p<.001; see S5 Text). Right: Contrast for non-preferred vs. preferred outcomes. The significant (non-significant) *a priori* midfrontal channels are again indicated by white (grey) discs (post-hoc Bonferroni corrected alpha=.017). (**b**) Feedback-related theta power did not vary as a function of biased learning. Left: Midfrontal theta power within the resulting time-window did not significantly increase for the

conditions where we observed biased learning at the behavioural level (rewarded Go: $\varepsilon_0+\kappa$, and punished NoGo: $\varepsilon_0 - \kappa$) relative to their unbiased counterparts (rewarded NoGo and punished Go: ε_0; BF_{01}=4.9). Right: Midfrontal theta power also did not significantly vary as a function of relative magnitude of instrumental learning. Here, we contrasted rewarded Go responses (enhanced learning; $\varepsilon_0+\kappa$) and punished Go with rewarded NoGo and punished NoGo (decreased learning; $\varepsilon_0 - \kappa$; BF_{01}=3.1).

Taken together, *cue-related* midfrontal theta power reflected the level of Pavlovian conflict and was associated with reducing the Pavlovian response tendencies, whereas *outcome-related* midfrontal theta reflected simply feedback-valence effects and was unrelated to the biased Go/NoGo learning. Together, these results suggest that the midfrontal cortex detects Pavlovian conflict and reduces the motivational biasing of action by subsequently modulating the Pavlovian response tendencies, rather than the bias in instrumental learning.

Midfrontal network dynamics reflect the ability to overcome Pavlovian response biases

In the previous section, we established that midfrontal theta power was enhanced when motivationally driven Pavlovian response tendencies conflicted with instrumental requirements, that is, on incongruent trials. We also showed that stronger trial-by-trial midfrontal theta activity within these incongruent trials were associated with reduced impact of the Pavlovian bias parameter, rather than reduced instrumental learning biases. We next hypothesized that the midfrontal cortex might instantiate this reduction of the Pavlovian response tendencies through functional connectivity with a network of task-relevant regions, specifically the dorsolateral prefrontal and motor sites, given that increased theta phase synchrony of these regions to the midfrontal cortex has been linked to conflict processing (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2012; van Driel et al., 2015, 2012). Accordingly, we reasoned that motivational congruency would differentially affect the intersite phase synchrony (ISPS) between the midfrontal channels and nodes in this network. Crucially, midfrontal ISPS has been proposed to be activity-dependent (Cohen, 2014), such that target sites become more responsive to the midfrontal signals when they are more active. Therefore, we disentangled the executing motor sites (i.e., contralateral Go responses) and the non-executing motor sites (i.e., ipsilateral Go responses and bilateral for NoGo responses) to account for putative differences in task activation due to the instantiation of overt motor responses.

Motivational congruency indeed modulated phase synchrony between the lateral prefrontal sites and the midfrontal cluster ($F_{1,27}$=6.3, p=.018; Figure 5), such that theta-band phase synchrony was strengthened for incongruent relative to congruent cues. This congruency effect was not significantly driven by either Go cues (Go-to-Avoid>Go-to-Win: t_{28}=-1.8, p=.089) or NoGo cues alone (NoGo-to-Win>NoGo-to-Avoid: t_{28}=1.7, p=.093). The midfrontal-prefrontal phase synchrony did not show significant main effects of Valence ($F_{1,27}$=.2, p=.623), or Required Action ($F_{1,27}$=.1, p=.730). Mirroring the midfrontal-prefrontal phase synchrony,

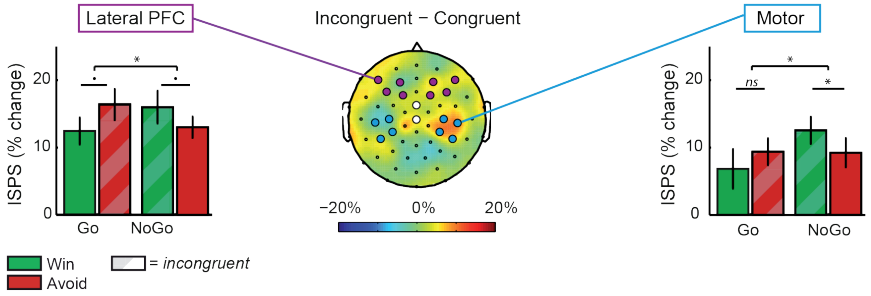
phase synchrony between the midfrontal and non-executing motor sites also increased during incongruent cues ($F_{1,27}$=5.3, p=.029; Figure 5), in the absence of a main effect of Valence ($F_{1,27}$<1, p=.856) and Required Action ($F_{1,27}$=2.3, p=.144). The motor synchrony significantly increased with motivational conflict during NoGo trials (t_{28}=2.5, p=.018, i.e. more phase synchrony on NoGo-to-Win than NoGo-to-Avoid trials), but non-significantly during the Go trials (t_{28}<1, p=.359). Taken together, we observed clear conflict-related changes in midfrontal theta phase synchrony with both the lateral prefrontal sites and the non-executing motor sites.

Figure 5 (right page). **Midfrontal intersite phase synchrony with lateral prefrontal and motor sites.** (**a**) Left: Midfrontal-lateral prefrontal phase synchrony increased for motivationally incongruent cues relative to congruent cues (p=.018). *p<.05; •p<.1. Middle: Topographic distribution of intersite phase synchrony (ISPS) for incongruent relative to congruent trials. The midfrontal channels showing the congruency effect of interest (see Figure 3) serve as t-weighted seed cluster (white discs). The lateral prefrontal and motor target channels are indicated with purple and blue discs respectively. Here we report the non-executing motor channels for simplicity (see S6 Text for an in depth discussion of the executing motor sites). Note that all subjects were right-handed, which might explain the somewhat lateralized intersite phase synchrony over the motor cortex. Right: Motor-midfrontal ISPS also increased with motivational incongruency (p=.029). (**b**) Model evidence relative to winning behavioural model M3c. Model evidence improves even further for the models where midfrontal-prefrontal ($\beta_{ISPS\text{-}PFC}$; M5a) or midfrontal-motor ($\beta_{ISPS\text{-}motor}$; M5b) phase synchrony scales the Pavlovian bias, compared to local, midfrontal theta power scaling the Pavlovian bias ($\beta_{\theta\text{-}power}$; M4a). The novel $\beta_{ISPS\text{-}PFC}$ and $\beta_{ISPS\text{-}motor}$ parameter estimates are negative (group-level both: 100%), indicating the synchrony relates to reduced Pavlovian biases. (**c**) Left: M5b model predictions (black lines; shaded areas indicate the standard error of the mean) resemble the behavioural data (coloured lines). The improvement in model predictions by the $\beta_{ISPS\text{-}motor}$ parameter is particularly apparent for the NoGo cues. Right: Effect of midfrontal-motor connectivity on the Pavlovian contribution for all individuals (coloured lines). Across the group (black line), stronger midfrontal-motor connectivity reflects improved ability to reduce the Pavlovian contribution.

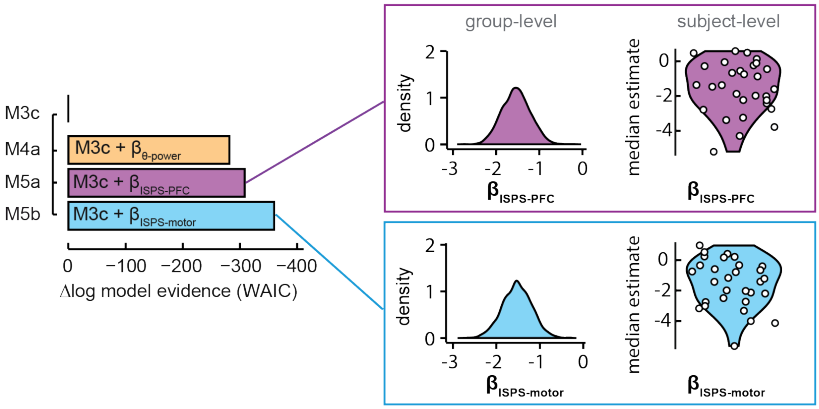
Surprisingly, we did not observe a significant congruency effect in the executing motor sites ($t_{28=}$1.7, p=.097), which was, if anything, in the opposite direction (Congruency x Motor Execution: $F_{1,27}$=10.2, p=.004; Valence x Required Action: $F_{2,54}$=4.0, p=.024). The absence of significant congruency effects for the Go responses seems to be in line with the cue-locked midfrontal theta power observations, where the congruency effect also appears stronger for the NoGo than the Go cues (Figure 3A). For a more elaborate discussion of the motor ISPS results for the executing motor sites, we refer to S6 Text.

We next assessed whether the conflict-related changes in midfrontal theta phase synchrony with the lateral prefrontal (M5a) and motor (M5b) sites could also help to explain trial-by-trial choice variability related to reducing the Pavlovian response tendencies under motivational conflict. To this end, we adopted a similar modelling approach as above, but using trial-by-trial synchrony estimates instead of power to scale the Pavlovian bias. Before doing so, we confirmed that the single-trial estimates replicate the trial-averaged results reported above (S7 Text). We then tested whether midfrontal-prefrontal phase synchrony could scale the impact of the Pavlovian bias for both Go responses, whereas the midfrontal-motor phase synchrony could scale the Pavlovian impact for the contralateral Go response option. Model evidence

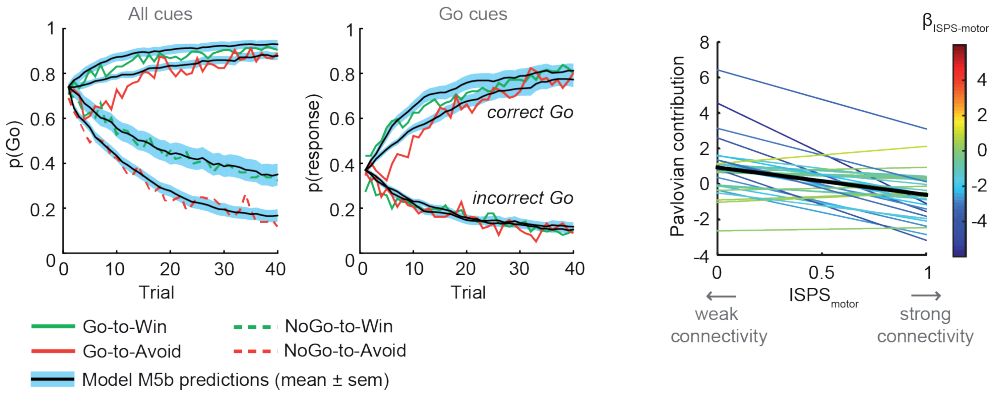
A. Midfrontal intersite phase synchrony



B. Model evidence and posterior connectivity weights



C. Model predictions



increased for both intersite phase synchrony models, compared with the winning theta-power-based model (M4a; $WAIC_{M4a}$=28503), where midfrontal-lateral prefrontal ($WAIC_{M5a}$=28477; R^2=40.6%) and midfrontal-motor phase synchrony ($WAIC_{M5b}$=28425; R^2=40.8%) modulated the Pavlovian bias. Although the trial-by-trial estimates of midfrontal theta power and intersite phase synchrony correlated positively (correlations ranging from +.12 to +.62; see S7 Text),

there was considerable unique variance within both measures (≥62%). Crucially, model evidence improved hardly any further when allowing the trial-by-trial estimates of midfrontal theta power to additionally modulate the Pavlovian bias in both synchrony models (both models: ΔWAIC=-7; $WAIC_{PFC}$=28470; $WAIC_{motor}$=28418). In other words, while theta intersite phase synchrony explained behaviour over and above midfrontal theta power ($\Delta WAIC_{PFC}$=-33; $\Delta WAIC_{motor}$=-85), midfrontal theta power barely explained any variance in addition to the phase synchrony. Thus, choices on incongruent trials were better explained when we allowed modulation of the Pavlovian response bias by intersite theta phase synchrony between the midfrontal and both task-relevant sites (lateral prefrontal and motor cluster). Alternative, target-dependent effects of intersite phase synchrony, such as lateral prefrontal synchrony modulating the goal representations and motor synchrony modulating motor excitability, reduced model evidence (S8 Text). See Figure 5 for model comparison and absolute model fit of the winning EEG model.

In this winning family of synchrony models, the novel β parameters, scaling the impact of the midfrontal-lateral prefrontal (β_{PFC}; M5a) and midfrontal-motor (β_{motor}; M5b) phase synchrony on the Pavlovian bias, have negative group-level estimates (both: 100% samples<0). These negative estimates indicate that midfrontal synchrony with these task-relevant areas is related to *weaker* impact of the Pavlovian response tendencies, such that stronger connectivity is associated with better ability to resolve the Pavlovian conflict (Figure 5). Altogether, these results indicate that trial-by-trial theta-band synchrony within a network of task-relevant regions provides a better explanation of the ability to resolve Pavlovian conflict than theta power over the midfrontal cortex alone. This is in line with recent theories suggesting that the midfrontal cortex might be responsible for detection of the conflict and hence the need for control, whereas the implementation of control is signalled by synchrony with downstream targets (other cortical and subcortical sites) (Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2012; van Driel et al., 2015, 2012).

Discussion

Motivation is a key driver of our actions. Here, we showed that both Pavlovian and instrumental learning mechanisms contribute to the motivational biasing of action, coupling action to reward and inaction to punishment. Theta power increased over the midfrontal cortex particularly when the prepotent, Pavlovian response tendencies conflicted with the instrumental task requirements. This conflict-related theta signal was associated with reduced Pavlovian response biases, rather than with reduced instrumental learning biases or enhanced specific instrumental responses. This conflict-related theta signal was accompanied by phase synchronization of the lateral prefrontal and motor sites to the midfrontal site, and these network dynamics predicted resolution of the Pavlovian conflict over and above local, midfrontal power.

In the motivational Go/NoGo learning tasks, subjects needed to learn to make one of two active 'Go' responses or withhold responding ('NoGo') in order to obtain rewards or avoid punishments. We replicated our previous finding that reward cues promoted actions non-selectively (Pavlovian response bias), while reward outcomes disproportionally facilitated credit assignment to selective actions (instrumental learning bias) (Swart et al., 2017). Conversely, punishment cues inhibited actions non-selectively, whereas punishment outcomes reduced credit assignment following inaction, thereby facilitating sustained inaction. For half the cues, these motivational biases were congruent with the instrumental requirements, whereas the other cues required a response that was incongruent with the motivational biases (i.e. NoGo to gain reward and Go to avoid punishment). Oscillatory theta power increased over midfrontal sites for the motivationally incongruent cues when subjects correctly performed instrumental responses, where higher levels of midfrontal theta power were associated with reduced motivational biasing of action, in line with work by Cavanagh and colleagues (Cavanagh et al., 2013). Here we assessed whether the reduced motivational biasing was driven by modulation of the Pavlovian response bias and/or the instrumental learning bias. To this end, we optimized the experimental paradigm to disentangle non-selective Pavlovian response tendencies from selective instrumental action values by including multiple active 'Go' response options, and showed that midfrontal theta power was associated with weaker contribution of the Pavlovian system (cf. M4a) rather weaker instrumental learning biases (cf. M4d), or enhanced contribution of the instrumental system (cf. M4b-c). Finally, motor and lateral prefrontal sites synchronized to the midfrontal site in the theta frequency range, and these network dynamics explained the trial-by-trial reduction of the Pavlovian contribution over and above local, midfrontal theta power (cf. M5).

The midfrontal cortex has long been linked to performance monitoring and the detection of conflict, realizing the need of control (Botvinick et al., 2001; Carter et al., 1998; Cavanagh and Frank, 2014; Cohen, 2014) and increasing the expected value of control (Shenhav et al., 2013). Detection of conflict has been suggested to be implemented as a coincidence detection in the midfrontal cortex, where co-activation of competing response alternatives putatively generates theta-band oscillations (Cohen, 2014). In line with these accounts, we observed that oscillatory theta power over the midfrontal cortex covaried with the trial-by-trial conflict between Pavlovian and instrumental controllers, which might reflect coincidence detection within the midfrontal cortex of the co-activation of the competing Pavlovian and instrumental response tendencies. Moreover, midfrontal theta power seemed to be particularly modulated by Pavlovian-instrumental conflict when subjects were able to perform the correct instrumental response (see S7 Text), suggesting that subjects might not have detected conflict between the Pavlovian and instrumental controllers on the incorrect trials, leaving them to follow their prepotent, Pavlovian response tendencies. Using computational model based analyses, we showed that these midfrontal theta signals were particularly predictive of the trial-by-trial

reduction of the Pavlovian response tendencies when Pavlovian and instrumental controllers conflicted (M4a). Note, however, that the midfrontal theta power did not seem to be a pure accuracy signal, since midfrontal theta power was less predictive of the specific instrumental action values (cf. M4b), most closely reflecting the correct responses. Furthermore, we did not observe a link between midfrontal theta power and the instrumental learning bias, which taken together suggests that the midfrontal cortex is not responsible for unbiased, normative behaviour per se, but rather detects conflict between competing response systems. The midfrontal theta signal is indeed often considered to be uninformative about what the resulting response *should be*, but rather conveys the signal that conflict resolution is needed (Cavanagh and Frank, 2014; Cohen, 2014). In other words, midfrontal theta power might signal conflict, potentially allowing for an increase in the decision threshold (Aron et al., 2016; Cavanagh et al., 2011; Frank et al., 2015; Herz et al., 2016; Kelley et al., 2018; Wiecki and Frank, 2013; Zavala et al., 2014), and thereby overcoming the Pavlovian response bias, but without directly selecting an instrumental response. Although we link our work to previous studies demonstrating the relation between midfrontal theta power and the decision threshold (i.e., using drift diffusion models for two response options) (Cavanagh et al., 2011; Frank et al., 2015; Herz et al., 2016), generalizing classic response conflict to Pavlovian-instrumental conflict, we could not directly test this relation in the current study, as these drift diffusion models are not suited to assess more than two response options.

The classical performance monitoring theory proposed that the need for control is signalled to the dorsolateral prefrontal cortex, which will then implement the control (Botvinick et al., 2001). More recent work suggested that the midfrontal cortex functions as a hub and signals the need for control to a wider network, including the lateral prefrontal cortex, motor cortex, ventral striatum, and subthalamic nucleus, in order to increase the decision threshold (Cohen, 2011; Cohen and Cavanagh, 2011; Cohen and Ridderinkhof, 2013; Nigbur et al., 2012; van Driel et al., 2015, 2012). Instead of assuming that the midfrontal cortex knows which areas to send the information to, the theta signal might be conveyed to all nodes in the network, while only the active areas are susceptible for the midfrontal theta-band input (Cohen, 2014), giving rise to activity-dependent functional connectivity. Accordingly, control-related functional connectivity has been observed between the midfrontal cortex and several target sites, depending on the task at hand (Cavanagh and Frank, 2014). Consistent with these theories, we observed a motivational conflict-related increase in phase synchrony between the midfrontal and motor sites, and between the midfrontal and lateral prefrontal sites. Note, however, that techniques with higher spatial resolution (e.g., fMRI) are needed to determine whether we can attribute the lateral prefrontal synchronization to the dorsolateral prefrontal cortex. Here, we reasoned that if control is instantiated by signalling the increased need for control to task-relevant sites (Cavanagh and Frank, 2014; Cohen, 2014), fluctuations in phase synchrony between the midfrontal and lateral prefrontal and motor sites would help to predict

the reduction of the Pavlovian contribution. These network dynamics indeed explained the reduction of the Pavlovian response tendencies over and above local, midfrontal theta power. Thus, stronger trial-by-trial midfrontal phase synchrony with the lateral prefrontal and motor sites predicted reduced Pavlovian response biases better than midfrontal theta power alone. It should be noted that, although the estimation of theta phase angles is orthogonal to the estimation of theta power, these estimations are not completely independent, that is, the phase estimation becomes more precise with more power. Yet, the intersite phase synchrony explains behavioural performance over and above local theta power, whereas adding power to the synchrony models hardly improves model fit. Thus, the distal phase synchrony could account for the variance explained by local power, while local power could not account for all variance explained by distal phase synchrony. Therefore, the synchrony results cannot purely be attributed to changes in power. Nevertheless, future studies could establish the contribution of network connectivity more independently by assessing connectivity measures that are independent of task-related local or network activation, such as structural or resting state connectivity.

It has been proposed that the conflict-related intersite phase synchronization would have differential computational effects depending on the function of the target circuit; the lateral prefrontal cortex might enhance the goal representations, whereas the motor cortex might increase the motor threshold (Cohen, 2014). Here, we did not observe evidence for these target-dependent effects. Specifically, we tested with alternative models whether synchronization of the lateral prefrontal sites was predictive of enhanced goal-representations, i.e. instrumental action values, whereas synchronization of the motor sites was predictive of increased motor thresholds (S8 Text). These alternative models were inferior to models that implemented a direct modulation of the Pavlovian response bias, which might reflect the increased decision threshold enabling non-hardwired decision systems to take over. Altogether, these results seem to indicate that the theta-band network dynamics reflect the signalling of the need for control, rather than signalling what specific computations are required within the target sites in order to implement the control. The resulting local computations could differ between target regions nonetheless, but are not reflected in the theta frequency band.

Finally, we have demonstrated increased midfrontal theta power for motivationally incongruent cues and linked trial-by-trial midfrontal theta power to the level of Pavlovian-instrumental conflict estimated from behaviour. However, cognitive factors other than conflict might also play a role during motivational incongruency. Midfrontal theta power has been linked to the more general notion of 'cognitive control' (Cavanagh and Frank, 2014), including for example conflict, novelty, punishment, error processing, and cognitive effort. While some factors (such as predicted response times and cue preferences) were orthogonal to motivational congruency in the current findings, we do not exclude the possibility that other factors (such as anxiety and cognitive effort allocation, which could be considered inherent to conflict resolution) might be important during motivational incongruency as well.

To summarize, in this study we replicate previous findings that i) the well-established motivational biasing of action arises partly from cue-based Pavlovian response biasing, and partly from outcome-based instrumental learning biases (Swart et al., 2017), and that ii) oscillatory theta power over the midfrontal cortex is associated with reduced motivational biasing of action (Cavanagh et al., 2013). The midfrontal theta activity covaried with the level of conflict between Pavlovian and instrumental responses, putatively reflecting the detection of conflict, and was not linked to the instrumental learning biases. This conflict-related midfrontal signal was specifically associated with reduced prepotent, Pavlovian response tendencies, without selecting the specific instrumental response per se. Synchronization of the lateral prefrontal and motor sites to the midfrontal site predicted the reduced contribution of the Pavlovian system even better than the local, midfrontal activity, which highlights the importance of investigating distributed, network processing in addition to local processing. Altogether, these findings suggest that the midfrontal cortex signals conflict to the network of task-related regions in order to putatively increase the decision threshold, and thereby overcome the prepotent, Pavlovian responses, and allow for goal-directed behaviour.

Methods

Subjects.

Thirty-four healthy adults participated in the study (aged 18-30 years, mean (SD)=23.2 (3.6); 27 females; right-handed; normal or corrected-to-normal vision). Exclusion criteria comprised a history of neurological or psychiatric disorders, use of psychotropic drugs, pregnancy, claustrophobia, and colour-blindness. Subjects signed informed consent prior to participation and received a financial compensation or study credits upon completion of the experiment. EEG data of four subjects were excluded due to excessive muscle artefacts in the EEG signal (see EEG data acquisition and preprocessing). All procedures were approved by the local ethics committee (CMO / METC Arnhem Nijmegen: CMO2014/288) and in accordance with the Helsinki Declaration of 1975.

Motivational Go/NoGo learning task.

We employed a motivational Go/NoGo learning task with multiple active response options (Swart et al., 2017) to dissociate non-selective Pavlovian activation from biased instrumental learning of selective responses. In this learning task, subjects need to learn to make Go or NoGo responses to maximize rewards for Win cues and minimize punishments for Avoid cues (Figure 1).

Trials start with presentation of a gem-shaped cue (1300ms), followed by a fixation cross (700ms), and feedback (1000ms). Four cues are followed by reward or neutral outcomes (Win cues), and four cues by punishment or neutral outcomes (Avoid cues). During cue presentation,

subjects make a button press with the left hand (Go-left) or right hand (Go-right), or withhold from responding (NoGo). Only one of these response options is considered correct per cue, and based on the response, subjects receive feedback. Correct responses are followed by a reward (Win cues) and a neutral outcome (Avoid cues) 80% of the time, and by a neutral outcome (Win cues) and a punishment (Avoid cues) otherwise. For incorrect responses, these probabilities are reversed. Based on the observed outcomes, subjects need to learn the optimal responses by trial-and-error. Trials end with an inter-trial interval (ITI), varying from 1000 to 1750ms in steps of 250ms.

Subjects are informed that i) each cue can be followed by either reward or punishment, ii) each cue has one optimal response, iii) feedback is probabilistic, and iv) the rewards and punishments are converted to a monetary bonus upon completion of the study. The monetary bonus ranged from 0 to 5 euro (mean=2.12, SD=1.70). In contrast to our previous study (Swart et al., 2017), cue valence is not instructed and can be learned from the feedback. In total, there are 8 cues with 40 trials per cue. After every 80 trials (~6min), subjects have a self-paced break. Each subject performed the task twice, using two independent cue sets. The order of the cue sets is counterbalanced, and cue identities are randomized. The second round of the task is followed by a forced-choice transfer phase (Cavanagh et al., 2013) (See S4 Text).

EEG data acquisition and preprocessing.

EEG data were acquired at 500Hz from 65 channels (using a Brain Products actiCAP system; http://brainproducts.com) placed according to the equidistant arrangement, under and above the left eye for vertical EOG, and lateral to the eyes for horizontal EOG. The ground was placed on the forehead and the left mastoid was used for online referencing. EEG data are preprocessed and analysed with the Fieldtrip software toolbox (Oostenveld et al., 2011) in MATLAB (The MathWorks). EEG data were re-referenced offline to the weighted average of the mastoids and the EEG signal of the reference electrode was recovered. Vertical and horizontal EOG electrodes were re-referenced into a bipolar montage. Data were high-pass filtered at 0.5Hz and epoched into segments starting 1.75s before cue onset and ending 1.5s after feedback offset. The epochs were made sufficiently long to avoid edge artefacts, resulting from time-frequency decomposition, in the critical times of interest. Epochs were linear baseline corrected using the 200ms prior to cue onset, and visually inspected for trial rejection. Trials were rejected when containing EMG or artefacts unrelated to brain activity, but not eye blinks. Four subjects had excessive EMG activity (>30% rejected trials) and were excluded from the EEG analyses. Of the remaining subjects, 1.7-23.6% of the trials were rejected (mean=10.5%). Independent component analysis (ICA) was performed over the EEG and vertical EOG data of the remaining epochs; components related to eye blinks or artefacts that were clearly distinguishable from brain activity were removed. One to five components were removed per subject (mean=2.5). Two subjects had one channel containing flat lines; these channels were

discarded and interpolated after ICA. Next, the EEG data were spatially filtered, using estimation of the surface Laplacian (Oostendorp and Van Oosterom, 1996; Perrin et al., 1989). The surface Laplacian filters out distant effects and accentuates local effects, thereby diminishing the effect of volume conduction on synchrony estimates (Srinivasan et al., 2007).

EEG time-frequency decomposition.

Cue- and response-locked time series were decomposed into their time-frequency (TF) representations using wavelet convolution. Here the time series are convolved with Morlet wavelets, i.e. sine waves convolved with a Gaussian, based on multiplication in the frequency domain. The frequencies ranged from 1 to 50Hz in 39 logarithmically spaced steps and the Gaussian width was fixed at 4 cycles. Wavelet convolution results in a complex signal from which power and phase are extracted for each TF point, down sampled to 40Hz. Condition specific power values were averaged over trials and dB baseline corrected using a condition averaged baseline condition (-250ms to -50ms relative to cue onset). Phase angles (φ) were used to compute intersite phase synchrony (ISPS), which is thought to reflect intersite functional connectivity (Fries, 2015; Siegel et al., 2012; van Driel et al., 2015). ISPS was computed for each TF point in the theta frequency range (4-8Hz) as following: $\text{ISPS} = \left| \frac{1}{N} * \sum_{n=1}^{N} e^{i} \varphi_{j,t,f} - \varphi_{k,t,f} \right|$ where *N* is the number of trials, *i* the complex operator, and *j* and *k* the seed and target channels. ISPS values are sensitive to the number of trials, and therefore we selected the same number of trials from each condition to compute ISPS. This trial selection was permuted 100 times, and ISPS values were averaged over permutations. Cells with fewer than 20 trials were discarded for the analysis, resulting in the exclusion of 2 subjects for this analysis. ISPS values can range from 0 (no phase synchrony) to 1 (identical phase angles), and were baseline transformed into percent signal change using a condition-averaged baseline (again -250ms to -50ms relative to cue onset).

For the single trial analyses, the Laplacian filtered EEG data were broadband filtered in the theta range (4-8Hz) and Hilbert transformed. Power time-series were extracted, z-transformed, averaged over the time-window, and t-weighted for the channels of interest (Figure 3: channel 1 and 2; 450ms to 650ms cue-locked). The t-weighting was based on the main contrast (incongruent – congruent). Phase angles were extracted to compute phase synchrony between the midfrontal seed channels and motor and prefrontal target channels during the same time-window: $\text{ISPS} = \left| \frac{1}{t} * \sum_{t=1}^{T} e^{i} \varphi_{j,n} - \varphi_{k,n} \right|$ Here, the seed channels were weighted identical to the power time-series, and the target channels were t-weighted by the corresponding main ISPS contrast (lateral prefrontal: Valence x Required Action; motor channels: Congruency x Motor Execution; Figure 5). The resulting power and ISPS values were inverse-transformed before use in the computational models.

EEG channel selection.

For all analyses we selected clusters of channels informed by prior work and fine-tuned the clusters in a data-driven manner (independent of the contrasts of interest) as scalp topographies can vary

considerably between subjects and studies. Note that we recorded EEG with an equidistant channel arrangement instead of the more commonly used 10-20 arrangement. To assess midfrontal theta power we selected two central channels corresponding to Cz and FCz in the 10-20 arrangement, as these channels showed a robust modulation of midfrontal theta power to response conflict in previous work (Cavanagh et al., 2014; van Driel et al., 2015). We then fine-tuned this midfrontal cluster based on the condition-averaged cue-locked data (orthogonal to the contrast of interest), by including one additional anterior channel, roughly corresponding to Fz (see Figure 6). With post-hoc t-tests, we assessed which of the channels contributed significantly to significant group-level cluster effects (Bonferroni corrected alpha=.017). We then combined these post-hoc significant channels using t-weighting based on the main contrast (incongruent – congruent; also see Statistical analysis), and entered the t-weighted trial-wise data in the computational models. To emphasize, the t-weighting was only performed *after establishing the effect of interest* at the group-level, and thus did not bias our results. We performed the t-weighting at the group-level rather than the subject-level to reduce proneness to noise and enhance generalizability. To assess phase synchronization of the lateral prefrontal cortex and motor cortex to the midfrontal cortex, we computed a t-weighted seed timeseries based on the significant midfrontal channels. We selected the following target clusters: i) eight dorsolateral prefrontal channels, including the dorsolateral peaks observed in the condition averaged data and extended laterally to surround the channels F5/6 in the 10-20 EEG montage (traditionally considered dorsolateral prefrontal channels (Cavanagh et al., 2009; Cohen and Cavanagh, 2011; van de Vijver et al., 2011)), and ii) four left and four right motor channel, for which we observed a clear lateralization based on the response hand independent of cue valence. See Figure 6 for the channel selection.

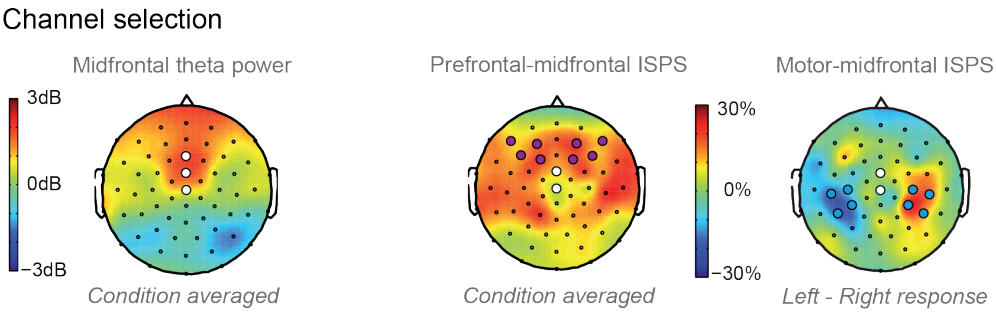


Figure 6. **Channel selection.**
Left: Topographic distribution of condition averaged theta power (4-8Hz). The three white discs indicate the midfrontal cluster. Right: Topographic distribution of intersite phase synchrony (ISPS) with the midfrontal channels serving as t-weighted seed. Only the two channels showing the effect of interest (see Figure 3) were included as seed channels. The lateral prefrontal target channels (purple discs) were selected based on the strongest condition averaged midfrontal phase synchrony and extended to include all channels surrounding F5/6 in the standard 10-20 montage, given that F5/6 are commonly assessed as dorsolateral prefrontal channels; the motor target channels (blue discs) were selected based on response lateralization (i.e. left – right responses).

Statistical analysis.

The behavioural data were analysed in line with Swart et al. (2017). We elaborate on the details of the statistical analyses and the computational models in S1 Text and S2 Text respectively.

To assess whether midfrontal theta power is related to reduced Pavlovian response biases, we analysed TF power for the midfrontal cluster and restricted the analysis to the frequency range of interest (4-8Hz) and to the response window, i.e. cue presentation. We employed a time-based permutation test (500 permutations) with the cue-locked, trial-averaged midfrontal theta power as dependent variable and the within-subject factor Congruency (congruent vs. incongruent); the Go-to-Win and NoGo-to-Avoid cues are considered to be motivationally congruent, as the Pavlovian response tendencies are in line with the instrumental requirements, whereas the NoGo-to-Win and Go-to-Avoid cues are considered to be motivationally incongruent. With post-hoc t-tests, we assessed which of the three midfrontal channels contributed significantly to the resulting time-window (Bonferroni corrected alpha=.017). Given that i) midfrontal theta power is known to strongly increase after errors(Cavanagh et al., 2012; van Driel et al., 2012), and ii) errors are more prevalent on motivationally incongruent than congruent trials, we assessed correct trials only to minimize the influence of error processing on the midfrontal theta signal. We report the analysis of incorrect trials in S7 Text. We repeated the analysis for response-locked power (Go cues only).

To test whether midfrontal theta power might also be related to reduced instrumental learning biases, we assessed feedback-related midfrontal theta power, which has been linked to reinforcement learning (Cavanagh and Frank, 2014; van de Vijver et al., 2011). In parallel to the cue-locked analysis, we selected the same midfrontal cluster and frequency range (4-8Hz) and restricted the analysis to the period of feedback presentation. To retrieve the time-window related to feedback processing, we first employed a time-based permutation test (500 permutations) with the feedback-locked, trial-averaged midfrontal theta power as dependent variable and the within-subject factor Outcome (Preferred vs. Non-preferred); reward following Win cues and neutral feedback following Avoid cues are considered to be preferred, whereas neutral feedback following Win cues and punishment following Avoid cues are considered to be non-preferred. Thus, the time-window resulting from this permutation test differentiated between the feedback. We then extracted power for the resulting TF window and contrasted the conditions with biased learning (enhanced learning for rewarded Go responses and decreased learning for punished NoGo responses) to their unbiased counterparts (rewarded NoGo and punished Go). Note that we only provide statistics for effects orthogonal to the main effect of Outcome, given that we selected the time-window based on the Outcome contrast, rendering statistics for this contrast invalid.

To assess phase synchronization of the lateral prefrontal cortex and motor cortex to the midfrontal cortex, we assessed the t-weighted phase synchronization between the midfrontal cluster and the lateral prefrontal and motor clusters. We extracted the ISPS values for the

significant time-window of the cue-locked permutation test using correct trials. We analysed ISPS using a repeated measures ANOVA with the Factors Valence (Win vs. Avoid cue) x Required Action (Go vs. NoGo) to assess if ISPS increased during motivationally incongruent cues. For the motor channels, we reasoned that ISPS might be differentially affected in the 'executing' motor cortex (i.e. contralateral to the Go response hand) relative to the 'non-executing' motor cortex (i.e. ipsilateral to the Go response hand). To elaborate, midfrontal functional connectivity is thought to be activity-dependent (Cohen, 2014), such that target sites become more susceptible to the midfrontal signals when they are more active. The motor cortex putatively becomes more activated during contralateral Go responses than during ipsilateral or NoGo responses. To this end, we split up the Go responses into contra- and ipsilateral, resulting in three levels for the factor Required Response (Go_{contra}/Go_{ipsi}/$NoGo_{bilateral}$). We then assessed for the significant Valence x Required Response interaction, whether this could be explained by an effect of motivational congruency in the non-executing motor channels (ipsilateral Go and bilateral for NoGo responses) vs. executing motor channels (contralateral Go).

Supplementary Files

S1 Text - Statistical analysis of behavioural data.

Behavioural data were analysed in line with Swart et al. (2017). In short, to assess the influence of motivational valence on behavioural activation, we first analysed Go vs. NoGo responses (irrespective of Go-left vs. Go-right) as a function of cue valence. Second, we tested whether motivational valence affected both correct and incorrect Go responses. To account for both between and within subject variability, choice data were analysed with logistic mixed-level models using the lme4 package in R (Bates et al., 2014; R Developement Core Team, 2015). These mixed models included the within subject factors Valence (Win vs. Avoid cue) and Required Action (Go vs. NoGo). The analysis of correct and incorrect Go responses included only Go cues, and hence only the factor Valence. Models included all main effects and interactions, and a full random effects structure (Barr, 2013; Barr et al., 2013). For completeness, we analysed reaction times (RTs) as a measure of behavioural vigour (see S3 Text).

S2 Text - Computational models.

First, we set out to replicate the finding that the Pavlovian response bias and instrumental learning bias contribute to the well-established asymmetric effects of valence on behavioural activation. To this end we employ the same computational modelling approach as previously (Swart et al., 2017). Using nested reinforcement learning models, we formally test whether i) *cue* valence biases behavioural activation in a Pavlovian manner, such that Win cues promote non-selective Go responses and Avoid cues promote NoGo responses, ii) *outcome* valence biases instrumental learning of (in)action; reward outcomes are more potent in reinforcing

active actions, whereas punishment outcomes are less potent in driving unlearning of holding back. Next, we extend the winning model with trial-by-trial estimates of theta power (model family M4), to assess whether these trial-by-trial neural measures increase our ability to explain behavioural responses. In particular, we assess whether on incongruent trials theta power is predictive of a) reduced Pavlovian response biases, b) enhanced impact of instrumental values, c) altered balance between these two systems, or d) reduced instrumental learning biases. Finally, we assess for the winning model, whether replacing local theta power with intersite phase synchrony in the theta range with putative target structures (lateral prefrontal cortex, motor cortex) further explains trial-by-trial behavioural responses (model family M5). Below we will describe the details of the models.

In all models, the probability of each response (*a*) is estimated based on computed action weights (*w*) using a softmax function:

$$p(a_t|s_t) = \left[\frac{\exp\left(w(a_t, s_t)\right)}{\sum_{a'} \exp\left(w(a', s_t)\right)}\right] \quad \text{Eq. 1}$$

Where *t* indicates the trial and *s* the state (i.e. the cue). In the simplest model (M1) the action weights are fully determined by the learned values of each action (Q-values). Action values are updated with the Q-value based prediction error, i.e. the deviation of the observed outcome from the expected outcome (standard delta-rule learning (Sutton and Barto, 1998), see Eq. 2). M1 contains two free parameters: a learning rate (ε) scaling the prediction-error, and feedback sensitivity (ρ) scaling the outcome value:

$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \varepsilon\left(\rho r_t - Q_{t-1}(a_t, s_t)\right) \quad \text{Eq. 2}$$

In Eq. 2, outcomes are reflected by *r,* where r ∈{-1,0,1}. As subjects can infer the cue valence upon the first reward or punishment outcome, neutral outcomes can only be evaluated as optimal (Avoid cues) or suboptimal (Win cues) after the first reward or punishment outcome. Accordingly, we tested whether subjects revalue previously experienced neutral outcomes as (sub)optimal upon the first reward or punishment. To this end, initial *Q*-values (Q_0) are set to ρ*0.5 for Win cues and ρ*-0.5 for Avoid cues (i.e. the initial expected outcome is 0.5 for Win cues and -0.5 for Avoid cues), whereas these Q-values affect behaviour only once the cue valence is known (Eq. 3). Thus, as an example, if a subject made 5 NoGo responses to a cue and receives 5 neutral outcomes, the subject will learn that NoGo leads to a neutral outcome, but, not knowing whether this is a 'Win' or 'Avoid' cue, does not know whether they are aiming to get a neutral outcome. If on the 6th trial the NoGo response is followed by a punishment or reward (because of the probabilistic feedback), the subject can suddenly deduce that the NoGo response was actually optimal (in case of punishment) or suboptimal (in case of reward), given that they know by instruction that cues can only lead to reward or punishment. Modelling this revaluation increases model evidence relative to model M1 without revaluation

of neutral outcomes (ΔWAIC=-37). Note that this revaluation adds a model-based component to our otherwise overall model-free learning system (cf. Eq.2).

$$q(a_t, s_t) = \begin{cases} Q(a_t, s_t) & if\ \{|r_{1\ldots t}|\}\ \ni\ 1 \\ 0 & else \end{cases} \quad \text{Eq. 3}$$

In M2, a Go bias parameter (*b*) was added to the action weights of Go responses to account for individual differences in the tendency to make Go responses independent of any cue. M3a then adds the influence of a Pavlovian response bias (π), where static Pavlovian values (*V*) contribute to the action weights:

$$w(a_t, s_t) = \begin{cases} q(a_t, s_t) + \pi V(s) + b & if\ a = Go \\ q(a_t, s_t) & else \end{cases} \quad \text{Eq. 4}$$

Here V=0.5 for Win cues, and V=-0.5 for Avoid cues. In this model, positive Pavlovian values increase the action weight of all Go responses, where π scales the weight of the Pavlovian values. As the cue valence can only be inferred upon the first reward or punishment, V(s)=0 beforehand. Alternatively, Pavlovian values might have been learned using delta-rule learning (Guitart-Masip et al., 2012): $V_t(s_t) = V_{t-1}(s_t) + \varepsilon\big(r_t - V_{t-1}(s_t)\big)$. However, prediction-error based learning of Pavlovian values greatly reduced model evidence (ΔWAIC=+323), and thus we proceeded with a static Pavlovian influence.

In M3b, we included an instrumental learning bias parameter (κ), to assess whether reward is more effective in reinforcing Go responses than NoGo responses, whereas punishment is less effective in unlearning NoGo responses than Go responses. In this model, κ increases the learning rate for rewarded Go responses and decreases the learning rate for punished NoGo responses:

$$\varepsilon = \begin{cases} \varepsilon_0 + \kappa & if\ r_t = 1\ \&\ a = go \\ \varepsilon_0 - \kappa & if\ r_t = -1\ \&\ a = nogo \\ \varepsilon_0 & else \end{cases} \quad \text{Eq. 5}$$

M3c included both the π and κ parameter to test whether the Pavlovian response bias and instrumental learning bias complementarily contribute to the observed motivational bias in action.

In the winning behavioural model (M3c), we tested whether midfrontal theta power covaried with the level of Pavlovian-instrumental conflict, which would be in line with the hypothesis that the midfrontal cortex is involved in the detection of motivational conflict. We computed the level of conflict as the difference in Q-values for the Pavlovian congruent versus incongruent responses. Thus, motivational conflict was computed as Q_{nogo} - mean(Q_{go}) for Win cues, and as mean(Q_{go}) - Q_{nogo} for Avoid cues.

After establishing the contribution of the cue- and outcome-based motivational biases,

we continued our computational modelling approach to assess the functional role of cue-locked midfrontal theta power in overcoming the motivational biases. We used competing models (M4a-d; cf. Cavanagh et al., 2013) to assess the potential mechanisms by which the midfrontal cortex might help to overcome the motivational biases. In these models, we quantified the impact of trial-by-trial estimates of midfrontal theta power (θ_t). The extended EEG models were estimated for all subjects that were included in the EEG analyses (i.e. excluding 4 subjects). To this end, we re-estimated the winning base model over the 30 EEG subjects, which left parameter inference unaffected, and used the resulting model evidence for model comparison in the EEG section.

In model M4a we tested whether midfrontal theta power relates to modulation of the impact of the Pavlovian bias on behavior:

$$w(Go'_t, s_t) = \begin{cases} Q(Go'_t, s_t) + (\pi + \beta * \theta_t) * V(s) + b & if\ conflict \\ Q(Go'_t, s_t) + \pi V(s) + b & else \end{cases} \quad \text{Eq. 6}$$

Here, the β parameter scales the impact of the midfrontal theta estimates on the Pavlovian response tendencies. Positive β estimates indicate that midfrontal theta power enhances the Pavlovian response tendencies, while negative β estimates reduce the Pavlovian response tendencies. Theta power only scales the impact of the Pavlovian bias on conflict trials (NoGo-to-Win, Go-to-Avoid), in line with previous report (Cavanagh et al., 2013).

In M4b we tested whether midfrontal theta power relates to modulation of the instrumental contribution. Here, the β parameter allows midfrontal theta estimates to scale the impact of the instrumental Q-values on the action weights:

$$w(Go'_t, s_t) = \begin{cases} (1 - \beta * \theta_t) * Q(Go'_t, s_t) + \pi V(s) + b & if\ conflict \\ Q(Go'_t, s_t) + \pi V(s) + b & else \end{cases} \quad \text{Eq. 7}$$

$$w(NoGo_t, s_t) = \begin{cases} (1 - \beta * \theta_t) * Q(NoGo_t, s_t) & if\ conflict \\ Q(NoGo_t, s_t) & else \end{cases}$$

In M4c we tested whether midfrontal theta power relates to modulation of the relative balance between the Pavlovian and instrumental control systems, using a Pavlovian-instrumental trade-off parameter τ instead of the Pavlovian bias parameter π. The β parameter allows midfrontal theta estimates to shift the balance between Pavlovian versus instrumental control:

$$w(Go'_t, s_t) = \begin{cases} (1 - (\tau + \beta * \theta_t)) * Q(Go'_t, s_t) + (\tau + \beta * \theta_t) * V(s) + b & if\ conflict \\ (1 - \tau) * Q(Go'_t, s_t) + \tau * V(s) + b & else \end{cases} \quad \text{Eq. 8}$$

$$w(NoGo_t, s_t) = \begin{cases} (1 - (\tau + \beta * \theta_t)) * Q(NoGo_t, s_t) & if\ conflict \\ (1 - \tau) * Q(NoGo_t, s_t) & else \end{cases}$$

The models M4a-c assess the mechanisms proposed by Cavanagh et al. (2013), and here we can optimally disentangle the Pavlovian and instrumental control systems due to the multiple

Go response options. Furthermore, in our previous work we established that motivational biases arise not only from cue-based Pavlovian mechanisms, but also from biased instrumental learning. We replicated this finding here showing the superiority of model M3c. This suggests that there is also a fourth mechanism by which midfrontal cortical control may help to respond correctly when Pavlovian and instrumental controllers conflict, namely that midfrontal theta power relates to reducing the instrumental learning bias. We test this hypothesis in model M4d, by allowing mid-frontal theta power to scale the motivational bias in learning rates:

$$\varepsilon_{rewarded\ Go} = \begin{cases} (1 - \beta * \theta_t) * (\varepsilon_0 + \kappa) + (\beta * \theta_t) * \varepsilon_0 & if\ conflict \\ \varepsilon_0 + \kappa & else \end{cases}$$

$$\varepsilon_{punished\ NoGo} = \begin{cases} (1 - \beta * \theta_t) * (\varepsilon_0 - \kappa) + (\beta * \theta_t) * \varepsilon_0 & if\ conflict \\ \varepsilon_0 - \kappa & else \end{cases} \quad \text{Eq. 9}$$

In contrast to M4a-c, here the β parameter was [0 1] constrained, such that the biased learning rates ($\varepsilon_{\text{rewarded Go}}$) and $\varepsilon_{\text{punished NoGo}}$) could regress towards the unbiased learning rate (ε_0), ensuring that the resulting learning rates could not go out of [0 1] bounds (see parameter constraints below). Accordingly, a positive β estimate indicates midfrontal theta power reduces the instrumental learning bias, rendering instrumental learning more unbiased.

Having established the computational mechanism through which midfrontal theta may reduce motivational biases (winning model M5a), we assessed whether this control might be instantiated by synchronization of the task-relevant regions to the midfrontal cortex. Current theories suggest that one role of the midfrontal cortex is to detect motivational conflict and 'alert' task-relevant regions to implement this control, through synchronization of the task-relevant regions to the midfrontal cortex (Cavanagh and Frank, 2014; Cohen and Cavanagh, 2011). These ideas suggest that perhaps the degree of synchronization with task-relevant regions, rather than local power, would be better predictors of the ability to reduce motivational biases. To this end, we assessed whether theta phase synchronization between midfrontal channels and the bilateral prefrontal channels (ISPS_{PFC}, M5a) or the contralateral motor channels ($\text{ISPS}_{\text{motor-contra}}$, M5b) scaled the Pavlovian bias, and explained behavior better than midfrontal theta power. Accordingly, we replaced θ_t in Eq. 6 with trial-by-trial phase synchrony in the theta band of midfrontal to lateral prefrontal and contralateral motor channels.

We used a sampling method for hierarchical Bayesian estimation of group-level and subject-level parameters. The group-level parameters (X) serve as priors for the individual-level parameters (x), such that $x \sim \mathcal{N}(X,\sigma)$. The hyperpriors for σ are specified by a half-Cauchy (Gelman, 2006) with a scale of 2. The hyperpriors for X are centered around 0 and weakly informative: $X_{\varepsilon,\kappa} \sim \mathcal{N}(0,2)$, $X_{\rho,b,\pi,\beta} \sim \mathcal{N}(0,3)$. All parameters are unconstrained, with the exception of ρ (positivity constraint implemented using an exponential transform), ε ([0 1] constraint implemented with

an inverse logit transform), and β in M5d ([0 1] constraint; i.e. β could suppress, but not reverse, the learning bias). To ensure that the effect of κ on ε (Eq.5) was symmetrical in model space (i.e. after inverse logit transformation to ensure [0 1] constraint), ε was computed as:

$$\varepsilon = \begin{cases} \varepsilon_0 = \text{inv.logit}(\varepsilon) & \\ \varepsilon_{\text{punished NoGo}} = \text{inv.logit}(\varepsilon - \kappa) & if\ \varepsilon_0 < .5 \\ \varepsilon_{\text{rewarded Go}} = \varepsilon_0 + (\varepsilon_0 - \varepsilon_{\text{punished NoGo}}) & if\ \varepsilon_0 < .5 \end{cases} \quad \text{Eq. 10}$$

$$\varepsilon = \begin{cases} \varepsilon_{\text{rewarded Go}} = \text{inv.logit}(\varepsilon + \kappa) & if\ \varepsilon_0 > .5 \\ \varepsilon_{\text{punished NoGo}} = \varepsilon_0 + (\varepsilon_0 - \varepsilon_{\text{rewarded Go}}) & if\ \varepsilon_0 > .5 \end{cases}$$

Model estimation was performed using Stan software in R (RStan)(Stan Development Team, 2016). Stan provides full Bayesian inference with Markov chain Monte Carlo (MCMC) sampling methods (Metropolis et al., 1953). The number of Markov chains was set at 4, with 200 burn-in iterations and 1000 post burn-in iterations per chains (4000 total). Model convergence was considered when the potential scale reduction factor $\hat{R}$ <1.1 for all parameters(Gelman and Rubin, 1992), and all models reached convergence accordingly. Model comparison was evaluated using the Watanabe-Akaike Information Criteria (WAIC)(Watanabe, 2010). WAIC is an estimate of the likelihood of the data given the model parameters, penalized for the effective number of parameters to adjust for overfitting. Lower (i.e. more negative) WAIC values indicate better model fit. As WAIC is reported on the deviance scale (Gelman et al., 2014), a difference in WAIC value of 2–6 is considered positive evidence, 6–10 strong evidence, and >10 very strong evidence (Kass and Raftery, 1995). We additionally provide a measure of explained variance (R^2) for the models, as R^2 might be considered more intuitive. However, WAIC is the most appropriate measure to compare models as WAIC penalizes for increasing model complexity. Moreover, WAIC takes into account how much variance a parameter could explain (for example, while the Pavlovian bias impacts all trials, the EEG model parameters only have an impact on the incongruent trials and can thereby explain less variance). In contrast, the R^2 values do not account for the number of parameters and the extent to which a parameter is restricted in explaining variance.

S3 Text - Reaction times.

For completeness, we analysed reaction times (RTs) as a measure of behavioural vigour. RTs were analysed with linear mixed-level models using the lme4 package in R (Bates et al., 2014; R Developement Core Team, 2015). First, we assessed RTs irrespective of accuracy, with a mixed model including the within subject factors Valence (Win vs. Avoid cue) and Required Action (Go vs. NoGo). Second, we assessed RTs for the Go cues as a function of accuracy. This mixed model included the within subject factors Valence (Win vs. Avoid cue) and Accuracy (correct vs. incorrect). RTs were ln-transformed to improve normality. Models included all main effects and interactions, and a full random effects structure (Barr, 2013; Barr et al., 2013).

The response times echoed the effects on proportion of Go responses. Learning was evidenced by shorter RTs for Go than NoGo cues (X^2_1=62.3, p<.001) and for correct relative to incorrect Go responses (X^2_1=88.3, p<.001). Motivational biasing was also evident from RTs, as subjects responded faster to Win vs. Avoid cues (X^2_1=98.6, p<.001) independent of the response requirements (X^2_1<1, p=.979) or accuracy (X^2_1=1.7, p=.198). Importantly, the effect of cue valence on RTs covaried with the effect on proportion of Go responses, such that subjects with a higher proportion of Go responses to Win cues also sped up more for Win cues ($R_{pearson}$=-.53, p=.001), suggesting that the same neural mechanisms underlying the motivational biases drove changes in both RT and choice.

S4 Text - Forced-choice transfer phase.

At the end of the learning task, subjects performed a forced choice transfer phase. In the transfer phase, cues from the last round appear on screen in pairs; subjects are requested to select the cue they found most rewarding. These explicit relative preferences provide a measure of the learned cue values. No feedback is presented at this stage, minimizing interference with the learned cue values. Cues are presented above and below the centre of the screen to be orthogonal to the left and right response requirements of the learning phase. All possible cue pairs are presented twice, with counterbalanced location, except for the pairs with cues from the same category (i.e. Go-to-Win/Go-to-Avoid/NoGo-to-Win/NoGo-to-Avoid). The transfer phase contained 48 trials in total.

All subjects indicated the Win cues as rewarding more often than the Avoid cues (t_{33}=36.5, p<.001), confirming that subjects learned the cue values. On top of that, subjects also preferentially indicated the Go cues over the NoGo cues as more rewarding (t_{33}=3.1, p=.004), even though the received outcomes did not differ significantly (t_{33}=1.7, p=.106). Thus, the preference of Go cues did not seem to reflect higher outcomes associated with these cues per se. In other words, the subjective values were boosted for Go vs. NoGo cues, which could not solely be explained by differential outcomes.

S5 Text - Feedback-related midfrontal theta power.

In the main text, we addressed modulations of feedback-related midfrontal theta power as a function of the biased instrumental learning (i.e. enhanced Go learning after reward and hampered NoGo unlearning after punishment). Here, we report a more extensive analysis of feedback-related power using a repeated measures ANOVA with the Factors Outcome (Preferred vs. Non-preferred) x Valence (Win vs. Avoid cue) x Required Response (Go vs. NoGo). Independent of the biased learning conditions, we observed clear effects of cue valence, such that feedback-related midfrontal theta power increased for Avoid relative to Win cues ($F_{1,29}$=35.6, p<.001). Additionally, the outcome effect (non-preferred vs. preferred outcomes) was stronger for the Avoid cues (Avoid cues: $F_{1,28}$=24.2, p<.001; Outcome x Valence: $F_{1,28}$=4.6, p=.042), though

also highly significant for the Win cues (Win cues: $F_{1,28}$=11.9, p=.002). Interestingly, midfrontal theta power did not significantly differ for neutral outcomes following a Win vs. Avoid cue ($F_{1,28}$=1.8, p=.189), despite their relative difference (i.e. non-preferred vs. preferred outcome). Thus, we observed well-established feedback-related modulations of midfrontal theta power (Cohen et al., 2011), yet this feedback-related theta power could not be linked to biased instrumental learning of Go/NoGo responses as reported in the main text.

S6 Text - Midfrontal-motor phase synchrony.

In the main text we reported our findings regarding midfrontal-motor phase synchrony. Specifically, we assessed whether midfrontal-motor ISPS was affected by motivational congruency, and whether this modulation depended on whether the motor site was associated with motor execution (contralateral Go) or no motor execution (ipsilateral Go or bilateral for NoGo). To elaborate, we reasoned that the motor cortex phase synchrony might behave differently for contra- and ipsilateral Go responses, as the contra- and ipsilateral motor cortex can be considered functionally different (i.e., the 'executing' vs. 'non-executing' motor site). Crucially, midfrontal functional connectivity has been proposed to be activity-dependent (Cohen, 2014), such that target sites become more responsive to the midfrontal signals when they are more active. Thus, executing and non-executing motor sites might differ in midfrontal-motor phase synchrony due to differences in task activation. To be able to assess such lateralization, we adapted the experimental setup such that subjects now needed to respond with the left and right hand, whereas in our previous study subjects responded with the index and middle finger of one hand (Swart et al., 2017). Finally, we reasoned that the motor cortex during NoGo responses would be most comparable to the ipsilateral motor cortex during Go responses, as these sites did not instantiate an overt Go response, and therefore grouped these together as the non-executing motor sites.

Following this line of reasoning, we started off with a non-directional ANOVA with the factors Valence x Required Response (Go_{contra} / Go_{ipsi} / $NoGo_{bilateral}$), and after establishing a significant Valence x Required Response interaction ($F_{2,54}$=4.0, p=.024), we continued with our planned contrast including the factors Congruence (congruent vs. incongruent) x Motor Execution (executing vs. non-executing) to assess whether the midfrontal-motor phase coherence showed a congruency effect for the executing and non-executing sites. Here we observed that ISPS with the non-executing motor increased with motivational incongruency ($F_{1,27}$=5.3, p=.029), resembling the midfrontal-lateral prefrontal phase synchrony, whereas phase synchrony with the executing-motor cortex showed, if anything, a trend in the opposite direction (Congruency x Motor Execution: $F_{1,27}$=10.2, p=.004); midfrontal-$motor_{contra}$ ISPS was marginally higher for the Go-to-Win cues than the Go-to-Avoid cues (t_{28}=1.7, p=.097). Thus, we observed that midfrontal-motor phase synchrony increased during motivational conflict for the non-executing motor sites, but, if anything, decreased for the executing (contralateral) motor sites.

The midfrontal-motor phase synchrony findings might be reconciled by considering activity-dependent functional connectivity. As mentioned above, it is thought that the midfrontal theta signals might be conveyed to all nodes in the network, whereas only active areas become susceptible to the midfrontal theta-band input (Cohen, 2014), giving rise to activity-dependent functional connectivity. As such, we do not need to assume that the midfrontal cortex "knows" to which areas to send the information in order to implement control. Accordingly, control-related functional connectivity has been observed between the midfrontal cortex and several target sites, depending on the task at hand (Cavanagh and Frank, 2014). Consistent with these theories, we also observed a motivational conflict-related increase in phase synchrony between the midfrontal and lateral prefrontal sites, and between the midfrontal and motor sites that did not instantiate an overt motor response. Surprisingly, however, we observed the opposite (non-significant) pattern for the motor sites contralateral to active 'Go' responses. Phase synchrony between the midfrontal and contralateral motor sites was marginally stronger in reward than punishment contexts, even though active 'Go' responding can be considered motivationally congruent with reward contexts. Although we acknowledge that this effect for the executing motor sites was not significant (and we should therefore be cautious to interpret this effect), we would nevertheless like to provide the following potential explanation for this surprising observation; That is, these findings could be reconciled by considering that reward cues facilitate behavioral activation (Dayan et al., 2006; Guitart-Masip et al., 2014a; Niv et al., 2007) through activation of the basal ganglia 'Go'-pathways (Collins and Frank, 2014; Hernandez-Lopez et al., 2000, 1997), increasing activation in the motor cortex (Chiu et al., 2014). The enhanced (reward-related) activation might have made the contralateral motor cortex more susceptible to the midfrontal theta-band signals, even though the midfrontal sites displayed weaker theta-band activation during the Go-to-Win trials. Of course such reward-related activation might also hold for the non-executing motor sites, but could nevertheless be most pronounced for the motor cortex that instantiates an overt motor response. To summarize, midfrontal-motor phase synchrony significantly increased with motivational conflict in the non-executing motor sites, but marginally decreased in the contralateral motor sites, potentially due to enhanced susceptibility for midfrontal connectivity resulting from reward-related activation of the contralateral motor cortex.

S7 Text - Single trial power and intersite phase synchrony.

In the main text, we showed that motivational conflict modulated trial-averaged midfrontal theta power and intersite phase synchrony with motor and prefrontal sites. Here we assessed whether the single-trial estimates could capture the main conflict modulations as observed at the trial-averaged level, before using the single-trial estimates in the computational models. Single-trial power and intersite phase synchrony (ISPS) estimates were analysed using mixed-model regression analyses. These models included the within subject factors Valence (Win vs.

Avoid) and Required Action (Go vs. NoGo) for the power and $ISPS_{midfrontal\text{-}prefrontal}$ analysis. For the $ISPS_{midfrontal\text{-}motor}$ analysis, we analysed the conflict modulation again as a function of executing vs. non-executing motor side. Accordingly, the model included the within subject factors Congruency (congruent vs. incongruent) and Motor Execution (executing vs. non-executing), where the effect of Congruency translates to the Valence x Required Action interaction in the other models. All models included a full random effects structure; the power and ISPS values were mean-corrected and inverse transformed to improve normality.

Results of the single-trial analysis of midfrontal theta power were in line with the trial-averaged analyses. The single-trial power values increased under motivational conflict (X^2_1=8.7, p=.003), in the absence of main effects of Valence (X^2_1=2.7, p=.098) and Required Action (X^2_1=1.3, p=.971). The modulation by valence was significant for the NoGo cues (X^2_1=11.3, p<.001), and not for the Go cues (X^2_1=.9, p=.348).

In the main analyses we excluded incorrect trials to minimize contamination by error-related signals, since both conflict detection and error processing have been linked to increased midfrontal theta power (Cavanagh and Frank, 2014; Cohen, 2014). The conflict detection signals commonly peak prior to the response, while the error processing signals peak after the response (Cohen, 2014). Although our time-window of interest (450-650ms) precedes the average response time (mean=753ms; range=594-1077ms), the power estimates might capture error processing signals nonetheless due to the temporal smoothing that is inherent in time-frequency analyses. We tested whether theta power increased on error trials, but observed the opposite pattern: the midfrontal theta power values increased for the correct trials relative to incorrect trials (X^2_1=17.1, p<.001), speaking against the contamination by error processing. Furthermore, motivational conflict did not significantly modulate the power estimates on incorrect trials (X^2_1=.2, p=.685), suggesting that midfrontal theta power particularly increases with motivational conflict when the conflict is correctly overcome. Note, however, that a mixed-level model with the full interaction (Valence x Required Action x Accuracy) did not converge. This is likely due to the low error rate particularly on congruent trials, leaving this interaction underpowered. To nonetheless assess this interaction, we circumvent this convergence issue by testing this interaction on the trial-averaged data using repeated measures ANOVA. This ANOVA indicated a significant three-way interaction (Valence x Required Action x Accuracy*:* $F_{1,26}$=6.6, p=.016), driven by a significant interaction for the correct trials (Valence x Required Action: $F_{1,29}$=9.6, p=.004) and a non-significant interaction for the incorrect trials (Valence x Required Action: $F_{1,26}$=1.2, p=.286). Three subjects lacked EEG data for the incorrect NoGo-to-Avoid trials and were therefore not included in this analysis. Given the low error rate on the motivationally congruent trials for some subjects, these results should be interpreted with caution. Taken together, the single-trial power estimates showed the modulation by motivational conflict, particularly when the conflict was successfully overcome, and did not show an error-related increase. This pattern of results could suggest that subjects might not

have detected the motivational conflict on the incorrect trials, leaving them to follow their prepotent, Pavlovian response tendencies. Therefore, we included the single-trial power estimates of both correct and incorrect trials in the computational modelling section.

Motivational conflict also modulated midfrontal-prefrontal phase synchrony at the single-trial level, such that the phase synchrony estimates increased with motivational conflict (X^2_1=5.4, p=.021). The phase synchrony modulation by motivational conflict showed only a significant simple effect for the NoGo cues (X^2_1=8.5, p=.003), and not for the Go cues (X^2_1=.3, p=.608). No other effects were significant (Valence: X^2_1=3.5, p=.062; Required Action: X^2_1<.1, p=.886).

The single-trial midfrontal-motor ISPS results again showed an effect of motivational conflict that depended on whether the motor site instantiated an overt Go response (Congruency x Motor Execution: X^2_1=4.1, p=.044). Thus, the midfrontal-motor ISPS increased with motivational conflict in the non-executing motor site (i.e. ipsilateral to Go responses and bilateral for NoGo responses), but decreased with motivational conflict in the executing motor site (i.e. contralateral to Go responses), although both simple effects were non-significant (non-executing: X^2_1=1.8, p=.176; executing: X^2_1=2.0, p=.158). Furthermore, single-trial midfrontal-motor phase synchrony increased for Win relative to Avoid cues across responses (X^2_1=4.1, p=.042). Single-trial midfrontal-motor phase synchrony did not significantly differ between the motor sites during NoGo responses and contralateral Go responses (X^2_1=1.0, p=.329), or ipsilateral Go responses (X^2_1=.2, p=.679). Altogether, the single-trial estimates for both power and intersite phase synchrony replicate the modulations by motivational conflict as observed at the trial-averaged level.

As we include the power and phase synchrony measures in competing computational models (family M4 and M5 respectively), we assessed how correlated the trial-by-trial estimates are, in order to assess whether there is sufficient unexplained variance between these measures. To this end, we computed the within-subject correlations between trial-by-trial midfrontal theta power and i) midfrontal-lateral prefrontal phase synchrony, and ii) midfrontal-motor phase synchrony per condition (see Fig S1). The average correlation between midfrontal theta power and midfrontal-lateral prefrontal phase synchrony was .38 (range: .22 to .62), and the average correlation for midfrontal theta power and midfrontal-motor phase synchrony was .32 (range: .12 to .53). Thus, even for the highest observed correlation (R=+.62; R^2=.38), there was still 62% unexplained variance between the power and phase synchrony measures. In other words, despite clear covariation between the power and phase synchrony measures, there was extensive unique variance.

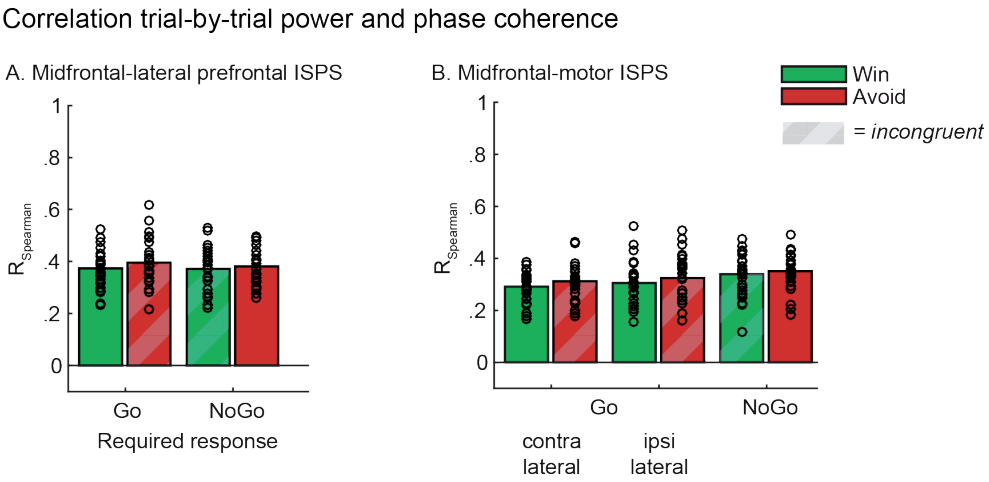


Fig S1. **Correlations between single-trial midfrontal theta power and intersite phase synchrony.** (**a**) Correlations for trial-by-trial midfrontal theta power and midfrontal-lateral prefrontal phase synchrony per subject (black circles), ranging from +.22 to +.62, where the lowest observation of unexplained variance still reaches 62%. (**b**) Correlations for trial-by-trial midfrontal theta power and midfrontal-motor phase synchrony per subject, ranging from +.12 to +.53.

S8 Text - Computational modelling: potential alternative mechanisms related to intersite phase synchrony.

In the analyses described in the main text, we assumed that the phase synchronization of the task-relevant clusters to the midfrontal cluster would impact the same system as local midfrontal theta power (M5). Alternatively, the phase synchronization might have target-dependent impact, where the lateral prefrontal synchrony might relate to the modulation of the goal representations (in this task, the instrumental action values), whereas midfrontal-motor synchrony might relate to modulation of the motor excitability. We tested these alternative mechanisms in a new set of models, M6. We used the same intersite synchrony measures, but now allowed the midfrontal-prefrontal synchrony to scale the impact of the instrumental controller (M6a; cf. Eq. 7) and midfrontal-motor synchrony to scale the contralateral action weight (M6b):

$$w(Go'_t, s_t) = \begin{cases} Q(Go'_t, s_t) + \pi V(s) + b + \beta * ISPS_{motor,contra} & if\ conflict \\ Q(Go'_t, s_t) + \pi V(s) + b & else \end{cases}$$

Model evidence reduced relative to the M5 models for these alternative synchrony models, where midfrontal-lateral prefrontal synchrony modulated the instrumental contribution ($WAIC_{M6a}$=28615), and midfrontal-motor synchrony modulated motor excitability ($WAIC_{M6b}$=28557). Altogether, model comparison favoured synchrony models where midfrontal-lateral prefrontal ($WAIC_{M5a}$=28477) and midfrontal-motor phase synchrony ($WAIC_{M5b}$=28425) modulated the Pavlovian bias. These findings are in line with the proposal that the midfrontal cortex signals the need to adjust the decision-threshold to the task-related network, in order to prevent impulsive, Pavlovian responses.

S1 Table. **Median [25–75 percentile] of subject-level parameter estimates in model space.**

	Behavioural models				
	M1	**M2**	**M3a**	**M3b**	**M3c**
ρ	9.9 [3.3 48.3]	8.9 [3.6 39.4]	7.9 [2.7 39.7]	6.1 [3.8 24.2]	6.9 [2.7 32.8]
ε_0	.03 [.01 .20]	.05 [.01 .19]	.06 [.02 .20]	.08 [.02 .18]	.07 [.02 .19]
b		.3 [.0 .5]	.3 [.0 .6]	.5 [.2 .7]	.5 [.1 .7]
π			.6 [-.2 1.2]		.4 [-.5 .9]
τ					
$\varepsilon_{\text{rewarded Go}}$ (ε_0+κ)				.55 [.11 .88]	.17 [.07 .49]
$\varepsilon_{\text{punished NoGo}}$ (ε_0-κ)				.005 [.002 .028]	.01 [.01 .07]
β					

	Extended EEG models					
	M4a	**M4b**	**M4c**	**M4d**	**M5a**	**M5b**
ρ	10.2 [3.3 40.2]	11.1 [3.8 40.0]	19.9 [6.9 40.2]	9.6 [3.4 35.3]	10.0 [3.3 39.2]	10.2 [3.4 40.8]
ε_0	.04 [.02 .19]	.05 [.02 .20]	.05 [.02 .20]	.06 [.02 .19]	.04 [.02 .18]	.04 [.02 .19]
b	.4 [.1 .7]	.4 [.1 .7]	.4 [.2 .7]	.4 [.1 .7]	.4 [.1 .7]	.4 [.1 .7]
π	.9 [-.0 1.3]	.3 [-.3 1.0]		.4 [-.4 .8]	.9 [-.0 1.3]	.9 [-.1 1.3]
τ			.2 [-.3 .5]			
$\varepsilon_{\text{rewarded Go}}$ (ε_0+κ)	.20 [.11 .64]	.17 [.06 .51]	.31 [.15 .68]	.35 [.18 .76]	.20 [.11 .65]	.19 [.10 .68]
$\varepsilon_{\text{punished NoGo}}$ (ε_0-κ)	.006 [.002 .047]	.012 [.005 .060]	.006 [.002 .024]	.004 [.002 .024]	.006 [.002 .052]	.006 [.002 .053]
β	-2.2 [-3.3 -.6]	-.2 [-.7 .4]	-.0 [-.2 .2]	.2 [.1 .4]	-1.4 [-2.4 -.3]	-1.3 [-2.7 -.4]

S2 Table. **Confidence of group-level parameter estimates (% of samples>0).**

	M1	M2	M3a	M3b	M3c	M4a	M4b	M4c	M4d	M5a	M5b
ρ	100	100	100	100	100	100	100	100	100	100	100
ε_0	0	0	0	0	0	0	0	0	0	0	0
b		99.9	99.8	100	100	100	100	100	100	99.9	100
π			99.9		95.5	99.5	92.6		92.7	99.7	99.8
τ								76			
κ				100	100	100	100	100	100	100	100
β						0	10.6	22.6	84.1	0	0

References

Aron, A.R., Herz, D.M., Brown, P., Forstmann, B.U., Zaghloul, K., 2016. Frontosubthalamic Circuits for Control of Action and Cognition. J. Neurosci. 36, 11489–11495. doi:10.1523/JNEUROSCI.2348-16.2016

Barr, D.J., 2013. Random effects structure for testing interactions in linear mixed-effects models. Front. Psychol. 4, 328. doi:10.3389/fpsyg.2013.00328

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4. R Packag. version. doi:citeulike-article-id:7112638

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. Psychol Rev. 108, 624–652. doi:10.1037/0033-295X.108.3.624

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D., 1998. Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. Science (80-.). 280, 747–749. doi:10.1126/science.280.5364.747

Cavanagh, J.F., Cohen, M.X., Allen, J.J.B., 2009. Prelude to and Resolution of an Error: EEG Phase Synchrony Reveals Cognitive Control Dynamics during Action Monitoring. J. Neurosci. 29, 98–105. doi:10.1523/JNEUROSCI.4137-08.2009

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., Frank, M.J., 2013. Frontal Theta Overrides Pavlovian Learning Biases. J. Neurosci. 33, 8541–8548. doi:10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. doi:10.1016/j.tics.2014.04.012

Cavanagh, J.F., Masters, S.E., Bath, K., Frank, M.J., 2014. Conflict acts as an implicit cost in reinforcement learning. Nat. Commun. 5, 5394. doi:10.1038/ncomms6394

Cavanagh, J.F., Wiecki, T. V, Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., Frank, M.J., 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat. Neurosci. 14, 1462–1467. doi:10.1038/nn.2925

Cavanagh, J.F., Zambrano-Vazquez, L., Allen, J.J.B., 2012. Theta lingua franca: A common mid-frontal substrate for action monitoring processes. Psychophysiology 49, 220–238. doi:10.1111/j.1469-8986.2011.01293.x

Chiu, Y.-C., Cools, R., Aron, A.R., 2014. Opposing Effects of Appetitive and Aversive Cues on Go/No-go Behavior and Motor Excitability. J. Cogn. Neurosci. 26, 1851–1860. doi:10.1162/jocn_a_00585

Cohen, M.X., 2014. A neural microcircuit for cognitive conflict detection and signaling. Trends Neurosci. doi:10.1016/j.tins.2014.06.004

Cohen, M.X., 2011. Error-related medial frontal theta activity predicts cingulate-related structural connectivity. Neuroimage 55, 1373–1383. doi:10.1016/j.neuroimage.2010.12.072

Cohen, M.X., Cavanagh, J.F., 2011. Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. Front. Psychol. 2. doi:10.3389/fpsyg.2011.00030

Cohen, M.X., Ridderinkhof, K.R., 2013. EEG Source Reconstruction Reveals Frontal-Parietal Dynamics of Spatial Conflict Processing. PLoS One 8. doi:10.1371/journal.pone.0057293

Cohen, M.X., Wilmes, K., van de Vijver, I., 2011. Cortical electrophysiological network dynamics of feedback learning. Trends Cogn. Sci. doi:10.1016/j.tics.2011.10.004

Collins, A.G.E., Frank, M.J., 2014. Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. 121, 337–66. doi:10.1037/a0037015

Davis, H., Wright, J., 1979. procedural and parametric variability in studies of conditioned suppression. Bull. Psychon. Soc. 14, 149–150.

Dayan, P., Niv, Y., Seymour, B., D. Daw, N., 2006. The misbehavior of value and the discipline of the will. Neural Networks 19, 1153–1160. doi:10.1016/j.neunet.2006.03.002

Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18. doi:10.3758/BF03199951

Duffy, E., 1962. Activation and behavior. New York.

Estes, W.K., 1943. Discriminative conditioning. I. A discriminative property of conditioned anticipation. J. Exp. Psychol. 32, 150–155. doi:10.1037/h0058316

Estes, W.K., Skinner, B.F., 1941. Some quantitative properties of anxiety. J. Exp. Psychol. 29, 390–400. doi:10.1037/h0062283

Frank, M.J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J.F., Badre, D., 2015. fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. J. Neurosci. 35, 485–494. doi:10.1523/JNEUROSCI.2036-14.2015

Fries, P., 2015. Rhythms for Cognition: Communication through Coherence. Neuron. doi:10.1016/j.neuron.2015.09.034

Gelman, A., 2006. Prior distribution for variance parameters in hierarchical models. Bayesian Anal. 1, 515–533. doi:10.1214/06-BA117A

Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016. doi:10.1007/s11222-013-9416-2

Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. doi:10.1214/ss/1177011136

Geurts, D.E.M., Huys, Q.J.M., Den Ouden, H.E.M., Cools, R., 2013. Aversive Pavlovian control of instrumental behavior in humans. J. Cogn. Neurosci. 25, 1428–1441.

Guitart-Masip, M., Duzel, E., Dolan, R., Dayan, P., 2014. Action versus valence in decision making. Trends Cogn. Sci. doi:10.1016/j.tics.2014.01.003

Guitart-Masip, M., Huys, Q.J.M., Fuentemilla, L., Dayan, P., Duzel, E., Dolan, R.J., 2012. Go and no-go learning in reward and punishment: Interactions between affect and effect. Neuroimage 62, 154–166. doi:10.1016/j.neuroimage.2012.04.024

Hauser, T.U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., Brem, S., 2014. The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. Neuroimage 84, 159–168. doi:10.1016/j.neuroimage.2013.08.028

Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A., Galarraga, E., 1997. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. J Neurosci 17, 3334–3342.

Hernandez-Lopez, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., Surmeier, D.J., 2000. D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLC[beta]1-IP3-calcineurin-signaling cascade. J. Neurosci. 20, 8987–8995. doi:20/24/8987 [pii]

Hershberger, W.A., 1986. An approach through the looking-glass. Anim. Learn. Behav. 14, 443–451. doi:10.3758/BF03200092

Herz, D.M., Zavala, B.A., Bogacz, R., Brown, P., 2016. Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus. Curr. Biol. 26, 916–920. doi:10.1016/j.cub.2016.01.051

Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109, 679–709. doi:10.1037/0033-295X.109.4.679

Huys, Q.J.M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R.J., Dayan, P., 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. PLoS Comput. Biol. 7. doi:10.1371/journal.pcbi.1002028

Kass, R., Raftery, A., 1995. Bayes Factors. J. Am. Stat. Assoc. 90, 773–795. doi:10.1080/01621459.1995.10476572

Kelley, R., Flouty, O., Emmons, E.B., Kim, Y., Kingyon, J., Wessel, J.R., Oya, H., Greenlee, J.D., Narayanan, N.S., 2018. A human prefrontal-subthalamic circuit for cognitive control. Brain 141, 205–216. doi:10.1093/brain/awx300

Marco-Pallares, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T.F., Rodríguez-Fornells, A., 2008. Human oscillatory activity associated to reward processing in a gambling task. Neuropsychologia 46, 241–248. doi:10.1016/j.neuropsychologia.2007.07.016

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092. doi:http://dx.doi.org/10.1063/1.1699114

Nigbur, R., Cohen, M.X., Ridderinkhof, K.R., Stürmer, B., 2012. Theta Dynamics Reveal Domain-specific Control over Stimulus and Response Conflict. J. Cogn. Neurosci. 24, 1264–1274. doi:10.1162/jocn_a_00128

Nigbur, R., Ivanova, G., Stürmer, B., 2011. Theta power as a marker for cognitive interference. Clin. Neurophysiol. 122, 2185–2194. doi:10.1016/j.clinph.2011.03.030

Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: Opportunity costs and the control of response vigor. Psychopharmacology (Berl). 191, 507–520. doi:10.1007/s00213-006-0502-4

Oostendorp, T.F., Van Oosterom, A., 1996. The surface Laplacian of the potential: Theory and application. IEEE Trans. Biomed. Eng. 43, 394–405. doi:10.1109/10.486259

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011. doi:10.1155/2011/156869

Pastötter, B., Dreisbach, G., Bäuml, K.T., 2013. Dynamic adjustments of cognitive control: oscillatory correlates of the conflict adaptation effect. J. Cogn. Neurosci. 25, 2167–78. doi:10.1162/jocn_a_00474

Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72, 184–187. doi:10.1016/0013-4694(89)90180-6

R Developement Core Team, 2015. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. 1, 409. doi:10.1007/978-3-540-74686-7

Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The expected value of control: An integrative theory of anterior cingulate cortex function. Neuron. doi:10.1016/j.neuron.2013.07.007

Siegel, M., Donner, T.H., Engel, A.K., 2012. Spectral fingerprints of large-scale neuronal interactions. Nat. Rev. Neurosci. 13, 20–25. doi:10.1038/nrn3137

Simon, J.R., Rudell, A.P., 1967. Auditory S-R compatibility: the effect of an irrelevant cue on information processing. J. Appl. Psychol. 51, 300–304. doi:10.1037/h0020586

Srinivasan, R., Winter, W.R., Ding, J., Nunez, P.L., 2007. EEG and MEG coherence: Measures of functional connectivity at distinct spatial scales of neocortical dynamics. J. Neurosci. Methods 166, 41–52. doi:10.1016/j.jneumeth.2007.06.026

Stan Development Team, 2016. RStan: the R interface to Stan, Version 2.9.0. [WWW Document]. URL http://mc-stan.org

Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. IEEE Trans. Neural Networks 9, 1054–1054. doi:10.1109/TNN.1998.712192

Swart, J.C., Froböse, M.I., Cook, J.L., Geurts, D.E.M., Frank, M.J., Cools, R., den Ouden, H.E.M., 2017. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. Elife 6. doi:10.7554/eLife.22169

Van de Vijver, I., Cohen, M.X., Ridderinkhof, K.R., 2014. Aging affects medial but not anterior frontal learning-related theta oscillations. Neurobiol. Aging 35, 692–704. doi:10.1016/j.neurobiolaging.2013.09.006

van de Vijver, I., Ridderinkhof, K.R., Cohen, M.X., 2011. Frontal oscillatory dynamics predict feedback learning and action adjustment. J. Cogn. Neurosci. 23, 4106–4121. doi:10.1162/jocn_a_00110

van Driel, J., Ridderinkhof, K.R., Cohen, M.X., 2012. Not All Errors Are Alike: Theta and Alpha EEG Dynamics Relate to Differences in Error-Processing Dynamics. J. Neurosci. 32, 16795–16806. doi:10.1523/JNEUROSCI.0802-12.2012

van Driel, J., Swart, J.C., Egner, T., Ridderinkhof, K.R., Cohen, M.X., 2015. (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. Cogn. Affect. Behav. Neurosci. 15, 787–807. doi:10.3758/s13415-015-0367-2

Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. J. Mach. Learn. Res. 11, 3571–3594.

Wiecki, T. V., Frank, M.J., 2013. A computational model of inhibitory control in frontal cortex and basal ganglia. Psychol. Rev. 120, 329–355. doi:10.1037/a0031542

Zavala, B.A., Tan, H., Little, S., Ashkan, K., Hariz, M., Foltynie, T., Zrinzo, L., Zaghloul, K.A., Brown, P., 2014. Midline Frontal Cortex Low-Frequency Activity Drives Subthalamic Nucleus Oscillations during Conflict. J. Neurosci. 34, 7322–7333. doi:10.1523/JNEUROSCI.1169-14.2014

Chapter 5

Deep brain stimulation of the nucleus accumbens releases aversive inhibition of behaviour in obsessive-compulsive disorder

I count him braver who overcomes his desires than who conquers his enemies;
for the hardest victory is over self.
- Aristotle

Contributing authors:
Swart JC, Gonzales-Rosa JJ, Lopez-Sosa F, Treu S, Barcia JA, Cools R, Strange B, & den Ouden HEM.

Abstract

The striatum has long been implicated in motivation, learning, and action. Recent theories have proposed that the striatum is particularly involved in the coupling of these processes, giving rise to motivational biases in the selection and learning of actions. Here we assess the causal role of the ventral striatum in the motivational biasing of action for the first time in humans by directly stimulating the nucleus accumbens with deep brain stimulation (DBS). Treatment-refractory obsessive-compulsive disorder patients (n=8) performed a motivational Go/NoGo task with concurrent EEG recordings. The subjects needed to learn to make Go or NoGo responses in order to win reward or avoid punishment, while DBS was switched ON vs. OFF in a cross-over within-subject design. As previously observed in healthy populations, performance was strongly affected by the cue valence, such that fewer and slower Go responses were made when avoiding punishment than when playing for reward. DBS attenuated the inhibitory influence of punishment cues on reaction times and marginally on the proportion Go responses when Go responses were required. For the Go cues where these Pavlovian response tendencies conflicted with the instrumental requirements, oscillatory theta (4-8Hz) activity increased over the midfrontal cortex. These putative conflict-related midfrontal theta responses were not significantly affected by DBS. Taken together, these results suggest that nucleus accumbens stimulation attenuates the motivational biasing of action, seemingly without affecting frontal control systems. These results causally implicate the human nucleus accumbens in the coupling of motivation and behavioural activation.

Introduction

The striatum has widely been implicated in a range of cognitive processes, namely motivation, learning, and action (Berridge and Robinson, 1998; Robbins and Everitt, 2007, 1996; Salamone et al., 2005). Neuropsychiatric disorders involving these processes have often been linked to altered striatal functioning, for example obsessive-compulsive disorder (OCD; Burguière et al., 2015), Parkinson's disease (Holthoff-Detto, 1997), and addiction (Pujara and Koenigs, 2014). In turn, treatment of these disorders regularly targets striatal functioning (Ahmari and Dougherty, 2015; MacDonald et al., 2011; Stoessl, 2008). Recent theories have proposed that the striatum is particularly involved in the coupling of motivation, learning, and action, giving rise to motivational biases in the selection and learning of actions (Collins and Frank, 2014; Guitart-Masip et al., 2014a; Swart et al., 2017). Here, we set out to assess the causal role of the human ventral striatum in the motivational biasing of action by direct stimulation of the nucleus accumbens.

To elaborate, anticipated rewards tend to facilitate taking action, whereas anticipated losses tend to facilitate holding back (Dickinson and Balleine, 1994; Guitart-Masip et al., 2014a; Huys et al., 2011). Similar biases have also been observed for the learning of (in)action based on reward and punishment outcomes (Swart et al., 2018, 2017). These motivational biases in

action have been theorized to arise through dopamine function in the basal ganglia (Collins and Frank, 2014), where particularly the striatum might subserve the motivational biases via the direct and indirect dopaminergic pathways (Dickinson et al., 2000; Hebart and Gläscher, 2015; Lex and Hauber, 2008; Taylor and Robbins, 1986, 1984; Wyvell and Berridge, 2000). In short, the mesolimbic peaks in striatal dopamine release, elicited by reward cues (Cohen et al., 2012; Day et al., 2007; Matsumoto and Hikosaka, 2009; Tobler et al., 2005), potentiate the direct D1 ('Go') pathway (Hernandez-Lopez et al., 1997), thereby promoting behavioural activation (DeLong and Wichmann, 2007; Mink and Thach, 1991). Conversely, the mesolimbic dips in striatal dopamine release, elicited by punishment cues, are thought to potentiate the indirect D2 ('NoGo') pathway (Hernandez-Lopez et al., 2000), thereby promoting behavioural inhibition. In other words, the input to the striatum, which depends on the motivational valence, putatively modulates activity in the (in)direct pathways and in that way biases behavioural activation.

The striatal coupling of the valenced input to the action-modulating output leads to our first prediction that dissociating the striatal in- and output will attenuate the motivational biases in action. We will test this hypothesis using deep brain stimulation (DBS). The exact mechanisms of action of DBS are still highly debated and likely involve a complex interplay of both inhibitory and excitatory effects, both locally and network-wide (Chiken and Nambu, 2016). Crucially, however, is that DBS is thought to disrupt the ongoing neural communication within the target site by dissociating the neural input and output (Chiken and Nambu, 2016). Based on this disruptive effect, we hypothesized that DBS will reduce the coupling of motivationally-driven input and action-related output in the nucleus accumbens, thereby reducing the motivational biasing of behavioural activation.

The motivational biases have been suggested to reflect the statistics of our environment and therefore to be beneficial by reducing computational load (Dayan et al., 2006). Yet, at times these biases can conflict with the behaviour required by our current goals, for example during delayed gratification or active avoidance. At those times, such motivational conflict needs to be detected and signalled to the task-relevant regions in order to facilitate goal-directed behaviour (Cohen, 2014). The midfrontal cortex has been linked to the detection of motivational conflict (Cavanagh et al., 2013; Swart et al., 2018), as reflected by increased midfrontal oscillatory activity in the theta frequency range (4-8Hz)(Cavanagh et al., 2013; Swart et al., 2018). Given that midfrontal oscillatory activity also covaries with continuous measures of conflict (Cohen and Cavanagh, 2011; Cohen and Donner, 2013; Swart et al., 2018), we secondly predicted that an attenuation of the motivational biases under DBS will consequently result in reduced conflict-related neural responses.

In this study, we assessed the effects of nucleus accumbens stimulation on the well-established motivational biases in action and on corresponding conflict-related neural signatures. To this end, we employed DBS in the bilateral nucleus accumbens of 8 OCD patients who receive DBS as part of their treatment. The participants performed a motivational Go/NoGo learning

task with alternating DBS ON/OFF and concurrent surface EEG recordings. We set out to test the hypotheses that DBS of the nucleus accumbens i) attenuates the motivational biases in behavioural activation and ii) consequently attenuates the conflict-related midfrontal theta responses.

Materials and methods

Subjects

We assessed 8 native Spanish patients (aged 21-50 years; 3 women; 7 right-handed) with bilateral deep brain electrodes (Medtronic 3391) situated in the nucleus accumbens. The two most distal contacts of each electrode were positioned within the posterior part of the nucleus accumbens (also see Stimulation procedure) with the aid of a preoperative MR scan fused with stereotactic frame-based CT imaging using standard clinical procedures (Nachev et al., 2015). The patients received DBS as treatment for drug- and therapy-resistant obsessive-compulsive disorder and were assessed during their routine check-ups in the hospital. Our sample was limited by the number of chronic OCD-DBS patients under treatment at Hospital Clínico San Carlos. See Supplementary Table 1 for an overview of the clinical evaluations and demographics per patient. The patients had normal or corrected to normal vision and did not report colour blindness. All patients gave informed consent, performed the motivational Go/NoGo learning task (see below) and a classical Go/NoGo task (to be published elsewhere), and received a reimbursement of EUR30,- upon completion of the study. The study was approved by the local ethical committee (CEIC Hospital Clínico San Carlos 10/131). During the completion of this thesis chapter, we assessed one more patient and the data of this patient will be included in our final publication.

Motivational Go/NoGo learning task

We employed a motivational Go/NoGo learning task (Figure 1) (cf. Guitart-Masip et al., 2011; Swart et al., 2017) where the required instrumental response (Go vs. NoGo) is orthogonal to the cue valence (Win vs. Avoid). Accordingly, there are 4 cue types in total (Go-to-Win, Go-to-Avoid, NoGo-to-Win, and NoGo-to-Avoid). Each trial starts with the presentation of a Win (green edge) or Avoid (red edge) cue. Subjects can press the spacebar (Go) or not (NoGo) during cue presentation and response-dependent feedback follows after a 0.5s fixation interval. Correct responses are followed by reward ('+100'; Win cues) or neutral outcome ('000'; Avoid cues) 70% of the time, and by neutral outcome (Win cues) and punishment ('-100'; Avoid cues) otherwise. These probabilities are reversed for incorrect responses. Go responses are considered correct for the Go-to-Win and Go-to-Avoid cue, whereas NoGo responses are considered correct for the NoGo-to-Win and NoGo-to-Avoid cue. Subjects can learn the correct responses by trial-and-error based on the feedback. Trials end with a fixation inter-trial interval (ITI) ranging from 1.25 to 2s in steps of 0.25s, randomly sequenced.

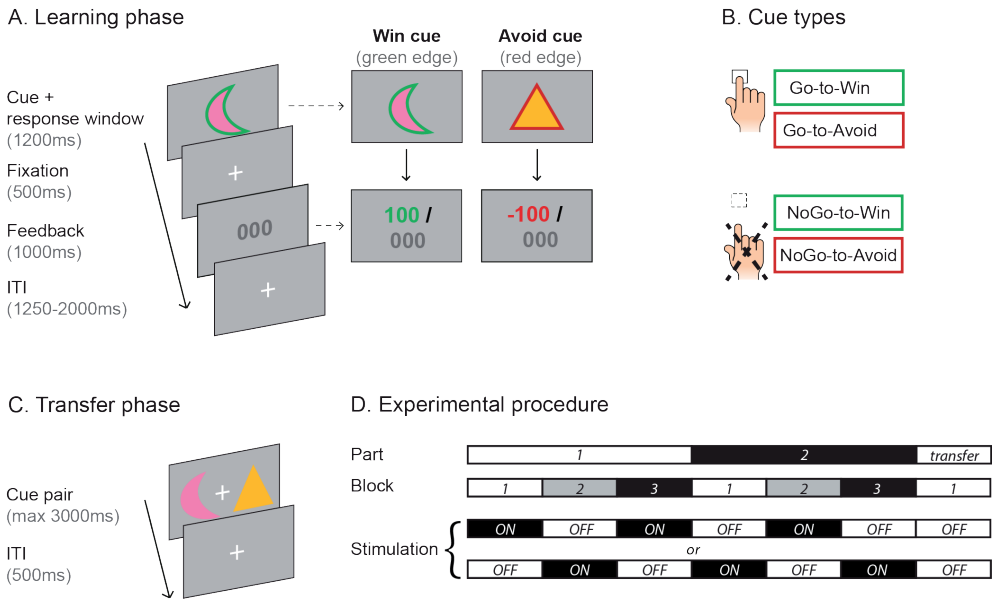


Figure 1. **Motivational Go/NoGo task.**
(**a**) Learning phase. Each trial starts with a Win (green edge) or Avoid (red edge) cue. During cue presentation, subjects can make a Go (press spacebar) or NoGo (not press) response. Correct responses are followed by reward (Win cues) and a neutral outcome (Avoid cues) in 70% of the times, and by a neutral outcome (Win cues) or punishment (Avoid cues) otherwise. For incorrect responses, these probabilities are reversed. Subjects should learn the correct responses by trial and error. Image adapted from (Swart et al., 2017). (**b**) The task contains 4 cue types; the Go-to-Win and Go-to-Avoid, for which the Go response is correct, and the NoGo-to-Win and NoGo-to-Avoid, for which the NoGo response is correct. (**c**) Transfer phase. Cues are presented in pairs and subjects are asked to select the most rewarding cue. (**d**) The learning phase was performed twice, using independent stimulus sets. Each part was divided into 3 blocks with self-paced breaks in between. Stimulation was interleaved ON and OFF, counter-balanced over parts and subjects.

Subjects perform the task twice, with a new stimulus set for the second part. All stimuli have unique shapes and colours well distinguishable from the red and green edge. The stimuli colours and shapes are randomly assigned to the 4 different cue types. Each cue is presented 60 times in each part, resulting in a total of 480 trials. The order of cue presentation is pseudorandom, with cues repeated once at most. The main task lasted approximately 40 minutes. Before start of the task, subjects are informed that i) each cue has one optimal response, ii) feedback is probabilistic, and iii) their wins and losses will be converted to a monetary bonus at the end of the experiment. Thus, we suggested a performance-dependent bonus to incentivize the patients, yet we disbursed a fixed amount of EUR30,- for ethical considerations (e.g., to prevent that more severely affected patients would receive a smaller reimbursement).

Upon completion of the main task, a short transfer phase follows (Cavanagh et al., 2013); cues from the second part are presented in pairs and subjects are asked to select the most

rewarding cue. Cues are paired with all other cues and all cue combinations are presented 8 times, resulting in a total of 48 trials. Importantly, the cues are presented without the coloured edge to measure the learned stimulus values and minimize interference by the explicit cue valences. The cue pair remains on screen until the response, with a maximum of 3s, followed by an ITI of 0.5s. If the subject responds too slowly, a message indicates to respond faster and the missed trial is repeated at the end. The transfer phase takes approximately 3 minutes in total.

Stimulation procedure

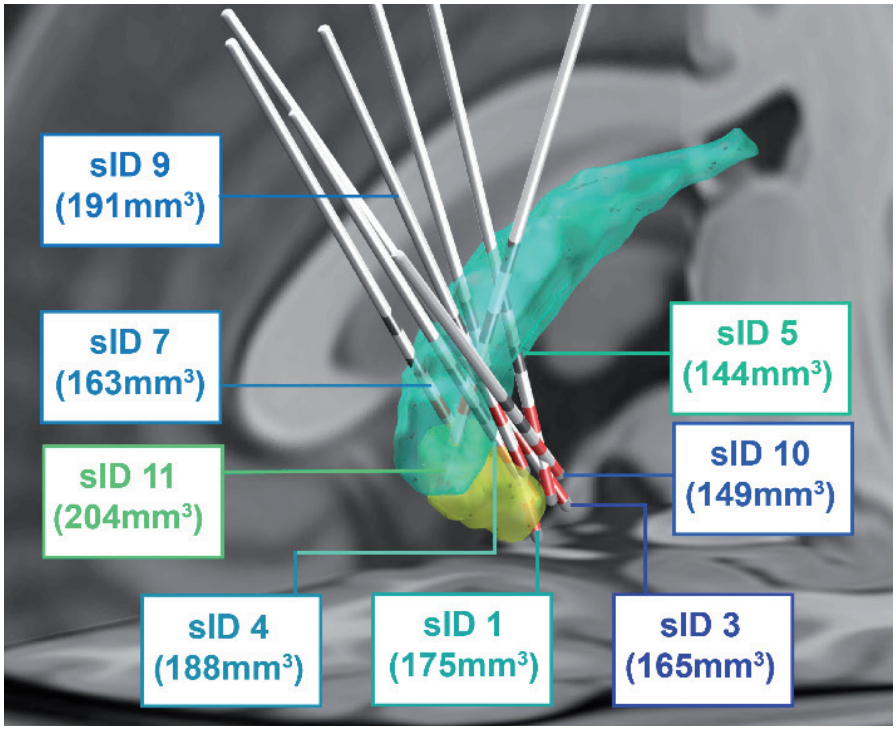
We assessed instantaneous effects of electrical stimulation of the nucleus accumbens during task performance by alternating blocks with (ON) and without (OFF) bilateral stimulation (3.5V; 130Hz; 60µs pulsewidth). We applied bipolar stimulation over the two nucleus accumbens contacts, with the most distal contact as cathode, in line with commonly used therapeutic settings (Benabid et al., 2009; Sturm et al., 2003) and previous reports (Nachev et al., 2015). Based on the stimulation parameters and electrode localization, the activated tissue was estimated using Lead-DBS (version 2.1; Horn and Kühn, 2015). See Figure 2 for the localization of the deep brain electrodes in MNI space and the estimated volumes of nucleus accumbens activation (VAT). Stimulation was switched off for at least two hours prior to testing and was switched ON/OFF one minute before start of each block. The stimulation order was blind to the subject, and counter-balanced across parts and subjects (Figure 1D). None of the patients could recall the employed stimulation order. The transfer phase was always OFF stimulation.

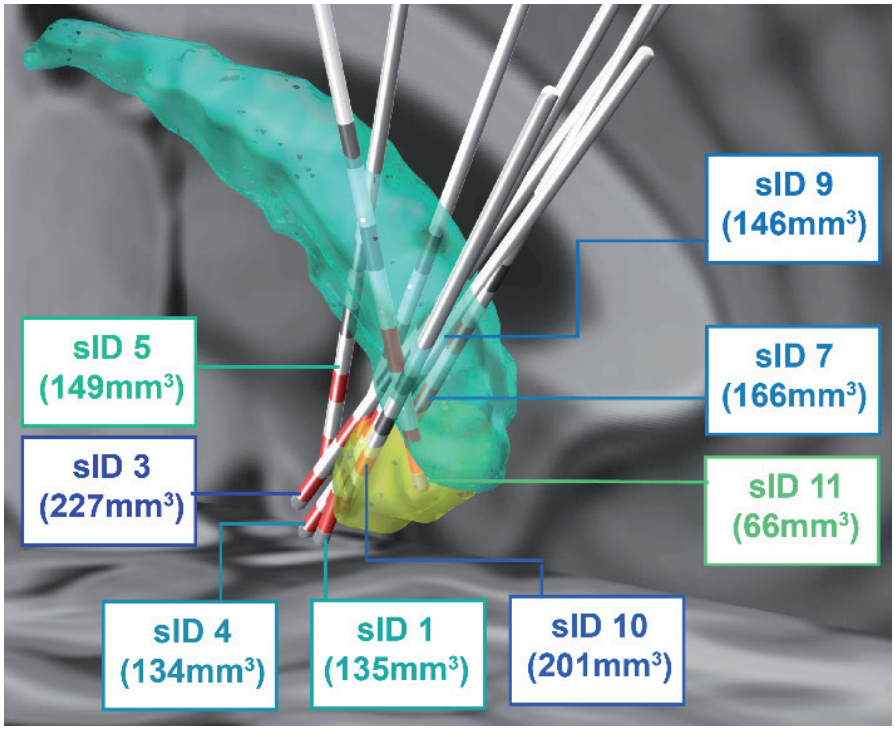
Figure 2 (right page). **Localization of the left (top) and right (bottom) deep brain electrodes in MNI space.** Volume of activated nucleus accumbens tissue (VAT) indicated per hemisphere for each patient/electrode. Throughout the manuscript we colour the individual subject data according to their average VAT. Nucleus accumbens and caudate are depicted according to the Harvard Oxford atlas (Desikan et al., 2006) in yellow and turquoise respectively. Volumes of activated tissue were estimated using Lead-DBS (Horn and Kühn, 2015). See Supplementary Table 2 for the MNI coordinates per contact and Supplementary Table 3 for the estimated activation of surrounding subcortical structures.

Statistical analyses: Task performance

Here we set out to assess the influence of nucleus accumbens stimulation on the motivational biasing of behavioural activation. To this end, we analysed Go vs. NoGo responses as a measure of behavioural activation, and reaction times (RTs) as a complementary measure of behavioural vigour.

Choice data were analysed with a repeated measures ANOVA in SPSS. First, we addressed whether there was overall evidence for task learning and motivational biasing of action by assessing whether the proportion of Go responses varied as a function of the required Go/NoGo action and of the cue valence. After establishing the presence of the standard task effects, we assessed whether DBS attenuated the valence effect on Go responses. Accordingly,


sID 9
(191mm³)
sID 7
(163mm³)
sID 11
(204mm³)
sID 4
(188mm³)
sID 1
(175mm³)
sID 3
(165mm³)
sID 5
(144mm³)
sID 10
(149mm³)


sID 5
(149mm³)
sID 3
(227mm³)
sID 4
(134mm³)
sID 1
(135mm³)
sID 10
(201mm³)
sID 9
(146mm³)
sID 7
(166mm³)
sID 11
(66mm³)

the ANOVA included the within subject factors Required Action (Go vs. NoGo), Valence (Win vs. Avoid cue), and DBS (ON vs. OFF). Given our small, yet highly unique sample, we focus on effect sizes in addition to statistical significance testing. We report partial eta squared (η_p^2) as a measure of effect size, where we interpret $\eta_p^2 > .14$ as large effects, $\eta_p^2 > .06$ as medium effects, and $\eta_p^2 > .01$ as small effects, in line with (Cohen, 1992, 1988).

RT data were analysed with linear mixed-level models using the lme4 package in R (Bates et al., 2014; R Developement Core Team, 2015). The mixed-level models account for both between and within subject variability, taking into account the number and consistency of RTs per subject. Although the choice data would ideally be analysed with logistic mixed-level models as well, these models did not converge for the current sample (presumably due to the limited variability in binomial data on the one hand and a small sample size on the other hand). RTs<100ms were discarded from the analysis and RTs were inverse transformed to improve normality. We limited the RT analysis to correct responses (i.e. RTs on Go cues), to reduce the model's effects structure and thereby increase statistical power, resulting in the within subject factors Valence (Win vs. Avoid cue) and DBS (ON vs. OFF). The mixed model included all main effects and interactions, and a full random effects structure (Barr, 2013; Barr et al., 2013). We estimated effect sizes for RTs based on the corresponding repeated measures ANOVA performed within SPSS as there is no clear consensus on the estimation of effect sizes for mixed-level models.

As mentioned above, the exact mechanisms of action of DBS are still highly debated, and likely involve a complex interplay of both inhibitory and excitatory effects, both locally and network-wide (Chiken and Nambu, 2016). Accordingly, it is unclear whether DBS in the nucleus accumbens will have an overall excitatory or inhibitory effect. However, we hypothesized that the disruptive effect of DBS on the ongoing striatal communication (Chiken and Nambu, 2016) will reduce the coupling of motivationally-driven input and action-related output, or in other words, will reduce the valence-specificity in Go responding and RTs. Thus, we hypothesized that disruption of the ongoing striatal communication with DBS would *reduce* the motivational biases, and therefore we assessed the DBS x Valence interactions with a one-sided test to increase the statistical power given our small sample. All one-sided tests are clearly indicated in the Results section. We exploratory assessed whether the DBS x Valence effect on p(Go) and RT covaries with i) the estimated volume of activated nucleus accumbens tissue by DBS (VAT) and ii) the reduction in OCD symptoms by chronic DBS (ΔY-BOCS; see Supplementary Table 1). We report the Pearson coefficient for these correlations.

Finally, we analysed the relative cue preferences during the transfer phase to confirm that the subjects indeed preferred the Win cues over the Avoid cues. To this end, we analysed how often each cue was selected during the transfer phase relative to chance, using the within-subject factors Valence and Required Action. As we contrasted the choices against chance level, the model for the transfer choices did not include an intercept. One subject did

not complete the transfer phase due to technical problems, and therefore we excluded this subject from the analysis of the transfer phase.

EEG recording, pre-processing, and time-frequency decomposition

Surface EEG was recorded during task performance with 61 (n=5) or 96 (n=3) channels placed according to the 10-20 system. The ground was placed at the chest and the EEG was online referenced to the left earlobe. Channels containing flat lines or persistent artefacts were discarded. In the 61 channel configuration, EEG data was sampled at 1024Hz and horizontal bipolar EOG was recorded with 2 additional channels placed lateral to the eyes. In the 96 channel configuration, EEG data was sampled at 500Hz and discarded channels were interpolated using the disjoint EEG channels, after which only the 61 joint channels were selected for further analysis.

EEG data were re-referenced to the weighted average of the EEG channels. Next, the data were band-pass filtered (0.5-100Hz) to remove slow drifts and the high frequency signal directly resulting from DBS (130Hz; see Stimulation procedure). Continuous EEG data was epoched into segments ranging from -2.25s to 4.7s relative to cue onset. These wide segments were selected to avoid edge artefacts, resulting from time-frequency decomposition, in the relevant time period. The epochs were linear baseline corrected, using the 200ms period prior to cue onset as baseline. The resulting epochs were visually inspected and epochs containing artefacts and excessive EMG noise were removed. An independent component analysis (ICA) was performed over the remaining epochs and components related to eyeblinks and artefacts unrelated to brain activity (mean=3, range=1-6) were removed. For the 61 channel configuration, missing channels were interpolated following ICA (1 channel for 1 subject). Trials with persisting artefacts following ICA were rejected, resulting in a total of ~38 rejected trials on average per subject (mean=7.9%, range=2-15%). Finally, the surface Laplacian was estimated over the EEG data, to filter out distal effects due to volume conduction, i.e. to attenuate the EEG signal so that local effects are better represented in the data (Oostendorp and Van Oosterom, 1996).

The preprocessed time series were decomposed into their time-frequency representations with wavelet convolution. The wavelets ranged from 1-40Hz, in logarithmically spaced steps, with the width ranging from 3 cycles (lowest frequency) to 7 cycles (highest frequency). The time-frequency data were down-sampled to 40Hz. Cue- and response-locked power values were extracted, trial-averaged, and dB baseline corrected with a condition-averaged baseline ranging from -250ms to -50ms relative to cue onset.

Statistical analyses: EEG

After establishing the effects of DBS on the motivational biases, we assessed whether DBS affected the neural signatures related to the detection of motivational conflict. In our previous work (Swart et al., 2018), we showed that theta power (4-8Hz) over the midfrontal cortex

(channels Cz, FCz) increased with motivational conflict, specifically when subjects successfully suppressed the Pavlovian response tendencies and performed the correct instrumental response. Accordingly, we analysed trial-averaged oscillatory power in the theta frequency range (4-8Hz) over the midfrontal electrode cluster {Cz FCz} for correct trials. First, we contrasted cue- and response-locked theta power for motivationally incongruent vs. congruent cues and employed time-wise permutation tests. For the permutation tests, we i) tested which of the time points showed a significant increase in midfrontal theta power for the incongruent vs. congruent cues, ii) created a reference distribution by randomizing the condition labels (500 permutations), and iii) tested the cluster size of significant time windows resulting from step 1 against the permutation distribution created in step 2. Next, we contrasted the resulting time-window for DBS ON vs. OFF using a t-test. Only conditions with ≥10 trials per condition per subject were included in the analysis. Given the low accuracy rate for the NoGo-to-Win trials (only 3 subjects had ≥10 correct trials) and the observation that the behavioural DBS effects appeared specific to the Go trials, we restricted our statistical analysis to the Go trials. Thus, we contrasted correct Go-to-Avoid (incongruent) with Go-to-Win (congruent) trials. Note that by contrasting the Go cues, we cannot disentangle congruency-related signals from valence-related signals in the current study.

For the behavioural analyses we hypothesised that DBS would reduce the motivational biases in Go responding and RTs by directly modulating striatal processing (Statistical analyses: Task performance). Accordingly, we hypothesized that the reduced biases would result in reduced neural markers of conflict detection under DBS, i.e. weaker midfrontal theta power to motivationally incongruent cues. An alternative explanation for reduced motivational biases under DBS, however, could be that DBS distally enhanced frontal control systems, which would be reflected by enhanced midfrontal theta signals to motivationally incongruent cues. To be able to differentiate between these alternative explanations, we assessed the DBS effects on midfrontal theta power with two-sided tests.

Results

General task performance

To address whether standard task effects were present within the sample of OCD patients, we first analysed the proportion of Go responses and response times (RTs) of correct Go responses independent of DBS. The participants indeed adjusted their Go/NoGo responses to the action requirements (Required Action: $F_{1,7}$=7.7, p=.028, η_p^2=.52), reflecting task learning (Figure 3). This action adjustment was marginally higher for Avoid ($F_{1,7}$=11.5, p=.011, η_p^2=.62) than for Win cues ($F_{1,7}$=2.1, p=.189, η_p^2=.23; Valence x Required Action: $F_{1,7}$=4.8, p=.064, η_p^2=.409). Note, that the accuracy was relatively low for the NoGo-to-Win cues in particular (mean[SD] correct = 21.7% [29.2%]), consistent with previous reports showing poorest performance and greatest

variability on the NoGo-to-Win trials (e.g., Cavanagh et al., 2013). Orthogonal to the action requirements, cue valence strongly biased Go responding (Valence: $F_{1,7}$=23.4, p=.002, η_p^2=.77), such that participants made more Go responses to Win than Avoid cues. Complementarily, the participants responded faster to Win than Avoid cues (Valence: X^2_1=55.2, p<.001, η_p^2=.71). These increased and faster Go responses when playing for reward vs. avoiding punishment is what we refer to as the motivational bias. Taken together, the current participants exhibit the standard task effects as we and others have previously reported in healthy samples in similar tasks (Guitart-Masip et al., 2014a; Swart et al., 2018, 2017), namely adjustment of Go/NoGo responses to the action requirements and cue valence biasing the Go/NoGo responses and RTs.

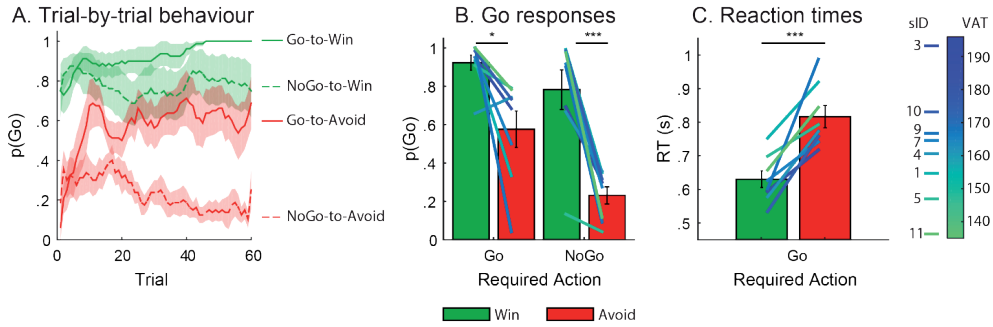


Figure 3. **General task performance collapsed over DBS ON/OFF.**
(**a**) Moving average of the proportion Go responses using a sliding window of 5 trials. Shaded areas represent the standard error of the mean (sem). The participants match their Go/NoGo responding particularly well to the action requirements on the congruent trials (i.e. Go-to-Win, NoGo-to-Avoid), whereas this action adjustment remains relatively poor on the incongruent NoGo-to-Win trials. (**b**) Average proportion of Go responses. Bars represent the group mean (±sem), and the coloured lines represent the individual participant data (coloured by the estimated volume of activated nucleus accumbens tissue, VAT). Participants made more Go responses for the Go than NoGo cues, indicative of task learning. Orthogonally, participants made more Go responses to Win than Avoid cues, reflecting the motivational bias. */*** indicates *p*-values smaller than .05/.001 respectively. (**c**) Average reaction time (RT) of correct Go responses. Participants made faster Go responses on the Go-to-Avoid than Go-to-Win trials (p<.001).

Before turning to the DBS effects, we analysed the choices in the transfer phase to assess whether the cue valences were learned adequately. In the transfer phase, cues were presented in pairs and participants were asked to select the most rewarding cue. All subjects selected win cues more often than avoid cues (Valence: $F_{1,6}$=338.0, p<.001, η_p^2=.98). This preference for Win cues was most pronounced for the NoGo cues (Valence x Required Action: $F_{1,6}$=24.1, p=.039, η_p^2=.54; NoGo cues: $F_{1,6}$=316.2, p<.001, η_p^2=.981), yet clearly present for the Go cues as well (Go cues: $F_{1,6}$=110.4, p<.001, η_p^2=.948). Thus, the explicit relative cue preferences well reflected the cue valences.

Nucleus accumbens stimulation attenuates the motivational bias in behaviour

After having established the presence of the standard task effects, we assessed whether stimulation of the nucleus accumbens attenuates the motivational biasing of action. To this end, we assessed whether DBS diminished the observed valence effect in Go responding and RTs. DBS significantly reduced the valence effect for RTs (DBS x Valence: X^2_1=3.1, p=.040, *one-sided*, η_p^2=.31), such that subjects responded relatively faster to Avoid cues under DBS (Figure 4). Thus, the subjects slowed down for Avoid cues relative to Win cues (see previous section), but to a lesser extent under DBS. One patient did not show this pattern for RTs, which was notably the patient with the smallest volume of activated nucleus accumbens tissue by DBS (VAT; Figure 4). DBS did not significantly alter RTs independent of the cue valence (DBS: X^2_1<1, p=.749, η_p^2=.21).

Complementary to the RTs, DBS numerically attenuated the valence effect on the proportion Go responses (DBS x Valence: $F_{1,7}$=1.1, p=.170, *one-sided*, η_p^2=.13). This medium-sized attenuation of the motivational bias was marginally significant for the Go cues ($F_{1,7}$=1.7, p=.058, *one-sided*, η_p^2=.20), such that DBS marginally increased Go responses to the Go-to-Avoid vs. Go-to-Win cues. Two subjects did not show this pattern of Go responses, which were again the patients with the smallest VAT (Figure 4). The attenuation of the valence effect under DBS was non-significant for the NoGo cues ($F_{1,7}$<1, p=.429, *one-sided*, η_p^2=.005). Yet, note that the three-way interaction (although of large effect size) did not reach significance (DBS x Valence x Required Action: $F_{1,7}$=1.6, p=.240, η_p^2=.19). DBS did not significantly affect the proportion Go responses independent of the cue type (DBS: $F_{1,7}$=1.6, p=.249, η_p^2=.19), or as a function of the required action in line with task learning (DBS x Required Action: $F_{1,7}$=2.1, p=.190, η_p^2=.23). Taken together, we observed medium-large effect sizes of DBS attenuating the impact of cue valence, which was consistent across RTs and Go responding for the Go cues.

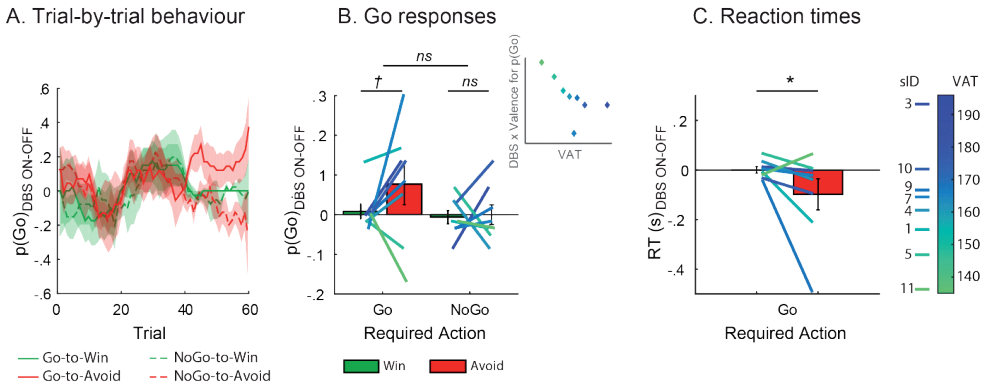


Figure 4. **DBS-induced Go responses and reaction times.**
(**a**) Moving average of the relative increase in the proportion Go responses under DBS. The differential effect of DBS on the Go-to-Win and Go-to-Avoid cues seems to arise particularly towards the end of the

experiment. (**b**) Under DBS, participants made marginally more Go responses for the Go-to-Avoid vs. Go-to-Win cues in particular, in line with reduced motivational biasing of action (p=.058). Notably, the two subjects who did not show this pattern had the smallest volume of activated nucleus accumbens tissue (VAT). For exploratory purposes only, we report whether the DBS x Valence effect on p(Go) covaried with i) VAT and ii) the reduction in OCD symptoms by chronic DBS (ΔY-BOCS; see Supplementary Table 1), yet we emphasize that the sample size is too small for a reliable estimation of these correlations. The DBS-induced reduction in the motivational biasing of Go responses significantly covaried with VAT (R=.81, p=.014), such that patients with larger volume of activated nucleus accumbens tissue showed a greater reduction in the motivational biases (see inset). This correlation did not hold for other subcortical tissues (e.g., caudate; see Supplementary Table 2 for an overview; all p>.6), or for ΔY-BOCS (R=-.31, p=.458). † indicates p<.1. (**c**) Participants also made faster Go responses for the Go-to-Avoid than Go-to-Win cues under DBS (p=.040), in line with reduced motivational biasing under DBS. The subject who did not show this pattern had again the smallest VAT. The DBS-induced reduction in the motivational biasing of response times did not significantly covary with VAT (R=.16, p=.702) or ΔY-BOCS (R=-.56, p=.147).

Midfrontal theta responses to motivational conflict are unaffected by nucleus accumbens stimulation

After observing that DBS attenuated the impact of cue valence on behaviour (significantly for RTs and numerically for Go responses), we assessed whether DBS also affected the neural signatures related to the presence of motivational conflict, reflected in oscillatory theta power (4-8Hz) over the midfrontal cortex (Cavanagh et al., 2013; Swart et al., 2018). We focused on the correct trials as these midfrontal theta signals are known to particularly surface in subjects and trials where control is successfully implemented (Cavanagh et al., 2013; Swart et al., 2018). Given that i) the DBS effect was most pronounced on the Go trials, and ii) five subjects had an insufficient number (<10) of correct NoGo-to-Win trials, we restricted our statistical analyses to the correct Go trials and display the EEG data for the three subjects with sufficient (≥10) correct NoGo-to-Win trials in Supplementary Figure 1. Note that one subject did not have sufficient (<10) correct Go-to-Avoid trials and was discarded from the primary EEG analyses.

Before assessing the effects of DBS, we first contrasted the correct Go-to-Avoid (incongruent) and Go-to-Win (congruent) trials independent of DBS to assess whether the sample of OCD patients exhibited the previously observed conflict effect (Cavanagh et al., 2013; Swart et al., 2018). We extracted oscillatory theta power for the midfrontal cluster and contrasted the Go trials using time-wise permutation testing. Midfrontal theta power significantly increased for the Go-to-Avoid trials relative to the Go-to-Win trials during one cue-locked time-window (275-475ms; p=.024; Figure 5A). The corresponding permutation test for the response-locked power also yielded a significant time-window (-1000 to -800ms, p=.048), yet these time points were too early to reflect a control-related signal, which commonly peak closer to the response. Given the average reaction time of ~.7s, this early condition effect might instead be driven by the RT difference for the Go-to-Win and Go-to-Avoid cues (i.e., capturing more pre- and post-baseline activity respectively).

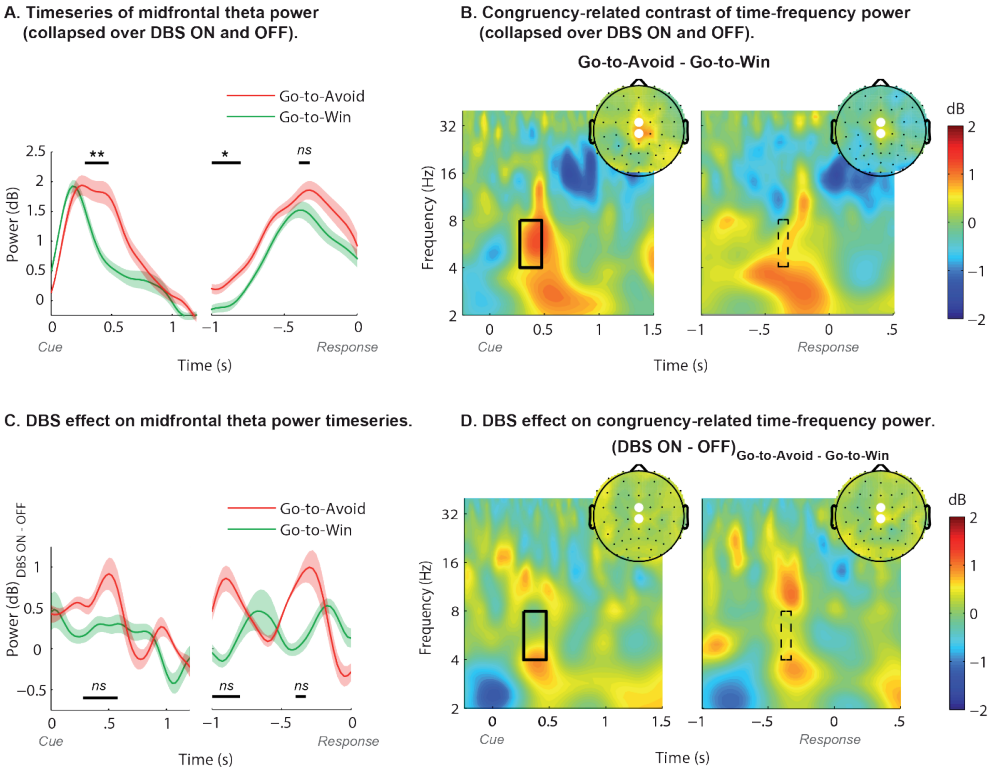


Figure 5. **Cue- and response-locked midfrontal oscillatory theta activity.**
(**a**) Midfrontal theta power collapsed over DBS ON and OFF (Cz+FCz, 4-8Hz). Time-wise permutation testing reveals one cue-locked (275 to 475ms; *p*=.024) time window during which midfrontal theta power significantly increased for the Go-to-Avoid (incongruent) vs. Go-to-Win (congruent) trials. The response-locked permutation test also yielded a significant response-locked time window (-1000 to -800ms, *p*=.048), yet this time window was too early to reflect a control-related signal, which commonly peaks closer to the response. Although the permutation test additionally indicated a time window closer to the response (-400 to -325ms), this window was not significant (*p*=.172). (**b**) Go-to-Avoid vs. Go-to-Win contrast for the midfrontal channels and resulting time window. The black frame indicates the time-frequency window displayed in the topoplot; the white discs indicate the midfrontal channels displayed in the time-frequency plot. (**c**) Relative increase in midfrontal theta power under DBS. DBS did not significantly alter midfrontal theta power in the cue-locked time window across the group (*p*=.205), nor as a function of the volume of activated nucleus accumbens tissue (R=-.64, *p*=.125) or the behavioural DBS x Valence effect (p(Go): R=.25, *p*=.583; RT: R=.15, *p*=.757). (**d**) Relative increase in the Go-to-Avoid vs. Go-to-Win contrast under DBS. See Supplementary Figure 1 for topoplots of individual subjects.

The response-locked permutation test additionally indicated a time-window closer to the response (-400 to -325ms), yet this time-window was not significant (*p*=.172). Both the response- and cue-locked permutations tests did not indicate any converse time-windows in which midfrontal theta power was increased for the congruent Go-to-Win relative to the incongruent Go-to-Avoid cue. Taken together, the results were consistent with a cue-locked

increase in midfrontal theta power to motivational conflict, in line with previous findings in healthy populations (Cavanagh et al., 2013; Swart et al., 2018). Note, however, that we could not disentangle the neural responses to motivational conflict and the cue valence in the current study as we only assessed the Go cues, yet the previous studies (Cavanagh et al., 2013; Swart et al., 2018) linked the midfrontal theta signals to motivational conflict, rather than to cue valence.

After establishing the increase in cue-locked midfrontal theta power for the Go-to-Avoid (incongruent) vs. Go-to-Win (congruent) cues, we extracted the power values for the resulting time-window and assessed the effect of DBS. DBS did not significantly attenuate the cue-locked midfrontal theta response to motivational incongruency across the group (t_6=1.4, p=.205, η_p^2=.252; Figure 5C). There were also no significant correlations between the DBS effect on cue-locked midfrontal theta response and behavioural DBS x Valence effect (p(Go): R=.25, p=.583; RT: R=.15, p=.757) or the estimated impact of DBS on the nucleus accumbens (R=-.64, p=.125). We also did not observe a significant main effect of DBS in the time-windows of interest (t_7=1.3, p=.249, η_p^2=.184), or during the baseline period (-250 to -50ms; t_7=1.1, p=.296, η_p^2=.154). Altogether, these data suggest that the midfrontal theta signals to the incongruent Go-to-Avoid vs. congruent Go-to-Win cues were unaffected by nucleus accumbens DBS.

Discussion

Here we studied the effects of nucleus accumbens stimulation on the well-established motivational biases in action (Cools et al., 2011; Guitart-Masip et al., 2014a) and motivational conflict-related neural signatures (Cavanagh et al., 2013; Cavanagh and Frank, 2014; Cohen, 2014; Swart et al., 2018). Cue valence had pronounced impact on the number of active Go responses and the associated reaction times, such that subjects made more and faster Go responses to reward (Win) cues than to punishment (Avoid) cues. When these response tendencies conflicted with instrumental requirements, midfrontal theta power increased when subjects made correct Go responses. DBS of the bilateral nucleus accumbens significantly attenuated the motivational bias in reaction times; subjects made relatively faster (correct) responses to the Avoid cues during DBS. This speeding of Go-to-Avoid responses was paralleled by a corresponding increase in the proportion of Go responses, yet this was only marginally significant for the Go cues. Intriguingly, an enhancing effect of DBS on the proportion of Go-to-Avoid responses did surface as a function of individual differences in the intensity of the stimulation, although clearly this between-subject effect has to be interpreted with caution given the small sample size. In contrast to these effects on behaviour, DBS of the bilateral nucleus accumbens left the midfrontal theta responses to motivational conflict unaffected.

The attenuation of the motivational biases (during Go responses) by striatal DBS provides causal support for current views of basal ganglia functioning (Collins and Frank, 2014); Reward cues elicit peaks in mesolimbic dopamine release, which potentiate the striatal direct 'Go' pathway and thereby promote behavioural activation. Similarly, punishment cues elicit dips in mesolimbic dopamine release, potentiating the indirect NoGo pathway and promoting behavioural inaction. Put simply, the striatal input (depending on cue valence) biases the striatal output (determining behavioural activation). Accordingly, we hypothesised that disrupting the striatal input-output coupling with striatal DBS (Chiken and Nambu, 2016) would reduce the influence of cue valence on behavioural activation. In line with our hypothesis, the influence of the cue valence on reaction times was significantly reduced during striatal DBS. This reduced influence of cue valence was particularly pronounced on the Go-to-Avoid cues, where subjects OFF DBS slowed down relative to the Go-to-Win cues, but to a lesser extent ON DBS. We observed a similar marginal attenuation in the proportion of Go responses, where subjects OFF DBS reduced Go responding for the Go-to-Avoid cues relative to the Go-to-Win cues, and to a lesser extend ON DBS. Note, however, that the three-way interaction (DBS x Valence x Required Action) was not statistically significant across the patient group.

The DBS-induced disinhibition on Go cues is generally consistent with previous findings in the same sample of patients (Nachev et al., 2015), where nucleus accumbens DBS disinhibited the active selection of risky, yet optimal choice options. OFF DBS these patients avoided high-risk options and were biased towards actively selecting suboptimal, low-risk options instead. DBS reduced this risk-avoidance bias, enabling active selection of instrumental choice options. Together with the current pattern of results, these findings suggest that DBS of the nucleus accumbens might release aversive inhibition of active behaviour.

We hypothesized that if DBS reduced the motivational biasing of action, the neural responses to motivational conflict would also decrease. Particularly the midfrontal cortex has been linked to detection of conflict between cue-based, Pavlovian response biases and goal-directed, instrumental actions, in order to signal this conflict within the task-related network for the implementation of control (Cavanagh et al., 2013; Cavanagh and Frank, 2014; Cohen, 2014; Swart et al., 2018). In our sample, we could extract these conflict-related signals on the correct Go trials, where cue-locked midfrontal theta power increased for the incongruent, Go-to-Avoid cues relative to the congruent, Go-to-Win cues. Yet, we did not observe a DBS-induced reduction of these conflict-related neural responses, suggesting that motivational conflict detection in the midfrontal cortex was not affected by DBS of the nucleus accumbens. We acknowledge that the current results do not allow us to disentangle the neural responses to conflict and cue valence, yet

previous studies linked the midfrontal theta signals to motivational conflict and not cue valence (Cavanagh et al., 2013; Swart et al., 2018). Nevertheless, our data suggest that midfrontal theta signals, commonly associated with conflict detection, are unaffected by DBS. One potential explanation for why this hypothesized reduction was absent is that the motivational bias remains strikingly potent, even following the reduction under DBS. Thus, the conflict between the motivational response tendencies and the instrumental requirements might remain too persistent to result in reduced neural conflict signals. In other words, the reduction in motivational biases by DBS might have been insufficient for some of the patients to be picked up by the midfrontal cortex, whereas a larger attenuation could have translated in reduced neural conflict signals.

Although our behavioural findings are highly consistent with reduced motivational biases due to DBS-induced disruption of striatal input-output coupling, these findings could potentially also be explained by enhanced cognitive control over behaviour. Striatal DBS does not only affect local, striatal processing, but also affects frontostriatal network connectivity (Figee et al., 2013). Thus, the reduced motivational biasing of action could potentially also reflect enhanced frontal control due to distal effects of DBS. However, if distal, frontal effects of DBS facilitated motivational conflict detection in the midfrontal cortex, and thereby facilitated the recruitment of control over behaviour, we would expect to see *enhanced* conflict-related midfrontal signals under DBS. Thus, the unaffected midfrontal theta signals suggest that our behavioural DBS observations are unlikely to reflect a distal enhancement of frontal control systems. Moreover, DBS of the subthalamic nucleus has been shown to hamper cognitive control and to leave the midfrontal theta responses unaffected as well (Cavanagh et al., 2011). Future studies might address the hypothesis, raised by the current study, that the DBS-induced reduction of aversive inhibition reflects (local) processing in subcortical brain regions, such as the nucleus accumbens itself, or its interaction with the amygdala.

Next to a disruption of the ongoing neural communication, DBS might have an overall inhibitory or excitatory effect depending on the balance of inhibitory (e.g., GABAergic) and excitatory (e.g. glutamatergic) neurons that are being stimulated (Chiken and Nambu, 2016). We reasoned that an overall excitatory or inhibitory effect of nucleus accumbens DBS might have a global, condition-independent effect on behavioural activation via changes in activity of the basal ganglia direct and indirect pathways. Although DBS numerically increased Go responding and reduced response times, these main effects were specifically driven by the Go-to-Avoid cues. Thus, we did not observe a convincing effect of DBS on *non-selective*, global motor activation. Nonetheless, nucleus accumbens DBS might have a net excitatory/inhibitory effect on local neuronal activity, yet such a potential net effect did not seem to result in overall changes in motor excitability.

Finally, given that we assessed a specific psychiatric population, namely treatment-refractory OCD patients, this raises the question whether the observed behavioural changes during DBS reflect normalization of behaviour. That is, DBS has previously been suggested to normalize both behavioural and neural processes in a range of psychiatric populations (e.g., Figee et al., 2013; Grubert et al., 2011; Kuhn et al., 2011; McCracken and Grace, 2009). However, so far OCD patients have not been compared with matched controls on the current experimental paradigm. Thus, future work is needed to elucidate to what extent the motivational biases in action are altered in OCD, in order to determine whether the observed DBS effects reflect normalization or disruption of the biases.

To conclude, here we assessed for the first time the causal involvement of the human ventral striatum in the motivational biasing of action by direct stimulation of the nucleus accumbens. Deep brain stimulation attenuated the motivational biases, putatively by releasing aversive inhibition of active responses, rather than enhancing distal control networks. We did not observe non-selective effects of DBS on behavioural activation, suggesting that stimulation of the nucleus accumbens does not affect motor excitability per se, but rather attenuates the coupling between motivation and behavioural activation.

Supplementary Files

Supplementary Table 1. **Patient demographics and presurgical clinical evaluation.**

	sID 1	sID 3	sID 4*	sID 5	sID 7	sID 9	sID 10	sID 11
Age	49	28	28	21	50	30	39	36
Gender	F	F	M	M	M	F	M	M
Handedness	R	R	R-L	R	R	L	R	R
Schooling (y)	15	15	11	16	9	12	15	12
Age at onset	12	7	10	10	9	8	22	8
Duration (y)	37	21	18	11	41	22	17	28
Diagnosis	OCD	OCD	Axis I: OCD Axis II: SPD	OCD	OCD	OCD	OCD	OCD
SCID-I Coding	300.3	300.3	300.3	300.3	300.3/305.00	300.3, 307.1, 301.9	300.3	300.3, 303.9, 292.9, 300.21
Comorbidity	None	None	Schizotypal personality disorder	None	Alcohol abuse	Anorexia Nervosa, Personality disorder not otherwise specified	None	Alcohol and other recreational drug dependence (cocaine) /Social phobia
Obsessions	Contamination/ Taboo thoughts (aggressive)	Taboo thoughts (religious, aggressive) / Contamination/ Doubts and checking	Symmetry / Ordering/ Taboo thoughts, (aggressive, sexual)	Doubts / Aggressive obsessive thoughts	Doubts and checking / Contamination	Contamination / Symmetry / Ordering, taboo thoughts / Hoarding	Contamination / Doubts	Taboo thoughts (religious, aggressive)

Continued - next page

Continued. **Supplementary Table 1 - Patient demographics and presurgical clinical evaluation.**

	sID 1	sID 3	sID 4*	sID 5	sID 7	sID 9	sID 10	sID 11
Compulsions	Cleaning	Cleaning	Ordering / Symmetry	Checking	Checking/ Cleaning	Cleaning	Cleaning / Checking	Avoiding behaviours / Checking
Drug therapy (mg/day)	Aripiprazole 10 Clonazepam 3 Valproic Acid 1000 Venlafaxine 225	Sertraline 200	Clomipramine 150 Sertraline 300 Oxitriptan 200	?	Clomipramine 225 Diazepam 5	?	?	Venlafaxine 225 Quetiapine 150
Previous therapies	CBT, STPP, GPT	Nil	TMS	?	CBT	?	?	Nil
Y-BOCS	34	31	13	24	38	30	37	36
BDI	27	24	22	27	37	18	37	51
STAI-S	40	44	24	43	53	45	47	58
STAI-T	49	45	40	49	48	51	42	60
CGI	6	6	6	5	6	6	6	6
GAF	30	25	25	50	25	20	20	25
%Y-BOCS ↓	47.22	41.38	15.39	4.17	15.79	-3.33	21.62	94.44

F/M: female/male. R/L: right/left. OCD: obsessive-compulsive disorder. SPD: Schizotypal personality disorder. SCID-I: Structured Clinical Interview for DSM-IV Axis I Disorders. CBT: cognitive behavioural therapy. STPP: short-term psychodynamic psychotherapy. GPT: group psychotherapy. TMS: transcranial magnetic stimulation. Previous therapies: Other previous psychotherapy or transcranial brain stimulation. Y-BOCS: Yale–Brown Obsessive Compulsive Scale. BDI: Beck Depression Inventory. STAI-S/T: State-Trait Anxiety Inventory–State/Trait version. CGI: Clinical Global Impression-severity scale. GAF: Global Assessment of Functioning. %Y-BOCS ↓: percentage reduction in Y-BOCS score following stimulation of the most distal nucleus accumbens contact. *This patient initially presented compulsion, which responded to pharmacotherapy. His score on the Yale–Brown Obsessive Compulsive Scale is in the mild range, which reflects the relative insensitivity of this scale to primarily obsessive symptomatology (Nachev et al., 2015).

Supplementary Table 2. **Localization of DBS contacts in MNI space (X/Y/Z).**

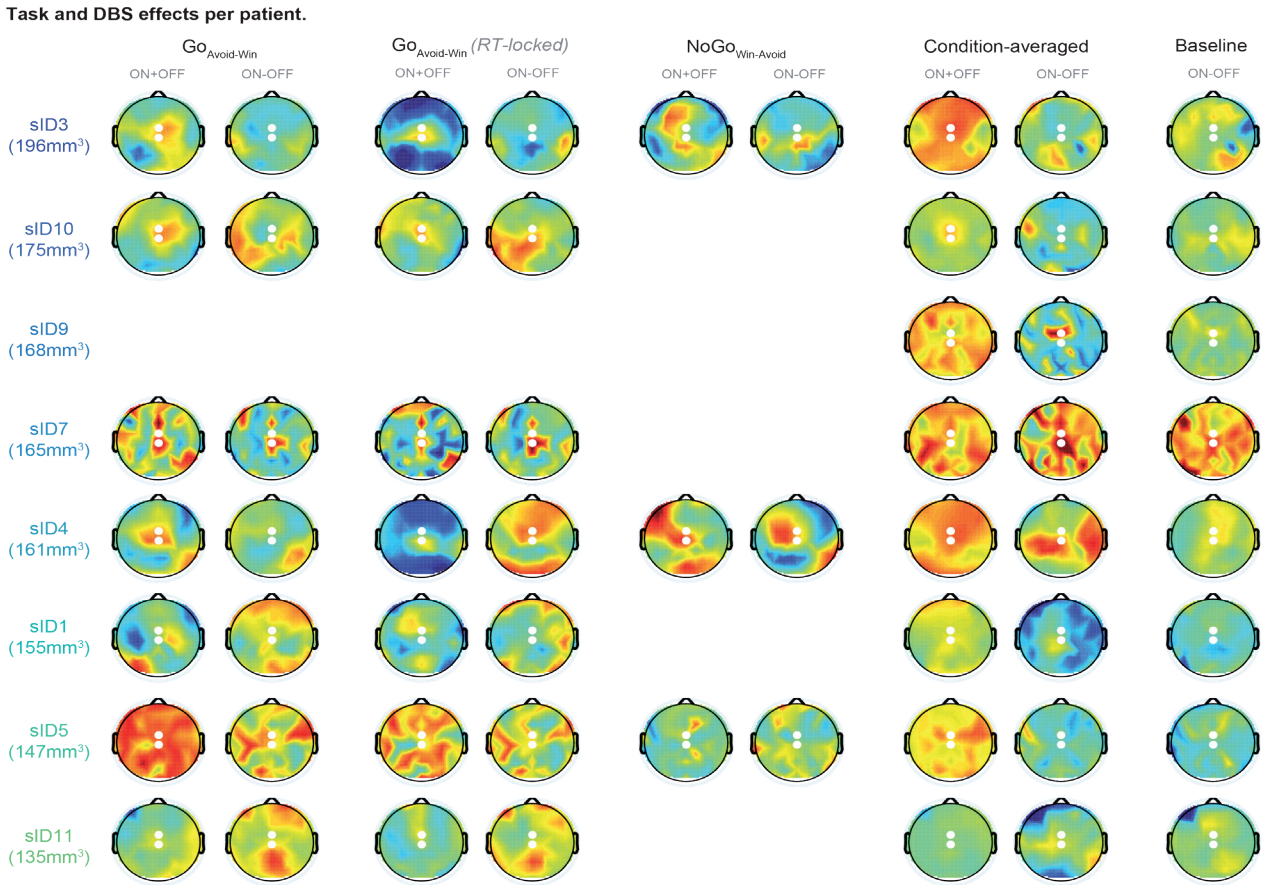
	sID 1	sID 3	sID 4	sID 5	sID 7	sID 9	sID 10	sID 11
MNI coordinates								
- left electrode								
C0	-6.3/2.2/-13.0	-7.7/0.2/-10.5	-6.2/2.4/-10.5	-8.5/1.9/-9.0	-4.2/9.0/-5.4	-10.9/7.4/-5.2	-7.5/0.4/-8.3	-7.0/13.1/-6.3
C1	-8.0/6.3/-5.0	-9.4/5.7/-4.1	-8.4/7.1/-5.2	-9.9/4.8/-2.0	-7.0/14.1/0.3	-13.7/12.0/0.8	-11.0/5.1/-4.3	-9.9/12.7/1.6
C2	-9.8/10.1/2.8	-11.0/11.2/2.3	-10.5/11.8/0.1	-11.4/7.7/5.1	-9.7/19.2/6.1	-16.5/16.7/6.7	-14.6/9.9/-0.4	-12.9/12.3/9.4
C3	-11.3/13.6/10.6	-12.7/16.7/8.6	-12.7/16.5/5.4	-12.8/10.7/12.2	-12.5/24.3/11.9	-19.3/21.3/12.7	-18.4/14.7/3.4	-15.8/11.9/17.2
- right electrode								
C0	7.2 /2.5/-12.1	5.6/-0.8/-10.8	2.9/-2.5/-14.2	4.2/-0.2/-8.1	9.6/13.6/-0.9	7.5/4.0/-9.5	7.0/3.8/-11.1	11.9/14.3/-4.5
C1	7.8 /7.2/-4.4	7.9/5.0/-5.1	5.2/2.5/-9.4	7.2/3.2/-1.9	12.1/19.1/5.2	9.7/9.0/-3.7	10.4/8.3/-6.6	13.8/13.8/3.7
C2	9.2 /11.2/3.2	10.2/10.8/0.6	7.4/7.5/-4.5	10.1/6.6/4.3	14.6/24.5/11.3	11.9/14.0/2.1	13.6/12.9/-2.0	15.7/13.3/11.9
C3	10.6/14.8/10.9	12.4/16.6/6.3	9.7/12.6/0.4	13.0/10.1/10.5	17.1/29.9/17.4	14.1/19.0/7.9	16.8/17.5/2.4	17.7/12.8/20.2
Distance to NAcc								
	2.4	4.3	4.6	4.6	4.8	3.5	2.4	4.2

All patients were implanted with bilateral deep brain electrodes (Medtronic 3391), containing four contacts (C0-C3). The two most distal contacts of each electrode (C0,C1) were positioned within the posterior part of the nucleus accumbens (NAcc) and were used for bipolar stimulation. The average distance (mm) between these bipolar contacts and the NAcc atlas structure (Harvard Oxford atlas; Desikan et al., 2006) is provided at the bottom of the table.

Supplementary Table 3. **Volume of activated tissues (mm³) estimated using Lead-DBS (Horn and Kühn, 2015).**

	sID 1	sID 3	sID 4	sID 5	sID 7	sID 9	sID 10	sID 11
Nucleus accumbens	155	196	161	147	165	168	175	135
Caudate	86	211	148	156	257	217	145	295
Bed nucleus of the stria terminalis	138	38	89	19	8	50	76	0
Basal forebrain	138	213	144	170	0	61	98	0
Fornix	72	190	101	178	0	15	83	0
Mamillary bodies	0	5	34	0	0	0	0	0

Nucleus accumbens and caudate according to the Harvard Oxford atlas (Desikan et al., 2006); Bed nucleus of the stria terminalis, fornix, and mammillary bodies according to the DISTAL atlas (Ewert et al., 2018); Basal forebrain according to the SPM Anatomy toolbox (Eickhoff et al., 2005).



Supplementary Figure 1. **Overview of subject-specific task and DBS effects on oscillatory theta power.**
Task contrasts for the resulting time-frequency windows (Figure 5A) are displayed independent of DBS (left images) and for DBS ON relative to DBS OFF (right images). Contrasts are cue-locked unless indicated otherwise. Subjects are shorted according to their estimated volume of nucleus accumbens activation.

References

Ahmari, S.E., Dougherty, D.D., 2015. DISSECTING OCD CIRCUITS: From ANIMAL MODELS to TARGETED TREATMENTS. Depress. Anxiety. doi:10.1002/da.22367

Barr, D.J., 2013. Random effects structure for testing interactions in linear mixed-effects models. Front. Psychol. 4, 328. doi:10.3389/fpsyg.2013.00328

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4. R Packag. version. doi:citeulike-article-id:7112638

Benabid, A.L., Chabardes, S., Mitrofanis, J., Pollak, P., 2009. Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. Lancet Neurol. doi:10.1016/S1474-4422(08)70291-6

Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Res. Rev. doi:10.1016/S0165-0173(98)00019-8

Burguière, E., Monteiro, P., Mallet, L., Feng, G., Graybiel, A.M., 2015. Striatal circuits, habits, and implications for obsessive-compulsive disorder. Curr. Opin. Neurobiol. doi:10.1016/j.conb.2014.08.008

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., Frank, M.J., 2013. Frontal Theta Overrides Pavlovian Learning Biases. J. Neurosci. 33, 8541–8548. doi:10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. doi:10.1016/j.tics.2014.04.012

Cavanagh, J.F., Wiecki, T. V, Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., Frank, M.J., 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat. Neurosci. 14, 1462–1467. doi:10.1038/nn.2925

Chiken, S., Nambu, A., 2016. Mechanism of Deep Brain Stimulation: Inhibition, Excitation, or Disruption? Neuroscientist 22, 1073858415581986-. doi:10.1177/1073858415581986

Cohen, J., 1992. A power primer. Psychol. Bull. 112, 155–159. doi:10.1037/0033-2909.112.1.155

Cohen, J., 1988. Statistical power analysis for the behavioral sciences. Stat. Power Anal. Behav. Sci. doi:10.1234/12345678

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482, 85–88. doi:10.1038/nature10754

Cohen, M.X., 2014. A neural microcircuit for cognitive conflict detection and signaling. Trends Neurosci. doi:10.1016/j.tins.2014.06.004

Cohen, M.X., Cavanagh, J.F., 2011. Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. Front. Psychol. 2. doi:10.3389/fpsyg.2011.00030

Cohen, M.X., Donner, T.H., 2013. Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. J. Neurophysiol. 110, 2752–2763. doi:10.1152/jn.00479.2013

Collins, A.G.E., Frank, M.J., 2014. Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. 121, 337–66. doi:10.1037/a0037015

Cools, R., Nakamura, K., Daw, N.D., 2011. Serotonin and dopamine: Unifying affective, activational, and decision functions. Neuropsychopharmacology. doi:10.1038/npp.2010.121

Day, J.J., Roitman, M.F., Wightman, R.M., Carelli, R.M., 2007. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat. Neurosci. 10, 1020–1028. doi:10.1038/nn1923

Dayan, P., Niv, Y., Seymour, B., D. Daw, N., 2006. The misbehavior of value and the discipline of the will. Neural Networks 19, 1153–1160. doi:10.1016/j.neunet.2006.03.002

DeLong, M.R., Wichmann, T., 2007. Circuits and circuit disorders of the basal ganglia. Arch. Neurol. 64, 20–24. doi:10.1001/archneur.64.1.20

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968–980. doi:10.1016/j.neuroimage.2006.01.021

Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18. doi:10.3758/BF03199951

Dickinson, A., Smith, J., Mirenowicz, J., 2000. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav. Neurosci. 114, 468–483. doi:10.1037/0735-7044.114.3.468

Figee, M., Luigjes, J., Smolders, R., Valencia-Alfonso, C.E., Van Wingen, G., De Kwaasteniet, B., Mantione, M., Ooms, P., De Koning, P., Vulink, N., Levar, N., Droge, L., Van Den Munckhof, P., Richard Schuurman, P., Nederveen, A., Van Den Brink, W., Mazaheri, A., Vink, M., Denys, D., 2013. Deep brain stimulation restores frontostriatal network activity in obsessive-compulsive disorder. Nat. Neurosci. 16, 386–387. doi:10.1038/nn.3344

Grubert, C., Hurlemann, R., Bewernick, B.H., Kayser, S., Hadrysiewicz, B., Axmacher, N., Sturm, V., Schlaepfer, T.E., 2011. Neuropsychological safety of nucleus accumbens deep brain stimulation for major depression: Effects of 12-month stimulation. World J. Biol. Psychiatry 12, 549–558. doi:10.3109/15622975.2011.583942

Guitart-Masip, M., Duzel, E., Dolan, R., Dayan, P., 2014. Action versus valence in decision making. Trends Cogn. Sci. doi:10.1016/j.tics.2014.01.003

Guitart-Masip, M., Fuentemilla, L., Bach, D.R., Huys, Q.J., Dayan, P., Dolan, R.J., Duzel, E., 2011. Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. J Neurosci 31, 7867–7875. doi:10.1523/JNEUROSCI.6376-10.2011

Hebart, M.N., Gläscher, J., 2015. Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. Psychopharmacology (Berl). 232, 437–451. doi:10.1007/s00213-014-3682-3

Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A., Galarraga, E., 1997. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. J Neurosci 17, 3334–3342.

Hernandez-Lopez, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., Surmeier, D.J., 2000. D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLC[beta]1-IP3-calcineurin-signaling cascade. J. Neurosci. 20, 8987–8995. doi:20/24/8987 [pii]

Holthoff-Detto, V., 1997. Functional Effects of Striatal Dysfunction in Parkinson Disease. Arch Neurol 76, 550–558. doi:10.1016/j.biopsych.2013.12.014

Horn, A., Kühn, A.A., 2015. Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. Neuroimage 107, 127–135. doi:10.1016/j.neuroimage.2014.12.002

Huys, Q.J.M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R.J., Dayan, P., 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. PLoS Comput. Biol. 7. doi:10.1371/journal.pcbi.1002028

Kuhn, J., Gründler, T.O.J., Bauer, R., Huff, W., Fischer, A.G., Lenartz, D., Maarouf, M., Bührle, C., Klosterkötter, J., Ullsperger, M., Sturm, V., 2011. Successful deep brain stimulation of the nucleus accumbens in severe alcohol dependence is associated with changed performance monitoring. Addict. Biol. 16, 620–623. doi:10.1111/j.1369-1600.2011.00337.x

Lex, A., Hauber, W., 2008. Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. Learn. Mem. 15, 483–491. doi:10.1101/lm.978708

MacDonald, P.A., MacDonald, A.A., Seergobin, K.N., Tamjeedi, R., Ganjavi, H., Provost, J.S., Monchi, O., 2011. The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: Support from functional MRI. Brain 134, 1447–1463. doi:10.1093/brain/awr075

Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459, 837–841. doi:10.1038/nature08028

McCracken, C.B., Grace, A.A., 2009. Nucleus accumbens deep brain stimulation produces region-specific alterations in local field potential oscillations and evoked responses in vivo. J Neurosci 29, 5354–5363. doi:10.1523/JNEUROSCI.0131-09.2009

Mink, J.W., Thach, W.T., 1991. Basal ganglia motor control. III. Pallidal ablation: normal reaction time, muscle cocontraction, and slow movement. J. Neurophysiol. 65, 330–351.

Nachev, P., Lopez-Sosa, F., Gonzalez-Rosa, J.J., Galarza, A., Avecillas, J., Pineda-Pardo, J.A., Lopez-Ibor, J.J., Reneses, B., Barcia, J.A., Strange, B., 2015. Dynamic risk control by human nucleus accumbens. Brain 138, 3496–3502. doi:10.1093/brain/awv285

Oostendorp, T.F., Van Oosterom, A., 1996. The surface Laplacian of the potential: Theory and application. IEEE Trans. Biomed. Eng. 43, 394–405. doi:10.1109/10.486259

Pujara, M., Koenigs, M., 2014. Mechanisms of reward circuit dysfunction in psychiatric illness: Prefrontal-striatal interactions. Neuroscientist. doi:10.1177/1073858413499407

R Developement Core Team, 2015. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. 1, 409. doi:10.1007/978-3-540-74686-7

Robbins, T.W., Everitt, B.J., 2007. A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). Psychopharmacology (Berl). doi:10.1007/s00213-006-0528-7

Robbins, T.W., Everitt, B.J., 1996. Neurobehavioural mechanisms of reward and motivation. Curr. Opin. Neurobiol. doi:10.1016/S0959-4388(96)80077-8

Salamone, J.D., Correa, M., Mingote, S.M., Weber, S.M., 2005. Beyond the reward hypothesis: Alternative functions of nucleus accumbens dopamine. Curr. Opin. Pharmacol. doi:10.1016/j.coph.2004.09.004

Stoessl, J., 2008. Potential therapeutic targets for Parkinson's disease. Expert Opin. Ther. Targets 12, 425–36. doi:10.1517/14728222.12.4.425

Sturm, V., Lenartz, D., Koulousakis, A., Treuer, H., Herholz, K., Klein, J.C., Klosterkötter, J., 2003. The nucleus accumbens: A target for deep brain stimulation in obsessive-compulsive- and anxiety-disorders, in: Journal of Chemical Neuroanatomy. pp. 293–299. doi:10.1016/j.jchemneu.2003.09.003

Swart, J.C., Frank, M.., Määttä, J.I., Jensen, O., Cools, R., den Ouden, H.E.M., 2018. Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. PLoS Biol.

Swart, J.C., Froböse, M.I., Cook, J.L., Geurts, D.E.M., Frank, M.J., Cools, R., den Ouden, H.E.M., 2017. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. Elife 6. doi:10.7554/eLife.22169

Taylor, J.R., Robbins, T.W., 1986. 6-Hydroxydopamine lesions of the nucleus accumbens, but not of the caudate nucleus, attenuate enhanced responding with reward-related stimuli produced by intra-accumbens d-amphetamine. Psychopharmacology (Berl). 90, 390–397. doi:10.1007/BF00179197

Taylor, J.R., Robbins, T.W., 1984. Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. Psychopharmacology (Berl). 84, 405–412. doi:10.1007/BF00555222

Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. Science 307, 1642–5. doi:10.1126/science.1105370

Wyvell, C.L., Berridge, K.C., 2000. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. J. Neurosci. 20, 8122–8130. doi:http://www.jneurosci.org/content/21/19/7831

Chapter 6

General discussion

All truths are easy to understand once they are discovered; the point is to discover them.
- Galileo Galilei

The primary aim of this thesis was to provide a deeper understanding of the neurocomputational mechanisms i) underlying the motivational biasing of action, and ii) involved in regulating dysfunctional motivational drives, allowing for adequate behaviour. Given the theorized role of striatal dopamine function in subserving the motivational biasing of action and the role of the medial frontal cortex in regulating dysfunctional response tendencies, I focused on these aspects specifically. To achieve these aims, we conducted several related experiments in which we employed a motivational Go/NoGo task. In the motivational Go/NoGo task, subjects need to learn to make Go responses (i.e. button presses) to the Go cues, and NoGo responses (i.e. refrain from button presses) to the NoGo cues. By making the correct responses, subjects are rewarded for half the cues (Win cues) and avoid punishment for the other cues (Avoid cues). Thus, based on the feedback, subjects can then learn the correct response for each cue. With this learning phase, we can probe to what extent people are biased towards making active, Go responses when pursuing reward and making inactive, NoGo responses when avoiding punishment. The learning phase was followed by a transfer phase in a subset of the experiments. During this phase, the preceding cues are presented in pairs and subjects needed to select their preferred cue out of the pairs. Thus, the transfer phase allowed us to probe the relative cue values after learning. In this chapter, I provide a summary of the main findings described in this thesis, discuss and integrate the most relevant findings, and highlight the relevance for future studies.

Main findings

Genetic carriers of dopamine-related pathogenic variants express altered subjective valuation, but unaffected motivational biases

Catecholamines (particularly dopamine) are strongly linked to motivation, learning and behavioural activation (Berridge and Robinson, 1998; Brozoski et al., 1979; Cools et al., 2009; Frank et al., 2004; Robbins and Everitt, 2007; Salamone et al., 2005; Schultz et al., 1997). Polymorphisms in dopamine-regulating genes have been implicated in these processes as well (Frank and Fossella, 2011), yet the cognitive effects of pathogenic variants in the tyrosine hydroxylase (TH) gene have never been studied. TH is the enzyme that is responsible for the conversion of the amino acid L-tyrosine into dopamine's direct precursor L-DOPA (Kurian et al., 2011), and a pathogenic variant in the TH gene putatively reduces dopamine function by limiting the synthesis of L-DOPA. In **chapter 2**, we assessed for the first time whether carriers of a pathogenic variant in the TH genes showed altered motivation and instrumental action. To this end, we employed a motivational Go/NoGo learning task, and compared 16 carriers of pathogenic TH variants with 20 education- and age-matched controls. In the initial learning phase of this task, subjects learnt to make Go or NoGo responses to cues that predict reward vs. punishment. In the final transfer phase, the subjects were presented with pairs of cues and chose the one they preferred, in the absence of reinforcement. In both the carriers and

matched controls, cue valence strongly biased Go/NoGo responding in the learning phase, such that subjects made more Go responses when playing for reward (Win cues) than when trying to avoid punishment (Avoid cues). This motivational bias in Go/NoGo responding was not significantly reduced in the carriers relative to the controls. In contrast, the carriers exhibited reduced impact of the valence on their subjective cue valuations during the transfer phase, specifically in the context of incongruency between the action requirements and the valence (i.e. NoGo-to-Win and Go-to-Avoid cues). These results suggest that the subjective valuation is altered in carriers of pathogenic variants in the TH genes, potentially due to catecholamine-dependent changes in reward expectations, whereas instrumental task performance appears unaffected. This pilot study provides a first insight into the cognitive consequences of carrying pathogenic variants in the TH gene, focusing on alterations in motivational biases in action and the reward valuation system.

Distinct Pavlovian and instrumental catecholaminergic mechanisms drive the motivational biases in action

Despite the absence of significant group differences in the motivational bias in **chapter 2**, catecholamines are generally known to modulate the impact of motivational cues on action (Taylor and Robbins, 1986, 1984). Such motivational biases have been proposed to reflect cue-based, 'Pavlovian' effects. In **chapter 3**, we assessed whether motivational biases may also arise from asymmetrical instrumental *learning* of active, Go and passive, NoGo responses following reward and punishment outcomes. Given the strong link between catecholamine transmission and motivated action, we also aimed to assess the effect of a catecholaminergic manipulation on these biases. To this end, a large sample of participants (N=106) performed a motivational Go/NoGo learning task twice, once under a catecholamine challenge (methylphenidate - MPH) and once on placebo. Based on previous literature of dopaminergic drug effects (Cools & D'Esposito, 2011, and Frank & Fossella, 2011 for reviews), we hypothesized that MPH effects on motivated action would covary with measures scaling with baseline dopamine function, namely working memory span (Cools et al., 2008) and trait impulsivity (Buckholtz et al., 2010). We presented an extended version of the motivational Go/NoGo task, including multiple Go response options, which allowed us to disentangle the impact of reward and punishment outcomes on the instrumental learning of selective responses from non-selective, Pavlovian response biasing. Computational analyses showed that motivational biases in Go/NoGo responding reflect both Pavlovian and instrumental effects: reward and punishment *cues* promoted generalized (in)action in a Pavlovian manner, whereas reward and punishment *outcomes* enhanced instrumental learning and unlearning of chosen actions. We replicate the presence of these cue- and outcome-based biases in **chapter 4**. These cue- and outcome-based biases were altered independently by the catecholamine enhancer melthylphenidate. Specifically, melthylphenidate modulated Pavlovian response biasing and altered the reward-

driven diffusion of credit assignment during instrumental learning. Methylphenidate's effect varied across individuals with a putative proxy of baseline dopamine synthesis capacity, working memory span, but did not significantly covary with trait impulsivity. Our study uncovers two distinct mechanisms by which motivation impacts behaviour, and helps refine current models of catecholaminergic modulation of motivated action.

Deep brain stimulation of the ventral striatum releases aversive inhibition of behaviour

The catecholaminergic modulation of motivated action is theorized to rely on dopamine's function in the (ventral) striatum (Dickinson et al., 2000; Hebart and Gläscher, 2015; Lex and Hauber, 2008; Taylor and Robbins, 1986, 1984; Wyvell and Berridge, 2000). In **chapter 5**, we assessed the causal role of the ventral striatum in the motivational biasing of action for the first time in humans by directly stimulating the nucleus accumbens with deep brain stimulation (DBS). In this study, 8 treatment-refractory obsessive-compulsive disorder patients performed the motivational Go/NoGo task with concurrent EEG recordings. During this task, DBS was switched ON vs. OFF in a cross-over within-subject design. As in all other chapters, the cue valences had pronounced impact on the number of active Go responses and the associated reaction times, such that subjects made fewer and slower Go responses to punishment ('Avoid') cues than to reward ('Win') cues. DBS of the bilateral ventral striatum attenuated this inhibitory influence of punishment cues on reaction times; subjects made relatively faster Go responses to the punishment cues during DBS. The corresponding attenuation in the proportion of Go responses was highly consistent in direction, yet the medium effect size did not result in statistical significance across the group. Rather did DBS seem to attenuate the inhibitory influence of punishment cues on the proportion Go responses proportional to the estimated DBS impact on the nucleus accumbens. These results causally implicate the human ventral striatum in the coupling of motivation and behavioural activation.

Midfrontal network dynamics reflect control over maladaptive motivational biases

In all chapters, the motivational biasing of behaviour was highly consistent: Reward biases towards action, punishment towards inaction. In **chapter 3**, we established that motivation exerts control over behaviour at least partly by eliciting Pavlovian responses, which can either match or conflict with instrumental action. We can overcome maladaptive motivational influences, putatively through frontal cognitive control (Cavanagh et al., 2013). However, the neurocomputational mechanisms subserving this control are unclear; does control entail upregulating instrumental systems, downregulating Pavlovian systems, or both? In **chapter 4**, we combined EEG recordings with the extended motivational Go/NoGo learning task (N=34), where multiple Go options enabled us to disentangle selective action learning from non-selective Pavlovian responses. In **chapter 4**, we replicated our **chapter 3** finding that both Pavlovian and instrumental learning mechanisms contribute to the motivational

biasing of action, coupling action to reward and inaction to punishment. Midfrontal theta-band (4-8Hz) activity covaried with the trial-by-trial level of conflict between the prepotent, Pavlovian response tendencies and the learned instrumental responses. This conflict-related theta signal was associated with reduced Pavlovian biases, rather than reduced instrumental learning biases or enhanced specific instrumental responses. This conflict-related theta signal was accompanied by phase synchronization of the lateral prefrontal and motor sites to the midfrontal site, and these network dynamics predicted the reduction of the Pavlovian biases over and above the local, midfrontal power. This work links midfrontal processing to detecting Pavlovian conflict, and highlights the importance of network processing in reducing the impact of maladaptive, Pavlovian biases. Moreover, this chapter presents the first work incorporating trial-by-trial phase synchronization in formal learning models, providing a proof of principle for the hypothesis-driven integration of dynamics EEG connectivity measures in computational models of behaviour.

Deep brain stimulation of the ventral striatum does not affect conflict-related midfrontal oscillatory dynamics

In the patient DBS study reported in **chapter 5**, we also recorded EEG during task performance and assessed if DBS of the bilateral nucleus accumbens altered the conflict-related neural responses that we observed in **chapter 4**. Given that the behavioural effect of DBS appeared specific to the Go cues and the number of correct NoGo-to-Win trials was insufficient, we restricted the EEG analysis to the Go cues. Independent of DBS, oscillatory theta activity increased over the midfrontal cortex for the cues where the Pavlovian response tendencies conflicted with the instrumental requirements (Go-to-Avoid > Go-to-Win), consistent with our **chapter 4** findings. DBS did not alter the putative conflict-related midfrontal theta responses across the group. These results suggest that DBS of the bilateral nucleus accumbens locally attenuated the motivational biases, rather than enhancing (distal) frontal control systems.

Interpretation of the findings

The motivational biasing of action is a highly robust phenomenon as demonstrated by the multiple experiments in this thesis. I assessed a range of populations (healthy students, obsessive-compulsive disorder patients, carriers of pathogenic variants in the TH genes, and a heterogeneous control population). Although there was considerable variability between subjects in the strength of the motivational biases, all tested populations consistently showed the motivational biases at least to some extent. Moreover, we consistently observed effects of motivational valence on both the proportion of Go responses and reaction times across different versions of the task (e.g., implicit vs. explicit cue valences, single vs. multiple Go response options, 70-80% feedback contingencies). These motivational effects on choice and

RT were complementary in nature (in contrast to a trade-off), as subjects made both *faster* and *more* Go responses when playing for reward than when preventing punishment. In sum, the motivational Go/NoGo task proved to be a robust tool to uncover these persistent biases in action.

Several of our findings were in line with a role for striatal dopamine in the motivational biasing of action. Based on prior evidence from theoretical and empirical work (Collins and Frank, 2014) indicating that bursts in striatal dopamine release elicited by reward cues potentiate the basal ganglia direct Go pathway, we predicted that reward cues would facilitate non-selective behavioural activation. Additionally, based on the hypothesis that bursts in striatal dopamine release elicited by reward outcomes potentiate recently activated connections in the same direct Go pathway, we put forward the novel hypothesis that reward outcomes would facilitate selective learning of the performed Go response. Comparable, yet opposite biases were hypothesised for punishment cues and outcomes, coupling punishment to inactive, NoGo responses. In **chapter 3** and **chapter 4** we found the first behavioural evidence for biased (outcome-based) instrumental learning in addition to the (cue-based) Pavlovian response biasing. In the other chapters we did not disentangle these cue- and outcome-based biases, as the paradigms did not include multiple Go response options (making it harder to disentangle selective instrumental learning from non-selective Pavlovian activation). Thus, we should be cautious to attribute the observed motivational biases in **chapter 2** and **chapter 5** to either Pavlovian or instrumental learning biases alone. However, in most experiments we made the cue valences explicit by giving the cues a green border (Win cues) or a red border (Avoid cues), and with these explicit cue valences we always observed motivational biases already on the first trial. Therefore, we can be quite confident that the motivational biases in **chapter 2** and **chapter 5** rely at least partly on Pavlovian mechanisms, but we should be cautious to either exclude (or take for granted) the contribution of instrumental learning biases. Altogether, the consistent observations of motivational biases in action, and the complementary cue- and outcome-based mechanisms, supported our novel behavioural hypotheses based on known striatal dopamine function.

Next to establishing the complementary contribution of Pavlovian and instrumental mechanisms in the motivational biases, other findings supported current views of striatal dopamine function described above and throughout this thesis. In **chapter 3** we demonstrated that challenging the dopamine system with the dopamine reuptake blocker methylphenidate modulated both the Pavlovian and instrumental biases, thereby causally linking catecholamine function to these Pavlovian and instrumental mechanisms. Furthermore, in **chapter 5** we showed that disrupting ventral striatal processing with deep brain stimulation reduced the behavioural inhibition elicited by aversive cues. The effect of striatal DBS was remarkably consistent for reaction times and the proportion Go responses, but did not reach significance for the Go responses in the current sample. These findings from **chapter 3** and **chapter 5** provided

causal evidence for the role of dopamine and the ventral striatum in the instrumental and Pavlovian biases, yet in **chapter 2** we did not observe a significant group difference between dopamine-related genetic carriers of pathogenic TH variants and matched controls. Note, however, that there was a small effect size for reduced motivational biases on the NoGo cues in the carriers, so we also cannot conclusively infer an absence of group differences. The carriers did express a reduced influence of the cue valences on their subjective cue preferences in the transfer phase, linking the pathogenic TH variants to subtle alterations in subjective valuation, in line with dopamine's theorized role in reward expectations. This pattern of findings might indicate that the transfer phase is more sensitive to subtle changes in the dopamine system, potentially reflecting the accumulated dopamine-dependent learning from reward prediction errors. In contrast, subjects might be able to compensate in the learning phase by recruiting working memory or additional prefrontal functions, particularly in the simple task version with four cues, making the learning phase relatively less sensitive to subtle, chronic changes in the dopamine system. Such relative task sensitivities would explain why the carriers of pathogenic TH variants showed altered subjective cue preferences, but unaffected motivational biases in action (**chapter 2**). Moreover, methylphenidate affected the motivational biases proportional to the putative dopamine-proxy working memory capacity, whereas this working memory dependency was absent under placebo (**chapter 3**), possibly illustrating the lower sensitivity to chronic states and higher sensitivity to pharmacological challenges. Unfortunately, our pharmacological study did not contain a transfer phase, impeding us to directly compare the dopamine sensitivity of the learning and transfer phase within a healthy population.

Finally, I assessed the role of the medial frontal cortex (MFC) in reducing maladaptive motivational biases. In **chapter 4** we particularly linked MFC theta activity to the amount of conflict between the Pavlovian and instrumental system, suggesting a role for the MFC in the detection of conflict between the response systems. In **chapter 5**, we replicated part of these findings by observing increased MFC theta responses to Pavlovian conflict (on the Go cues) within the sample of DBS patients. Yet, the reduction in motivational response tendencies under DBS was not accompanied by altered MFC theta signals, resembling the unaffected MFC theta responses under STN DBS in the presence of behavioural changes in cognitive control (Cavanagh et al., 2011). These combined findings were strikingly consistent with my previous work (van Driel et al., 2015), linking MFC theta responses to the anticipation and detection of classic response conflict (i.e. where stimulus-induced response tendencies conflict with instrumental task requirements). Thus, the MFC might be responsible for detecting conflict between multiple activated response systems, irrespective of the identity (e.g., motivational, affective, instrumental) of the sources that are causing the conflict. Our results showed that these MFC theta responses were specifically related to the reduction of maladaptive Pavlovian biases, and did not seem to relate to changes in the instrumental system. Thus, we reasoned that the MFC might detect conflict in order to halt prepotent responses, potentially by raising

the decision threshold , and allow for the slower, more adaptive systems to take over (Aron et al., 2016; Cavanagh et al., 2011; Cavanagh and Frank, 2014; Cohen, 2014; Frank et al., 2015). Motor and lateral prefrontal sites synchronized with the MFC, which was even more directly related to the reduction of maladaptive Pavlovian biases, suggesting a role for network-wide communication in the regulation of maladaptive response tendencies. We could not link the MFC theta responses to (reductions in) the instrumental learning biases, which might be because these learning biases do not conflict with other response systems. At least, it seems redundant to assume that we have both a biased and unbiased instrumental system, which would independently track learned action values and would then conflict with each other. Taken together, the MFC does not seem to reduce motivational biases per se, but rather detects and signals conflict between activated response systems in order to allow for adaptive systems to take control.

Limitations

The work described in this thesis has provided valuable insight into the neurocomputational mechanisms subserving the motivational biasing of action on the one hand, and the resolution of maladaptive biases on the other hand. These studies, however, also had limitations that should be considered when interpreting the work.

First of all, the specificity of our findings: I linked the results from the TH genetic study (**chapter 2**) and the methylphenidate study (**chapter 3**) to theoretical views of dopamine function, yet other catecholamines (i.e. noradrenaline and adrenaline) might have played a role as well, as the TH enzyme and methylphenidate non-specifically affect the catecholamine system. Moreover, methylphenidate was administered systemically and was not restricted to the striatum, nor are the consequences of the TH pathogenic variants. Similarly, I linked our deep brain stimulation findings (**chapter 5**) to ventral striatal functioning, yet other structures might be also modulated by the stimulation (e.g., surrounding subcortical structures, or cortical structures via passing white matter tracts).

Second, the TH genetic study (**chapter 2**) was observational in the sense that we did not randomly assign subjects to the groups. This observational nature is inherent to human genetics studies, yet consequently the groups might differ on other relevant factors, which might also explain the observed group differences. I assumed that the groups would differ in dopamine synthesis as a direct consequence of the pathogenic genetic variant, and that the observations could be attributed to the altered dopamine function. However, we did not measure dopamine synthesis, and therefore the extent of dopamine reduction in the carriers is unknown.

Third, DBS provides a valuable tool for assessing the causal involvement of neural structures, but for ethical reasons we cannot employ DBS in healthy populations. Therefore,

the DBS study (**chapter 5**) involved a psychiatric population, and we do not know to what extent the mechanisms of actions are similar in obsessive-compulsive patients as in a healthy population.

Finally, the localization of our EEG results (**chapter 4** and **chapter 5**) is limited, and future studies (for example employing fMRI) are needed to provide more confidence in the localization of the EEG results. To address the localization of our EEG findings and assess striatal signals in a healthy popilation, we have set up a follow-up study (not covered in this thesis) in which we record simultaneous EEG and fMRI.

Future outlook

The work reported in this thesis presents several research opportunities for future neuroscientific studies. First, having established the complementary contribution of Pavlovian and instrumental mechanisms to the motivational biases, a first step would be to further disentangle the neural underpinnings: Do the same neural structures underpin the Pavlovian and instrumental mechanisms, or do they rely on different neural pathways? Are these neural structures purely situated in the basal ganglia? Second, the relation between network-wide connectivity and reduced maladaptive Pavlovian biases invites us to optimize the localization of the results: Does the lateral prefrontal connectivity reflect connectivity with the dorsolateral prefrontal cortex? To what extent is the connectivity restricted to the lateral prefrontal and motor sites? Does the MFC also signal to subcortical sites? Finally, and perhaps most exciting, how do the different structures interact with each other at the neural level? Does the MFC receive input from the basal ganglia and only signal conflict to other task-relevant regions, or does the MFC also modulate activity within the basal ganglia? How is control subsequently implemented in the target structures, i.e. what do the specific computational implementations look like? We have started to address these questions in our combined EEG-fMRI study, and with these questions I hope to inspire and invite other researchers to follow up on our studies and extend our line of work.

On a higher level, the gained insights into the biased instrumental learning system has potential implications for behavioural modification, relevant to e.g. psychological treatments and pedagogy. For example, it might be more effective to motivate students to study for good grades (positive outcomes) than for avoiding failure (avoiding negative outcomes). Such framing effects are known to have strong impact on behaviour (Kahneman and Tversky, 1979), and our work refines these framing effects by showing the reliance on active vs. inactive behaviour. Moreover, future research could assess the implementation in psychological treatment for example of addiction behaviour. Here our findings suggest, albeit very speculatively, that it might be more effective to encourage someone to take *alternative actions* (in order to get more positive outcomes), rather than to encourage someone to refrain from

their addiction behaviour per se, as we might be better in attributing the positive outcomes to our alternative behaviours than to refraining from our addiction behaviour. Here our findings are of particular relevance, as so far it has been assumed that our motivational biases purely arise from a Pavlovian system, whereas we have demonstrated that classes of (in)action are better learned from certain outcomes.

Finally, the theorized role of the catecholamine system and of frontostriatal functioning might help to improve our clinical understanding. That is, the reliance of the motivational biases on the catecholamine system can help to better understand, and potentially even predict, cognitive side effects of dopaminergic treatment. Our pharmacological study thereby emphasizes once more the importance of taking into account individual differences in pharmacological treatment. Similarly, the reliance of the biases on frontostriatal functioning is relevant for a better understanding of psychiatric disorders in which frontostriatal functioning is affected. Moreover, it has been suggested that either too much or too little Pavlovian biases are at the heart of a number of psychiatric disorders (Garbusow et al., 2016; Geurts et al., *in prep.*; Heinz et al., 2016; Watson et al., 2014), where our findings further suggest that these symptoms could arise from too much or too little instrumental biases as well.

Concluding remarks

Motivation is a hallmark of human cognition and throughout this thesis I have shown that motivation strongly drives our behaviour. In many cases motivation is beneficial for the realization of our goals, as it generally facilitates the behaviour that is required for achieving those goals. In this work I addressed the neurocomputational mechanisms that underlie how our motivation can irrationally drive our actions on the one hand, and the neural mechanisms that allow us to reduce the impact of irrational motivational drives when these drives become dysfunctional. Here, I demonstrated for the first time the dissociable contribution of Pavlovian and instrumental mechanisms to the well-established motivational biases in action that couple reward to active responses and punishment to inaction. I theoretically linked both these Pavlovian and instrumental mechanisms to striatal dopamine function, and established the susceptibility of these mechanisms to a catecholamine challenge (methylphenidate) and disruption of the striatum (deep brain stimulation). I related the medial frontal cortex particularly to reduced impact of the Pavlovian biases, but not the instrumental learning biases, when the motivational biases became dysfunctional. I propose that the medial frontal cortex serves as a hub, detecting and signalling conflict between the Pavlovian and instrumental system to recruit control over the dysfunctional Pavlovian biases. To conclude, with this work I have shed light on neurocomputational mechanisms related to acquiring and overcoming motivational biases in action.

References

Aron, A.R., Herz, D.M., Brown, P., Forstmann, B.U., Zaghloul, K., 2016. Frontosubthalamic Circuits for Control of Action and Cognition. J. Neurosci. 36, 11489–11495. doi:10.1523/JNEUROSCI.2348-16.2016

Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Res. Rev. doi:10.1016/S0165-0173(98)00019-8

Brozoski, T., Brown, R., Rosvold, H., Goldman, P., 1979. Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. Science (80-.). 205, 929–932. doi:10.1126/science.112679

Buckholtz, J.W., Treadway, M.T., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., Baldwin, R.M., Schwartzman, A.N., Shelby, E.S., Smith, C.E., Kessler, R.M., Zald, D.H., 2010. Dopaminergic Network Differences in Human Impulsivity. Science (80-.). 329, 532. doi:10.1126/science.1185778

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., Frank, M.J., 2013. Frontal Theta Overrides Pavlovian Learning Biases. J. Neurosci. 33, 8541–8548. doi:10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. doi:10.1016/j.tics.2014.04.012

Cavanagh, J.F., Wiecki, T. V, Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., Frank, M.J., 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat. Neurosci. 14, 1462–1467. doi:10.1038/nn.2925

Cohen, M.X., 2014. A neural microcircuit for cognitive conflict detection and signaling. Trends Neurosci. doi:10.1016/j.tins.2014.06.004

Collins, A.G.E., Frank, M.J., 2014. Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. 121, 337–66. doi:10.1037/a0037015

Cools, R., D'Esposito, M., 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol. Psychiatry. doi:10.1016/j.biopsych.2011.03.028

Cools, R., Frank, M.J., Gibbs, S.E., Miyakawa, A., Jagust, W., D'Esposito, M., 2009. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J. Neurosci. 29, 1538–43. doi:10.1523/JNEUROSCI.4467-08.2009

Cools, R., Gibbs, S.E., Miyakawa, a, Jagust, W., D'Esposito, M., 2008. Working memory capacity predicts dopamine synthesis capacity in the human striatum. J.Neurosci. 28, 1208–1212. doi:10.1523/JNEUROSCI.4475-07.2008

Dickinson, A., Smith, J., Mirenowicz, J., 2000. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav. Neurosci. 114, 468–483. doi:10.1037/0735-7044.114.3.468

Frank, M.J., Fossella, J. a, 2011. Neurogenetics and pharmacology of learning, motivation, and cognition. Neuropsychopharmacology 36, 133–152. doi:10.1038/npp.2010.96

Frank, M.J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J.F., Badre, D., 2015. fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. J. Neurosci. 35, 485–494. doi:10.1523/JNEUROSCI.2036-14.2015

Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science (80-.). 306, 1940–1943. doi:10.1126/science.1102941

Garbusow, M., Schad, D.J., Sebold, M., Friedel, E., Bernhardt, N., Koch, S.P., Steinacher, B., Kathmann, N., Geurts, D.E.M., Sommer, C., Müller, D.K., Nebe, S., Paul, S., Wittchen, H.U., Zimmermann, U.S., Walter, H., Smolka, M.N., Sterzer, P., Rapp, M.A., Huys, Q.J.M., Schlagenhauf, F., Heinz, A., 2016. Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. Addict. Biol. 21, 719–731. doi:10.1111/adb.12243

Geurts, D.E.M., den Ouden, H.E.M., Janssen, L., Swart, J.C., Froböse, M., Cools, R., Speckens, A.E.., n.d. Mindfulness-based cognitive therapy enhances behavioural inhibition to aversive Pavlovian stimuli in patients with ADHD.

Hebart, M.N., Gläscher, J., 2015. Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. Psychopharmacology (Berl). 232, 437–451. doi:10.1007/s00213-014-3682-3

Heinz, A., Schlagenhauf, F., Beck, A., Wackerhagen, C., 2016. Dimensional psychiatry: mental disorders as dysfunctions of basic learning mechanisms. J. Neural Transm. doi:10.1007/s00702-016-1561-2

Kahneman, D., Tversky, A., 1979. Kahneman & Tversky (1979) - Prospect Theory - An Analysis Of Decision Under Risk. Econometrica 47, 263–292. doi:10.2307/1914185

Kurian, M.A., Gissen, P., Smith, M., Heales, S.J.R., Clayton, P.T., 2011. The monoamine neurotransmitter disorders: An expanding range of neurological syndromes. Lancet Neurol. doi:10.1016/S1474-4422(11)70141-7

Lex, A., Hauber, W., 2008. Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. Learn. Mem. 15, 483–491. doi:10.1101/lm.978708

Robbins, T.W., Everitt, B.J., 2007. A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). Psychopharmacology (Berl). doi:10.1007/s00213-006-0528-7

Salamone, J.D., Correa, M., Mingote, S.M., Weber, S.M., 2005. Beyond the reward hypothesis: Alternative functions of nucleus accumbens dopamine. Curr. Opin. Pharmacol. doi:10.1016/j.coph.2004.09.004

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science (80-.). 275, 1593–1599. doi:10.1126/science.275.5306.1593

Taylor, J.R., Robbins, T.W., 1986. 6-Hydroxydopamine lesions of the nucleus accumbens, but not of the caudate nucleus, attenuate enhanced responding with reward-related stimuli produced by intra-accumbens d-amphetamine. Psychopharmacology (Berl). 90, 390–397. doi:10.1007/BF00179197

Taylor, J.R., Robbins, T.W., 1984. Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. Psychopharmacology (Berl). 84, 405–412. doi:10.1007/BF00555222

van Driel, J., Swart, J.C., Egner, T., Ridderinkhof, K.R., Cohen, M.X., 2015. (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. Cogn. Affect. Behav. Neurosci. 15, 787–807. doi:10.3758/s13415-015-0367-2

Watson, P., Wiers, R.W., Hommel, B., De Wit, S., 2014. Working for food you don't desire. Cues interfere with goal-directed food-seeking. Appetite 79, 139–148. doi:10.1016/j.appet.2014.04.005

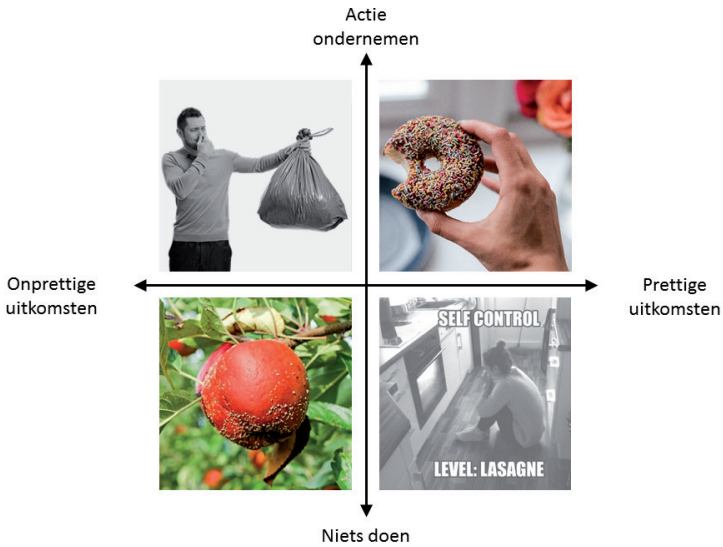
Wyvell, C.L., Berridge, K.C., 2000. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. J. Neurosci. 20, 8122–8130. doi:http://www.jneurosci.org/content/21/19/7831

Appendix

Nederlandse samenvatting

Inleiding

We maken elke dag enorm veel keuzes. Volgens Google maar liefst 35.000! Dit zijn natuurlijk niet allemaal levensbepalende keuzes als "*Wat voor werk zal ik doen na mijn promotie?*", maar ook simpelere keuzes als "*Zal ik wel of geen koffie nemen?*". We kunnen deze keuzes niet allemaal overdenken en zullen soms (onbewust) moeten terugvallen op vuistregels die de keuzes makkelijker voor ons maken. Eén zo'n belangrijke vuistregel maakt gebruik van onze verwachtingen: Zo zijn we geneigd actie te ondernemen als we een prettige uitkomst verwachten, terwijl we juist terughoudend zijn als we een onprettige uitkomst verwachten (zie Figuur 1). Deze vuistregel is vaak heel handig en helpt ons meestal goede keuzes te maken. Denk bijvoorbeeld aan het opeten (*actie ondernemen*) van appetijtelijk eten (*prettige uitkomst*) en het laten staan (*terughoudend zijn*) van bedorven voedsel (*onprettige uitkomst*). Deze vuistregel is echter niet altijd handig. Soms moeten we bijvoorbeeld wachten met eten (*ons inhouden*) als de maaltijd nog in de oven staat en kunnen we het bedorven voedsel maar beter weggooien (*actie ondernemen*). In deze gevallen werkt onze vuistregel ons dus tegen. In mijn proefschrift heb ik enerzijds onderzocht hoe onze hersenen ervoor zorgen dat deze specifieke vuistregel tot stand komt en anderzijds hoe we deze vuistregel kunnen onderdrukken wanneer deze ons tegenwerkt.

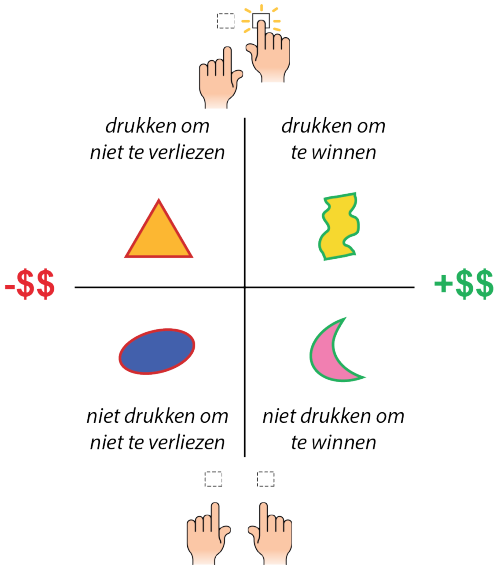


Figuur 1. **(On)pretttige verwachtingen sturen de activatie van ons gedrag.**
Onze verwachtingen over prettige vs. onprettige uitkomsten hebben een tegenovergestelde werking op de activatie van ons gedrag. Zo zetten prettige verwachtingen ons veelal aan tot het nemen van actie en zo laten we ons door negatieve verwachtingen juist vaak inhouden. Deze vuistregel maakt het makkelijk om bijvoorbeeld appetijtelijk voedsel op te zoeken en weg te blijven van bedorven voedsel. Deze

makkelijke situaties zijn boven in kleur weergegeven. Het koppelen van actie aan prettige uitkomsten en nietsdoen aan onprettige uitkomsten is vaak heel nuttig, maar in sommige gevallen kunnen we beter het tegenovergestelde gedrag vertonen. Zo kan je soms een vies huis voorkomen door het bedorven voedsel weg te gooien en wordt een lasagne nog lekkerder als je voldoende lang wacht. Onze vuistregel maakt deze situaties moeilijker (aangegeven in zwart-wit).

Het onderzoek

Om de vuistregel te onderzoeken, heb ik met mijn collega's een computerspel ontwikkeld waarin we bovenstaande situaties in een spelvorm nabootsen (zie Figuur 2). In dit computerspel tonen we afwisselend positieve afbeeldingen die aangeven dat de speler geld kan winnen (*prettige uitkomst*) en negatieve afbeeldingen die aangeven dat de speler geld kan verliezen (*onprettige uitkomst*). Voor de helft van de positieve afbeeldingen moet de speler op een toets drukken (*actie ondernemen*) om het geld daadwerkelijk te krijgen, terwijl de speler voor de andere positieve afbeeldingen juist niet moet drukken (*zich inhouden*) om het geld te krijgen. Ditzelfde principe geldt ook voor de negatieve afbeeldingen: Voor de helft van de negatieve afbeeldingen moet de speler op een toets drukken om *geen geld te verliezen*, oftewel om de onprettige uitkomst te *voorkomen*, terwijl de speler voor de andere negatieve afbeeldingen juist niet moet drukken om niet te verliezen. We vertellen de spelers niet wat zij idealiter voor elke afbeelding zouden moeten doen en dus moeten de spelers er tijdens het spel zelf achter komen wat de beste reacties zijn.



Figuur 2. **Het computerspel.**
In ons computerspel krijgen deelnemers verschillende positieve en negatieve afbeeldingen te zien waarvoor ze óf op een knop moeten drukken (actie ondernemen) óf niets moeten doen (zich inhouden). Voor de positieve afbeeldingen (met groene rand) kunnen de deelnemers geld winnen als ze de juiste reactie geven, terwijl ze voor de negatieve afbeeldingen (met rode rand) dan juist voorkomen dat ze geld

verliezen. De deelnemers moeten tijdens het spel zelf leren wat de beste reactie is voor elke afbeelding. Hoe beter de deelnemers het spel spelen, hoe meer geld ze winnen. Met dit spel kunnen we de vuistregel meten, oftewel in hoeverre de deelnemers geneigd zijn om actie te ondernemen als ze geld kunnen verdienen en zich inhouden als ze geld kunnen verliezen.

Ik heb dit computerspel gebruikt in vier verschillende onderzoeken die ik rapporteer in hoofdstuk 2 tot en met 5 van dit proefschrift. In elk van deze onderzoeken zagen we dat de deelnemers geneigd waren om actie te ondernemen voor de positieve afbeeldingen en zich juist inhielden voor de negatieve afbeeldingen. Voor de helft van de positieve en negatieve afbeeldingen hielpen deze neigingen de deelnemers dus inderdaad om de beste keuze te maken. Voor de andere afbeeldingen zorgden de neigingen er echter voor dat de deelnemers juist minder goede keuzes maakten. Met behulp van ons computerspel konden we de vuistregel dus blootleggen en verder onderzoeken. In de volgende sectie bespreek ik beknopt en per overkoepelend onderwerp wat deze onderzoeken inhielden en wat de bijbehorende bevindingen waren.

Bevindingen

Automatisme of aangeleerd?

In dit proefschrift hebben we onderzocht of de vuistregel berust op een automatisch proces of dat de vuistregel ook kan ontstaan door een leerproces (hoofdstuk 3 en 4). Veelal werd er namelijk vanuit gegaan dat de vuistregel op een automatisch proces berust; dat de vuistregel als het ware zo is vastgelegd in onze hersenen. Mijn collega's en ik redeneerden echter dat de vuistregel ook zou kunnen ontstaan doordat mensen verschillend leren van prettige en onprettige uitkomsten. Om deze verschillende verklaringen te kunnen onderscheiden en toetsen, ontwikkelden we wiskundige modellen. Uit deze modellen bleek dat de vuistregel voor een deel automatisch gebruikt werd, dus zelfs al wanneer de deelnemers de afbeeldingen voor de eerste keer zagen. Met behulp van de modellen toonden we echter voor het eerst aan dat de deelnemers daarnaast inderdaad verschillend leerden van de uitkomsten. Zo schreven de deelnemers positieve uitkomsten sneller toe aan hun acties en waren ze juist minder goed in het wijten van negatieve uitkomsten aan hun inactiviteit. Hierdoor hielden deelnemers zich vaker in ook al hadden zij hierdoor eerder geld verloren. Uit deze bevindingen konden we concluderen dat de vuistregel dus deels over tijd wordt aangeleerd, wat een vernieuwend inzicht in het veld was.

Neurochemische modulatie van de vuistregel

In hoofdstuk 2 en 3 hebben we het verband onderzocht tussen de vuistregel en een groep chemische stoffen die worden gebruikt voor communicatie in de hersenen: Catecholamines. Catecholamines zijn belangrijk voor het verwerken van beloningen, maar ook voor het

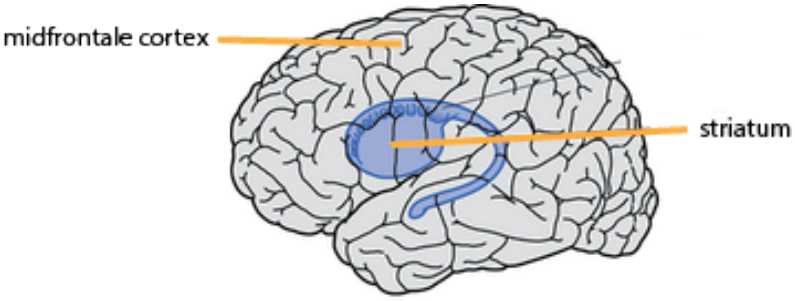
ondernemen van actie. Allereerst onderzochten we het verband tussen catecholamines en de vuistregel door de catecholamine-huishouding te veranderen (hoofdstuk 3). Hiervoor lieten we dezelfde gezonde studenten eenmaal ons computerspel spelen na het innemen van catecholamine-beïnvloedende medicatie en eenmaal na placebomedicatie. We vergeleken vervolgens hun gedrag na de catecholamine-beïnvloedende medicatie met hun eigen gedrag na placebomedicatie. Zo zagen we dat de catecholamine-beïnvloedende medicatie niet bij iedereen hetzelfde effect had. Het effect hing af van een maat die samenhangt met de natuurlijke catecholamine aanmaak, namelijk het kortetermijngeheugen ('werkgeheugen'). Bij deelnemers met een goed kortetermijngeheugen, die veelal van nature meer catecholamines aanmaken, versterkte de medicatie de vuistregel; zij gingen vaker actie ondernemen als ze geld konden winnen en deden vaker niets als ze geld konden verliezen. Deze vuistregel was dus nuttig voor de helft van de afbeeldingen, maar werkte tegen voor de andere helft van de afbeeldingen. Bij de deelnemers met een minder goed kortetermijngeheugen, die veelal minder catecholamines aanmaken, had de catecholamine-beïnvloedende medicatie het tegenovergestelde effect en verzwakte de invloed van de vuistregel; zij deden vaker niets als ze geld konden winnen en gingen vaker actie ondernemen als ze geld konden verliezen. Deze bevindingen suggereerden dus dat het effect van de catecholamine-beïnvloedende medicatie op de vuistregel afhing van de natuurlijke catecholamine aanmaak.

Vervolgens onderzochten we het verband tussen catecholamines en de vuistregel ook door een specifieke genetische testgroep te vergelijken met een controlegroep (hoofdstuk 2). De genetische testgroep bestond uit mensen die een genetische samenstelling hebben die de aanmaak van catecholamines beperkt. De controlegroep bestond uit een vergelijkbare groep zonder afwijkende genetische samenstelling. Tegen onze verwachtingen in, zagen we geen aantoonbare verschillen tussen de testgroep en de controlegroep in de mate waarop zij de vuistregel uitten. Samengenomen hebben we het verband tussen catecholamines en de vuistregel dus met name kunnen laten zien nadat mensen catecholamine-beïnvloedende medicatie kregen, maar niet als zij geen medicatie kregen.

Hersengebieden en de vuistregel

Catecholamines zijn onder meer werkzaam in het striatum, een evolutionair gezien 'oud' hersengebied (zie Figuur 3). In dit hersengebied worden beloningen verwerkt en vindt ook het selecteren van acties plaats. We denken daarom dat catecholamines de vuistregel beïnvloeden via het striatum. We onderzochten de relatie tussen het striatum en de vuistregel in hoofdstuk 5. Hiervoor hadden we een bijzondere onderzoeksgroep, namelijk Spaanse patiënten die diepe hersenstimulatie in het striatum kregen als behandeling tegen obsessief-compulsieve stoornis. Deze patiënten speelden ook ons computerspel terwijl ze afwisselend wel of geen diepe hersenstimulatie ontvingen. Deze patiënten gebruikten dezelfde vuistregel als onze gezonde deelnemers: zij ondernamen sneller en vaker actie voor de positieve afbeeldingen en

zij ondernamen langzamer en minder actie voor de negatieve afbeeldingen. Deze vuistregel werd afgezwakt wanneer we het striatum verstoorden met de diepe hersenstimulatie: de patiënten gingen relatief sneller actie ondernemen voor de negatieve afbeeldingen wanneer dat inderdaad de bedoeling was. Ze werden dus beter in het spel doordat ze beter hun vuistregel konden negeren als dat nodig was! Dankzij deze bijzondere patiëntengroep konden we dus laten zien dat het striatum belangrijk is voor het tot stand komen van de vuistregel.



Figuur 3. **Midfrontale cortex en het striatum.**
Schematisch zijaanzicht van een brein. De linker lijn wijst de evolutionair gezien 'jongere' midfrontale cortex aan. Het blauwe gebied weerspiegelt het evolutionair gezien 'oudere' striatum, dat diep in de hersenen ligt. Afbeelding aangepast van iKnowledge (clinicalgate.com/the-basal-ganglia).

Onderdrukking van de vuistregel

Tot slot heb ik in dit proefschrift onderzocht hoe we de invloed van de vuistregel kunnen beperken wanneer deze ons niet helpt om de goede keuze te maken (hoofdstuk 4 en 5). Denk eraan, de vuistregel helpt in veel gevallen om goede keuzes te maken, maar soms is het nodig om het tegenovergestelde te doen. We hebben gekeken of in deze situaties de vuistregel werd ingeperkt door de frontale cortex (zie Figuur 3). De frontale cortex is een evolutionair gezien 'nieuw' hersengebied, dat belangrijk is voor 'hogere' cognitieve functies zoals het monitoren en beheersen van gedrag. Om de frontale cortex te onderzoeken hebben we hersenactiviteit gemeten met elektro-encefalografie (EEG) terwijl de deelnemers ons computerspel speelden. Specifieke trage hersengolven werden sterker wanneer deelnemers doorhadden dat ze hun vuistregel niet moesten volgen. We maten deze hersengolven over de midfrontale cortex, een subgebied van de frontale cortex. De midfrontale cortex bleek echter niet in isolatie te werken, maar communiceerde met andere hersengebieden die betrokken waren bij het uitvoeren van het computerspel. De mate van communicatie tussen deze hersengebieden bleek het beste te voorspellen of de deelnemers in staat waren het tegenovergestelde gedrag aan de vuistregel te vertonen wanneer dat nodig was.

Oplettende lezers zullen zich wellicht herinneren dat we in hoofdstuk 5 naast EEG ook nog diepe hersenstimulatie hebben gebruikt en zich afvragen of dat invloed had op de

midfrontale hersengolven. Dat was niet het geval. Hieruit konden we opmaken dat de diepe hersenstimulatie met name het tot stand komen van de vuistregel in het striatum verminderde en dat de diepe hersenstimulatie niet zozeer van invloed was op de frontale controlesystemen.

Conclusie

Samengevat heb ik in dit proefschrift vier onderzoeken gepresenteerd met verschillende invalshoeken naar dezelfde vraag: hoe zorgen onze hersenen ervoor dat een specifieke vuistregel tot stand komt, oftewel hoe zorgen onze hersenen ervoor dat prettige uitkomsten ons aanzetten tot actie en onprettige uitkomsten ons doen inhouden? Allereerst heb ik laten zien dat de vuistregel deels tot stand komt door leerprocessen en deels door automatische processen. Daarnaast heb ik aangetoond dat het veranderen van de catecholamine-huishouding van invloed is op de vuistregel en dat het verstoren van het striatum de vuistregel vermindert. Deze bevindingen ondersteunen dat catecholamines en het striatum ervoor zorgen dat de vuistregel tot stand komt. Anderzijds heb ik onderzocht hoe we van onze vuistregel kunnen afwijken indien het tegenovergestelde gedrag gewenster is. Hiervoor heb ik laten zien dat specifieke langzame hersengolven over de midfrontale cortex toenemen als mensen opmerken dat er een mismatch is tussen de vuistregel en het gewenste gedrag. De mate van communicatie tussen de midfrontale cortex en andere betrokken hersengebieden was vervolgens belangrijk voor het daadwerkelijk kunnen vertonen van het gewenste gedrag. Met deze inzichten hoop ik bij te dragen aan mijn wetenschappelijke veld en collega-onderzoekers te inspireren om onze onderzoekslijn te vervolgen.

Acknowledgements

Many different people have contributed to this thesis in many different ways, and the following acknowledgements will for sure not be exhaustive. I am grateful for all support – mental, conceptual and technical - that I have received during my scientific career.

First and foremost, I'd like to thank my scientific parents **Hanneke den Ouden**, **Roshan Cools**, **Michael Frank**, **Mike Cohen** and **Joram van Driel**. Thank you for believing in me, and for always being patient with me. You all have been an inspiration to me and I'm honoured to have had you as my marvelous examples. **Mike** and **Joram**, you've turned my first baby steps into a success and without a doubt I owe my career to you. **Hanneke** and **Roshan**, you have been excellent supervisors in every possible way. You've inspired, coached, mentored, supported, comforted, and sometimes frustrated, but above all, you've raised my work to another level. **Hanneke**, I'm honoured to have been your first primary PhD student, and to have grown so much together. **Michael**, you were by far the best unofficial supervisor I could ever wish for and I'm glad that I can finally call you my promotor in the end. Your range of expertise is truly unprecedented. Thank you so much for having shared that with me.

All members of the *Learning and Decision-Making lab* (**Hanneke**, **Annelies**, **Johannes**, **Vanessa**, **Jessica**, **Martin**, **Emma**, **Margot**, **Ahnjili**) and the *Motivational and Cognitive Control lab* (**Roshan**, **Monja**, **Jessica**, **Annelies**, **Dirk**, **Jennifer**, **Lieneke**, **Monique**, **Mirjam**, **Mieke**, **Joost**, **Esther**, **Iris**, **Sophie**, **Bram**, **Guillaume**, **Eliana**, **Payam**, **Andrew**, **Marpessa**, **Lieke**, **Lieke**, **Ruben**, **Danae**, **Patricia**, **Britt**, **Sean**, **Desiree**, **Daan**, **Ilke**, **Joyce**, **Martin**, **Verena**, **Romain**, **Ruth**), together you've created my scientific safe haven. Thank you for the nurturing and inspiring discussions.

Jessica, **Emma**, **Rozemarijn**, **Suzanne** and **Johannes**, thank you for bearing with me in my first steps as being a supervisor. You all did such an amazing job and I could not be more proud of you!

Of course this thesis would not have existed without my wealth of wonderful collaborators and participants. First of all, the ambitious Ritalin group (**Jennifer**, **Monja**, **Dirk**, **Michael**, **Hanneke**, **Roshan**), with whom I decided to collect over 200 drug data sets within three months of time. Special thanks to **Monique** and **Peter** for medical assistance, **Jessica** for excellent screening assistance, **Sean** for CMO advice, and most notably, **Matthew**, for removing that one comma in my code, which made all computational modeling possible! **Bryan** and **Juan**, thank you for taking me on board in the DBS studies and for welcoming me into the *Laboratory for Clinical Neuroscience*. **Javi** and **Fer**, you were my Spanish safety net. Thank you so much for the "amazing headaches", the after-hospital tapas, pinchos, and (unsuccessfully) declined shots.

You and **Lafina** have brightened up my visits to Madrid. **Svenja**, you have been an amazing addition to the team. Thank you for being so extremely helpful during the last steps of my PhD. **Ole Jensen** and **Rene Scheeringa**, thank you for sharing your EEG(-fMRI) expertise with me, and **Tessa van Wassenberg** and **Michèl Willemsen**, for introducing me into the world of genetics. All technical and administrative support (**Paul**, **Jose**, **Marek**, **Jessica**, technical and administrative support staff), you have made my life at the DCCN so very easy. The **corona** and **manuscript committee**, thank you for evaluating my work and joining me at the defense ceremony. Finally, I'd like to thank **all participants** for enduring my experiments. Particularly the **Madrid participants**, you are all so brave! I am both utterly impressed and thankful that you contributed to my work and science in general. I wish you all the best.

I have been very lucky with my roomies, who turned my office into my home. Above all, **Jill** & **Freek**, you are the best! I am going to miss you and all of our quirky habits tremendously. **Lucia**, you might not be an official roomie, but you bring life to the Donders! **Bob**, I'm glad you joined the Donders as an almost-roomie, both literally and topic-wise. **Rodolfo**, **Leticia** (or was it 'Sabine'?), & **Tim**, I enjoyed all our international analyses and discussions. **Jim**, your company was short, but much appreciated.

My old and new friends, you mean everything to me. **Paula**, I hope you will always stay my baby sister, even now we've grown up to be so 'burgerlijk'. My "intro"-friends (**Pia**, **Nicolien**, **Lesley**, **Raquel**, **Gloria**, **Lotte**, & **Carmen**), it's awesome that we've become closer every year. Let's keep the holidays, spa weekends, sinterkersts, dinners and drinks going. **Pia**, your hugs have always been there whenever I needed them. I'm thankful you're staying at my side until the end as my paranimf. **Anna** and **Nienke**, thank you for the fifthy shades of food; it's about time for another dinner-date! The old Spinoza-crew (**Lukas**, **Noor**, **Berno**, **Joanne** & **Tomas**), you've brightened up all ICON and NVP conferences. Thank you all for staying in my life after I've left Amsterdam. **Jan Harm**, **Jitske**, **the Gert-crew**, **Freek**, **Jill**, and **Monja**, you made Nijmegen my new home.

I am also very grateful for my family, who supported me every step of the way. My direct family **Pauline**, **Paul**, and **Danny**, you've supported whatever decision I've ever made. **Ronald** and **Guido**, despite how life has turned out, I will always see you as my family.

Monja, my partner in salsa and science, I'm so glad we found each other during the MPH study. I'm very thankful that I could always drop by for coffee and for all our dinner dates. Without you I would not have made it.

Angelina, words would never be enough. You are the very best of me.

List of publications

Published / in press:

Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. Swart JC, Frank MJ, Määttä JI, Jensen O, Cools R, & Den Ouden, HEM (2018). In: *PLOS Biology*, https://doi.org/10.1371/journal.pbio.2005979. IF=10.

Catecholaminergic modulation of the avoidance of cognitive control. Froböse MI, Swart JC, Cook JL, Geurts DEM, Den Ouden HEM, & Cools R (2018). In: Journal of Experimental Psychology: General, http://dx.doi.org/10.1037/xge0000523. IF=4.

Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. Swart JC, Froböse MI, Cook JL, Geurts DEM, Frank MJ, Cools R*, & Den Ouden HEM* (2017). *shared last author. In: *eLife*. https://doi.org/ 10.7554/eLife.22169. IF=8.

(No) Time for Control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. Van Driel J*, Swart JC*, Ridderinkhof KR, Egner T, & Cohen MX (2015). *shared first author. In: *Cognitive, Affective and Behavioral Neuroscience*. https://doi.org/10.3758/s13415-015-0367-2. IF=3.

Acute serotonin depletion releases motivated inhibition of response vigour. Den Ouden, HEM, Swart JC, Schmidt K, Fekkes D, Cools R (2014). In: *Psychopharmacology*. https://doi.org/10.1007/s00213-014-3762-4. IF=4.

Submitted:

Cook JL, Swart JC, Froböse MI, Diaconescu A, Geurts DEM, Den Ouden HEM, & Cools R. Catecholamine challenge uncovers distinct mechanisms for direct versus indirect, but not social versus non-social, learning. Preprint available at: https://doi.org/10.1101/303982.

Mindfulness based cognitive therapy enhances suppression of behaviour by aversive cues in patients with ADHD. Geurts DEM, Den Ouden HEM, Janssen L, Swart JC, Froböse MI, Cools R, & Speckens AEM.

About the author

Jennifer Swart was born in Zwijndrecht, the Netherlands on December 1st, 1989. She studied Psychology at the University of Amsterdam, specializing in Cognitive Neuroscience and Clinical Neuropsychology. During her bachelors, she completed the honours programme and studied abroad for one semester at the University of Melbourne, Australia. She completed the BSc programme *bene meritum* in 2011, and continued with the Research Master Psychology. Here she first joined an MRI project focusing on individual differences, and continued working as an MR operator alongside her studies. Finally, she set up an EEG study under supervision of dr. Joram van Driel and dr. Mike X Cohen, focusing on the neural signatures related to cognitive control and conflict detection. In 2013 she graduated *cum laude* and started working as a research assistant under supervision of dr. Hanneke den Ouden and prof. Roshan Cools at the Donders Institute for Brain, Cognition, and Behaviour. Here she started investigating the pharmacological underpinnings of motivational control over behaviour. Together with Hanneke and Roshan, she successfully applied for an NWO Research Talent grant, which allowed her to continue as a PhD candidate in the lab from 2014 onwards. Her PhD project focused on motivational biases in behavioural activation, with primary interest in i) basal ganglia dopamine function as a putative mechanism for the acquisition of motivational biases, and ii) midfrontal electrophysiological processes as a putative mechanism for surmounting *maladaptive* motivational biases. The results of her PhD project are described in this thesis, and were also presented at several (inter)national conferences. Additionally, she was involved in several side-projects, such as a mindfulness, pharmacological, online, and combined EEG-fMRI study, but also in the centre-wide implementation of electronic data capture software (CastorEDC), the supervision of several MSc / PhD / BSc students and research assistants, and several teaching activities. After her PhD candidacy, Jennifer continued her research career at the Central Bureau of Statistics, focusing on the Dutch healthcare.

Donders Graduate School for Cognitive Neuroscience

For a successful research Institute, it is vital to train the next generation of young scientists. To achieve this goal, the Donders Institute for Brain, Cognition and Behaviour established the Donders Graduate School for Cognitive Neuroscience (DGCN), which was officially recognised as a national graduate school in 2009. The Graduate School covers training at both Master's and PhD level and provides an excellent educational context fully aligned with the research programme of the Donders Institute.

The school successfully attracts highly talented national and international students in biology, physics, psycholinguistics, psychology, behavioral science, medicine and related disciplines. Selective admission and assessment centers guarantee the enrolment of the best and most motivated students.

The DGCN tracks the career of PhD graduates carefully. More than 50% of PhD alumni show a continuation in academia with postdoc positions at top institutes worldwide, e.g. Stanford University, University of Oxford, University of Cambridge, UCL London, MPI Leipzig, Hanyang University in South Korea, NTNU Norway, University of Illinois, North Western University, Northeastern University in Boston, ETH Zürich, University of Vienna etc.. Positions outside academia spread among the following sectors: specialists in a medical environment, mainly in genetics, geriatrics, psychiatry and neurology. Specialists in a psychological environment, e.g. as specialist in neuropsychology, psychological diagnostics or therapy. Positions in higher education as coordinators or lecturers. A smaller percentage enters business as research consultants, analysts or head of research and development. Fewer graduates stay in a research environment as lab coordinators, technical support or policy advisors. Upcoming possibilities are positions in the IT sector and management position in pharmaceutical industry. In general, the PhDs graduates almost invariably continue with high-quality positions that play an important role in our knowledge economy.

For more information on the DGCN as well as past and upcoming defenses please visit:
http://www.ru.nl/donders/graduate-school/phd/

DONDERS
INSTITUTE

ISBN 978-94-6284-173-4

Radboud University Radboudumc