Neurochemistry of reward-based decision-making

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Luca Franziska Kaiser
aus Kahl am Main (Bayern)

Düsseldorf, May 2022

aus dem Institut für Experimentelle Psychologie,
Biologische Psychologie des Entscheidungsverhaltens
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Gerhard Jocham

2. Prof. Dr. Christian Bellebaum

Tag der mündlichen Prüfung:

14.09.2022

für meinen Opa

Danksagung – Acknowledgements

An dieser Stelle bleibt mir die Möglichkeit auf meine Promotionszeit zurückzublicken. Ich bin sehr glücklich, ein solch umfangreiches Projekt mit dieser Arbeit abschließen zu können und für alle Erfahrungen und Erkenntnisse, die ich sammeln durfte, sehr dankbar. Durchaus bin ich mir bewusst, dass ich diese Zeilen ohne die Hilfe sehr vieler Menschen heute nicht schreiben könnte. Ich möchte diesen Personen im Folgenden deshalb von ganzem Herzen danken.

Ich danke meiner Familie, Valentin, Flori, Pepe, Britta, Jessi und Lena, für Eure fortwährende Unterstützung. Lieber Vali, danke, dass du mir mit deinem glücklichen Lachen jeden Tag zeigst, wie schön das Leben sein kann und mir so geholfen hast manche Herausforderungen der zurückliegenden Zeit zu meistern. Flori, ich danke dir, dass du trotz zahlreicher Umzüge und Fernbeziehungen um die ganze Welt immer an meiner Seite warst, mehr an mich glaubst als ich selbst, mich mit ungesunden Snacks versorgst und mit bayrischem Dialekt zum Lachen bringst. Mama & Papa, ich danke euch für euren immensen Support während der gesamten Zeit. Ob als wiederholte Umzugshelfer durch die gesamte Republik oder Vali-Aufpasser während wochenendlicher Lehrveranstaltungen-ihr habt mich und diese Dissertation an vielen Stellen gerettet. Jessi und Lena, ich danke euch für das kritische Lesen meiner Arbeiten, für überraschende Blumenlieferungen und dass ihr oft die Kulinarik in unserem Haushalt gerettet habt. Kurzum-ich liebe euch sehr.

Lieber Gerhard, ich danke dir sehr für die Betreuung meiner Arbeit und die andauernde Unterstützung. Deine Begeisterung für die Wissenschaft hat mich sehr beeindruckt und ich denke gerne an eine sehr unterhaltsame und lehrreiche Zeit zurück. Danke, dass du immer an mich geglaubt hast. Ich freue mich sehr von weiteren wissenschaftlichen Erkenntnissen von Dir und den CoCos zu hören!

Lieber Lennart, du standest fast meine gesamte Dissertationszeit an meiner Seite. Ich danke dir, dass ich mit dir über jedes Problem zu jeder Zeit reden konnte. Dass du zu jedem noch so abwegigen Gedanken eine schlaue Erklärung hast und vor allem auch bereit bist, diese bis ins kleinste Detail bei Hummus-Nudeln und Bier zu diskutieren. Mir bleibt nur zu sagen: „Und sorgt es auch für Augendrehn---Berufswunsch? Irgendwas mit Fame!“ Ich wünsche dir allen Fame dieser Welt auf deinem weiteren Weg-du hast es dir mehr als verdient.

Lieber Theo, auch wenn sich unsere Wege leider durch unseren Umzug nach Düsseldorf trennten, habe ich eine Menge von dir lernen können. Ich danke dir für deine Unterstützung und denke sehr gerne an unsere spannenden Diskussionen und unser gemeinsames Büro zurück. Danke, dass du in so vielen Bereichen ein offenes Ohr für uns hattest und uns so durch eine herausfordernde Zeit geholfen hast.

Dear Anne, I would like to thank you from the bottom of my heart for the great opportunity to work in your lab. I have learned a lot during this time and your enthusiasm as well as incredible and inspiring ability to approach complex scientific problems served as a great example to me. At this point, I also want to thank the University of Magdeburg: Without your funding for Women in Science, working in Berkeley would have been impossible for me.

Liebe Julia, durch unseren Umzug ist unsere wissenschaftliche Zusammenarbeit leider weniger geworden. Ich bin dir umso dankbarer, dass wir es weiterhin schaffen uns regelmäßig auszutauschen und uns zu unterstützen. Du gibst einem immer- und zwar egal wo- das Gefühl zuhause zu sein und das hat mir sehr geholfen.

Dear Mario, dear Karsta, dear Camila, dear Marina, I want to thank you for sharing many happy und unhappy PhD moments with me. Mario, your incredible diligence, huge heart and willingness to help others (also without returns) has inspired me a lot. Your amazing sense of humor and our shared love for fancy drinks has definitely made my PhD experience much better. Dear Karsta and Camila, I want to thank you for sharing your PhD moments over coffee, beer and vegan dinner at Botanica. Dear Marina, working, travelling and sharing time and thoughts with you always was great fun. Your commitment to support women in STEM is inspiring.

Liebe Renate, auch wenn wir leider keine weiteren Daten gemeinsam aufnehmen konnten, ist dein Einsatz und deine immer freundliche Hilfsbereitschaft sehr bewundernswert. Ich danke dir, dass du durch dein unglaubliches Engagement das Leben vieler Wissenschaftler um einiges besser machst. Danke Stefan und Nina, für euren Support beim Messen der MEG Daten, das geduldige Beantworten von Fragen und eure Hilfsbereitschaft.

Dear Alex, during my PhD I had to work with bash scripting for the first time. Honestly, in the beginning the whole process felt like needing to deal with a black box which can break at any time. Overcoming this fear probably would not have been possible without your ongoing support and great patience in answering every (probably not all that smart) questions. Here, I would also like to thank Adina from the bottom of my heart for the best possible and most accessible coding intro I ever had. Your enthusiasm for open science and great commitment to help others is inspiring.

Lieber Daniel, lieber Lukas, ich danke euch für zwei ersten tollen PhD Jahre, die ohne unser gemeinsames Büro wahrscheinlich nur halb so unterhaltsam gewesen wären. Ich habe sehr gerne in unserem Team zusammen gearbeitet und wünsche euch für die Zukunft das Allerbeste.

Dear Coconuts, thank you for sharing the last period of my PhD and many great memories. Anna, you thorough approach to handle scientific problems has inspired me. I am looking forward for all berries to be picked and I hope they are either sweet or good for wine! Dear Christiane, your Tante Christiane Laden helped me a lot through the last time. Thank you for your ongoing support and being the good soul of the office. Dear Monja, it was a great pleasure to share an office (or pumping room) with you. Thanks for always bringing fun and optimism to the lab. Thank you Hannah, my first office mate. Thank you for sharing your immense physical and mathematical knowledge and your amazing self-baked treats. Dear Eduard, thanks for being you and your help with continuous Linux struggles. I will miss the smell of tea and apple oatmeal for sure. Thank your Armin for walks around the park and sharing the control experience with me. Dear Mani and Lina, unfortunately the latest home-office requirements never allowed us to work together in the office and we only had a small overlap of being a coconut. I hope that your time in the lab will be as great as mine.

Danke Halla, Steffi, Christina, Georg und Joshua für eure Hilfe beim Datenerheben und Experimente vorbereiten. Liebe Steffi, liebe Halla, für euren Einsatz am Wochenende bin ich euch ganz besonders dankbar sowie für die vielen interessanten Gespräche während unserer langen Messungen.

Tani, Verena, Miri, Sascha, Jenny, Maja & Bobby: Ich danke euch, dass ihr trotz weiter Entfernungen und wenig Zeit immer an meiner Seite steht. Ich hoffe, dass ich euch auch immer ein guter Freund sein kann und ihr von mir das gleiche Maß an Unterstützung und Wertschätzung erfahrt.

Zusammenfassung

Im täglichen Leben müssen wir oft zwischen Optionen unterschiedlichen Wertes entscheiden, um unsere individuellen Vorteile und Ressourcen zu maximieren. Diese Entscheidungen können uns auf verschiedene Art und Weise begegnen. In der vorliegenden Arbeit befassen wir uns mit zwei Arten von belohnungsbasierten Entscheidungen, *patch-leaving decisions* und *value-guided choice*. Ein Beispiel für eine *patch-leaving decision* aus dem täglichen Leben ist die Entscheidung, den Arbeitsplatz zu wechseln. Hierbei muss eingeschätzt werden, wie zufrieden ich mit meiner derzeitigen Position bin. Außerdem muss der Wert möglicher Alternativangebote abgewogen werden. Ein Arbeitsplatzwechsel kann mit Nachteilen verbunden sein, beispielsweise einem aufwändigen Umzug. Während einer *patch-leaving decision* muss folglich bestimmt werden, ob die Ressourcen in der momentanen Umgebung zufriedenstellend sind, oder ob man durch einen Umgebungswechsel Vorteile erwarten kann, auch wenn dieser potentiell mit vorübergehenden notwendigen Investitionen verbunden ist. Bei der *value-guided choice* hingegen wählt der Handelnde zwischen aktuell verfügbaren Optionen. Als anknüpfendes Beispiel könnte man die Wahl einer neuen Wohnung nach einem Umzug betrachten. Hier muss aufgrund unterschiedlicher Faktoren, wie zum Beispiel der Lage und des Preises, bestimmt werden, welche die aktuell beste Option ist. Wie an diesem Beispiel zu sehen ist, setzt sich der Wert einer Option häufig aus verschiedenen Dimensionen zusammen, die integriert werden müssen, um den Gesamtwert zu bestimmen und die Optionen vergleichen zu können.

Bisherige Forschungen legen nahe, dass das menschliche belohnungsorientierte Entscheidungsverhalten von verschiedenen Neurotransmitter-systemen beeinflusst wird. In dieser Arbeit wird zunächst der Einfluss der exzitatorischen-inhibitorischen Balance (E/I Balance), bestimmt durch relative Konzentrationen des exzitatorischen Neurotransmitters Glutamat und des inhibitorischen Botenstoffes GABA, auf das Entscheidungsverhalten bei oben beschriebenen Entscheidungstypen betrachtet. Zweitens wird ein möglicher Einfluss des Dopamins auf das Verhalten in *patch-leaving decisions* diskutiert. Insgesamt deuten die präsentierten Ergebnisse auf einen Zusammenhang zwischen der E/I Balance im dorsalen anterioren cingulären Kortex beziehungsweise im ventromedialen präfrontalen Kortex mit interindividuellen Unterschieden im Verhalten bei *patch-leaving*

decisions respektive der *value-guided choice* hin. Zudem legen wir anhand früherer Forschung dar, wie das dopaminerge System potentiell den Einfluss vorhandener Alternativoptionen auf die Entscheidungsfindung in *patch-leaving decisions* verändert. Unsere Ergebnisse deuten auf eine komplexe Modulation des belohnungsbasierten Entscheidungsverhaltens durch unterschiedliche Neurotransmittersysteme hin und zeigen zum ersten Mal einen empirischen Zusammenhang zwischen der E/I Balance und dem *patch-leaving* Verhalten im Menschen.

Abstract

Humans need to choose between options of different reward value on a regular basis to maximize individual gains. These decisions can be encountered in various ways. In the present work, we address two types of reward-based decisions, patch-leaving decisions, and value-guided choice. Patch-leaving decisions require humans to balance the benefits of staying in a current environment and leaving for a potential richer one even though leaving commonly comes with a cost. An example from everyday life would be the decision to change jobs. Here, one needs to assess how satisfied they are with their current position. In addition, the value of potential alternatives needs to be considered as well as potential costs associated with switching such as needing to move to another city. During value-guided choice, on the other hand, the agent chooses between currently available options. As an example, one could consider the choice of a new apartment after moving. Here, based on different factors, such as location and price, it must be determined which is best current option. As can be seen from the example, the overall value often is composed of different dimensions that need to be integrated to compare options during value-guided choice. Previous research suggests that many different neurotransmitters contribute to reward-based decision-making. In this work, we first consider the influence of the excitation-inhibition balance (E/I balance), determined by relative concentrations of the excitatory neurotransmitter glutamate and the inhibitory neurotransmitter GABA, on behaviour in both types of decisions. Second, we discuss a potential modulatory influence of dopamine on patch-leaving decisions. Overall, the results presented suggest dissociable contributions of the E/I balance in the dorsal anterior cingulate cortex and ventromedial prefrontal cortex to interindividual differences in behavior during patch-leaving decisions and value-guided choice, respectively. In addition, we outline how dopamine possibly affects the influence of alternative options on decision-making during patch-leaving based on earlier research. Our results suggest a complex modulation of reward-based decision-making behaviour by different neurotransmitter systems and show an empirical link between E/I balance and patch-leaving behaviour in humans for the first time.

List of abbreviations

vmPFC	-	Ventromedial prefrontal cortex
dACC	-	Dorsal anterior cingulate cortex
E/I balance	-	Excitation-Inhibition balance
fMRI	-	Functional magnetic resonance imaging
MEG	-	Magnetoencephalography
RPE	-	Reward prediction error
BIC	-	Bayesian information criterion
DDM	-	Drift diffusion model
MRS	-	Magnetic resonance spectroscopy
MRI	-	Magnetic resonance imaging
dlPFC	-	Dorsolateral prefrontal cortex

Contents

1. Introduction

Given limited resources, humans and other animals need to decide between options of different value to maximize individual rewards and ensure survival on a daily basis. To understand how humans make these reward-based decisions is a fundamental problem in many areas of research, such as machine-learning, economics and cognitive neuroscience (Sutton and Barto, 1998; Leiser and Azar, 2008; Glimcher and Fehr, 2013).

In this work, we focus on two classes of reward-based decisions, value-guided choice and patch-leaving decisions. In value-guided choice problems, decisions are made between current available options of differing value. The value of an option can be constituted of different dimensions that need to be integrated to a common value to decide advantageously. Moreover, we concentrate on patch-leaving decisions on when to leave a resource, or patch, of depleting value to start exploiting another option.

Given the ubiquitous importance of efficient reward-based decision-making, understanding the neural basis of how values are processed and compared has a long history of investigation. When choosing between different options in the above-described decision problems, representations of their respective value are reflected in neural activity across the brain (Boorman et al., 2009; Kolling et al., 2012; Hunt et al., 2013; Hikosaka et al., 2014). However, to understand whether this activity emerges as an actual reflection of the value comparison itself or as an epiphenomenon, recent research investigated possible mechanistic models underlying decision-making behaviour (Wang, 2002; Hunt et al., 2012; Le Heron et al., 2020). Various neurotransmitter systems have been discussed as possible candidate mechanisms influencing how values comparisons are conducted within the brain as well as affecting eventual choice behaviour (Wang, 2002; Hunt et al., 2012; Jocham et al., 2012; Le Heron et al., 2020). The present work aims at contributing to previous research by investigating the role of neurochemical systems underlying and influencing value-guided choice and patch-leaving decisions. In the following introductory paragraph, I will give an overview of these decisions and highlight key empirical findings with respect to their potential neuronal correlates as well as possible influences of the neurotransmitters GABA, glutamate and dopamine.

1.1 Value-Guided Choice

During value-guided choice, decision-makers need to find the current best available option to decide advantageously. Therefore, they need to identify the dimensions constituting the overall value of an option, integrate them and compare them against potential alternatives. As an example, if I want to make the best possible choice for today's dinner, I have to decide among different places available in my neighbourhood. If I like pizza better than Sushi, but heard of someone having a bad experience at my nearby pizza place, I would need to integrate my love for pizza with the risk of a ruined dinner to compare both options.

Empirical evidence suggests the ventromedial prefrontal cortex (vmPFC) as one key region reflecting value comparisons during value-guided choice (Hunt et al., 2012). A system involved in guiding value comparisons would, at first, be expected to be influenced by the overall value of available options (Hunt et al., 2012; Strait et al., 2014). However, finding value correlates does not necessarily implicate that the respective region is guiding the value comparison process (Hunt et al., 2012). Intuitively, an array of different cognitive functions is assumed to be closely related to option values, such as value-dependent attentional allocation (Maunsell, 2004; Kunar et al., 2017) or the preparation of specific actions necessary to obtain the respective option value (Wunderlich et al., 2009). It is therefore difficult to determine whether the respective signals relate to value itself (O'Doherty, 2014), the choice process (Hunt et al., 2012) or if an explicit representation of value even exists in the brain (Yoo and Hayden, 2018). A seminal paper (Hunt et al., 2012) therefore investigated what activity is fundamental to the value comparison process by using a theoretical model of neural activity based on known neurobiology (Wang, 2002). In brief, their simulations predict an early reflection of overall value proceeding to a value-difference signal, the relative advantage of the chosen over the unchosen option, in regions implementing a value comparison process. A system implementing the value comparison process would therefore not only be expected to track the value of going out for pizza and sushi, but also their respective value difference (Hunt et al., 2012; Papageorgiou et al., 2017). Finally, at the end of the comparison process it is expected to covary with the value of the chosen option (Strait et al., 2014) suggesting that it serves to produce a choice. The authors report that activity in the vmPFC (and posterior superior parietal lobule) matches with model predictions in

early trials, highlighting their importance for value comparisons during value-guided choice (Hunt et al., 2012). Moreover, neural correlates of all of these key components were found in vmPFC activity previously. Option values covary with vmPFC activity when participants are asked to choose between concurrently presented options (Lim et al., 2011; Kolling et al., 2012). Additionally, vmPFC activity reflects the respective value difference (Boorman et al., 2009; Philiastides et al., 2010; Hunt et al., 2012; Jocham et al., 2014) between options and covaries positively with the value of the chosen option (Wunderlich et al., 2009, 2010; Kolling et al., 2012) as well as negatively with the value of the unchosen option (Boorman et al., 2009). Importantly, valuation signals in vmPFC can be found independent of the eventual motor action needed to obtain the respective choice (Wunderlich et al., 2009, 2010) and even when participants were not actively indicating their preferred option (Lebreton et al., 2009; Levy et al., 2011). They have therefore been interpreted as signalling an abstract representation of a decision process between stimuli of different value (Hunt et al., 2013; Jocham et al., 2016).

1.2 Patch-Leaving

While it is important to understand how humans and other animals take comparative, binary value-guided decisions those are not the only types of choice we are faced with in everyday life. Additionally, they are probably not the only type of decision that have shaped our decision-making systems evolutionary (Stephens, 2008; Adams et al., 2012; Pearson et al., 2014; Mattson, 2019). Hunter-gatherer cultures presumably rather rarely faced the luxury to choose *which* animal to hunt but rather *whether* the prospect of better hunting grounds elsewhere is worth the associated risk and resources of leaving their current environment. These decisions, where one needs to decide whether to leave their current resource (or patch), despite potential disadvantages, have been coined as patch-leaving decisions in the past (Charnov, 1976; Hayden et al., 2011; Wolfe, 2013).

In order to decide advantageously in this class of choice-problems, one does not only need to consider currently available evidence, but also integrate evidence from the past. It would not be beneficial for the hunter to stay in their current hunting fields if they knew that they could find a herd elsewhere across the hill, whereas the decision might be different when surrounded by wasteland. Even though this example might sound contrived from a perspective of modern life, patch-leaving

decisions have been of high relevance now and then. The decision to change jobs requires to compare how content you are with your current job (based on previous evidence and your expectations for the future) and compare that to other jobs on the market while considering potential risks (of not getting the job) and costs (for example of moving to another city).

Previous research suggests that activity in the dorsal anterior cingulate cortex (dACC) signals key information guiding patch-leaving decisions (Hayden et al., 2011; Kolling et al., 2016). Neural signalling in the dACC correlates with the average value of potential alternatives as well as the cost of leaving in humans (Kolling et al., 2012). Also in primates, neurons in the dACC signal the relative value of leaving a depleting patch for an alternative one and those signals are modulated by the respective travel time required to move between patches (Hayden et al., 2011). Furthermore, activity in the dACC tracks the reward history (Holroyd and Coles, 2008; Bernacchia et al., 2011; Wittmann et al., 2016), an information necessary to judge the overall quality of the current environment. Activity in ACC does not only track the value of different choice options but is also predictive of whether those signals lead to behavioural change (Fouragnan et al., 2019) such as switching to an alternative option (Wittmann et al., 2016). In addition to that, activity in the dACC reflects discerned trends in the environment and could thereby allow an estimation of future reward availability within an environment (Wittmann et al., 2016).

It is unclear whether the previously described neural correlates of value information reflect computations fundamental to the comparison of option values or emerge as an epiphenomenon. Just because activity in one region correlates with value difference or the relative value of leaving, this does not necessarily mean that this region is implementing the choice process itself. In analogy, if one registers a power failure after a defective fuse they would not assume that it is producing current because one knows about the underlying mechanisms where the fuse functions as an electrical safety device. With respect to decision-making, theoretical ideas on *how* values are compared can be captured in formalized theoretical models which generate artificial neuronal signals and make predictions on expected choice behaviour. Those predictions can then be compared to empirical data to test whether the observed signals match model predictions (Hunt et al., 2012) and confirm the assumed candidate mechanism. One of these mechanistic models assumes that

value comparisons crucially depend on the excitation-inhibition balance (E/I balance) within neural circuits (Wang, 2002; Hunt et al., 2012; Jocham et al., 2012; Kolling et al., 2016) which are consequently expected to affect the speed of value integration and the upcoming choice pattern in a predictable manner. In the current work, we empirically test whether individual relative concentrations of the excitatory neurotransmitter glutamate and the inhibitory neurotransmitter GABA contribute to behaviour during patch-leaving and value-guided choice. The mechanistic key predictions and previous empirical findings regarding the influence of the E/I balance on individual choice behaviour are highlighted in the following.

1.3 Value-Guided Choice, Patch-Leaving and the E/I Balance

One theoretical mechanism assumed to underlie value comparisons is *competition by mutual inhibition* (Wang, 2002, 2008; Hunt et al., 2012; Strait et al., 2014; Kolling et al., 2016). According to the assumptions of this model, neuronal pools representing choice options are excited according to their input value and possess recurrent excitation. Those neuronal pools engage in a mutual inhibition process until activity remains in the eventual “winning” pool and a choice for that option is being made (Wang, 2008; Hunt et al., 2012; Kolling et al., 2016). As an example, if I am to compare a highly valuable option A to a lower value option B, there will be a much greater excitement in neurons representing A, which can easily compete with the low excitement in pool B via inhibitory connections. By suppressing activity in pool B the model will converge to represent option A. It thereby predicts a progression of neural signals from reflecting the overall value of choice options to their respective difference (during the competition) (Hunt et al., 2012), the key information guiding value-guided choice .

The theoretical assumptions of these models have previously been captured in a biophysically plausible neural network model to simulate artificial neural signals that would be expected from a system based on competition by mutual inhibition (Hunt et al., 2012). Functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) signals in the vmPFC match predictions derived from competition by mutual inhibition (Jocham et al., 2012; Hunt et al., 2012). Furthermore, single-unit recordings from macaques vmPFC revealed activity consistent with mutual inhibition, such as anti-correlated tuning curves for different choice options as well as a progression towards signalling the chosen option (Strait

et al., 2014). Additionally, since this class of models relies on inhibitory and excitatory connections, testable hypotheses can be derived regarding the influence of excitatory to inhibitory neurotransmitters, such as glutamate and GABA, on value comparisons during value-guided choice (Jocham et al., 2012). If the overall recurrent excitation in the network is high, it is expected that the model converges faster to a choice at the expense of accuracy in choosing the higher valued option (Hunt et al., 2012). Indeed, humans were found to be more reliable in selecting the higher valued option with a lower E/I balance in vmPFC (Jocham et al., 2012). The previous results support competition by mutual inhibition as a candidate mechanism underlying value comparisons during value-guided choice and thereby highlight the E/I balance as a potential variable underlying interindividual differences in decision-making behaviour.

Even though a similar mechanism is assumed as a candidate mechanism underlying other types of choice, such as patch-leaving decisions (Kolling et al., 2016), supporting empirical evidence regarding the E/I balance in humans is still lacking. During patch-leaving, it is hypothesised that the value of potential alternative options is represented by a neuronal pool interacting with neurons reflecting the costs of leaving. In addition, the current patch value excites a neuronal population possessing self-excitation communicating with the alternative population via inhibitory neurotransmission (Kolling et al., 2016). Again, the account predicts that depending on the levels of excitation to inhibition, the network would eventually move to an attractor state in which neurons representing staying with the current status quo or leaving for an alternative are active and a choice for that option is being made (Kolling et al., 2016). Relationships between inhibitory and excitatory neurotransmission and the decision to switch away from the current course of action have been reported previously. In animals, inhibitory transmission contributes to behavioral adaptation in changing environments (Cho et al., 2020) as well as to the decision to leave or stay a current environment (Kvitsiani et al., 2013). While fMRI signals in humans match predictions from the model (Kolling et al., 2012; Kolling et al., 2016), an examination of the proposed relationships with E/I balance is still missing, and is an integral part of the present work.

Reward-based choices were shown to be influenced by many different neurotransmitter systems previously (Constantino et al., 2017; Burke et al., 2018)

and disentangling specific contributions is an important field of ongoing research. Given that neurotransmitters influence each other on multiple levels, the dopaminergic system has for instance been shown to directly modulate levels of cortical inhibition (Seamans et al., 2001; Gorelova et al., 2002; Winterer and Weinberger, 2004), complex neurophysiological interactions underlying choice behaviour are to be expected. The present work highlights a potential role of the dopaminergic system during patch-leaving decisions in addition to the role of the E/I balance in influencing choice behaviour. Previous research regarding this relationship is outlined in the following paragraph.

1.4 Patch-Leaving and Dopamine

To decide whether or not to leave a current environment, it is essential to know which costs are associated with leaving and whether the expected value benefits in an alternative environment are worth paying that cost. Therefore, the current and alternative patch's reward history as well as switch costs are essential to make an informed decision. The reward history potentially allows to extract trends within the environment to evaluate what to expect after leaving, which is the key information to decide whether to stay or to leave (Wittmann et al., 2016).

The mesolimbic dopamine system plays a crucial role in signaling this kind of information. Dopamine has been found to be involved in overcoming costs in order to obtain rewards (Day et al., 2010). Additionally, dopamine is involved in tracking the average reward rate of the environment (Niv et al., 2007; Hamid et al., 2016), providing estimates of available future rewards (Day et al., 2010; Hamid et al., 2016) as well as in signaling the proximity and value of distant rewards (Howe et al., 2013). It is further involved in signaling the reward history (Bayer and Glimcher, 2005; Santesso et al., 2009; Bromberg-Martin et al., 2010).

The previous evidence suggests a role of dopaminergic signaling in patch-leaving scenarios. Indeed, behaviour in a patch-foraging task is related to the medication status of Parkinson patients, a disease characterized by a depletion of dopamine (Höglinger et al., 2004). More specifically, unmedicated Parkinson patients stayed longer in their respective patch of diminishing returns as compared to control participants (Constantino et al., 2017) and dopamine replacement therapy alleviated this effect. Similarly, the dopamine agonist cabergoline affects behaviour in a patch foraging task. Dopamine agonism influences the weighting of background

reward rates, the reward rate of potential alternatives after leaving the current patch, on behaviour. Participants were shown to leave patches in poor environments earlier under cabergoline, suggesting an increased perceived richness of the environment following dopamine receptor stimulation (Constantino et al., 2017; Le Heron et al., 2020).

1.5 Objective and Hypotheses

In the present work, we sought to investigate how different neurotransmitters relate to reward-based decision-making behaviour. Specifically, we asked how the ratio of excitatory glutamate to inhibitory GABA (E/I balance) relates to choice behaviour during value-guided choice and patch-leaving decisions. Building on previous work, we hypothesised dissociable contributions of the E/I balance in vmPFC and dACC on value-guided choice and patch-leaving, respectively. To the best of our knowledge, it has not been tested before whether the E/I balance relates to decision-making behaviour in a task combining different types of choice in a mechanistically plausible manner. Additionally, in a theoretical review, we highlight a potential role of dopamine in patch-leaving decisions, commenting on a recent publication (Le Heron et al., 2020). The key findings of the study are summarized, and we provide a discussion on the significance of the work based on previous evidence.

2. Methods

The following section provides a short overview of the main methods that were used in the present work. The reader is invited to refer to our original publications (Kaiser et al., 2021; Marzecová et al., 2021) for a more technical and detailed description.

2.1 Patch-Leaving and Value-Guided Choice Paradigm

In the present work, we set out to design a task combining two choice problems. Firstly, participants were making a patch-leaving choice in every trial. Here, they were presented with two patches of different value. They always resided in one of two patches and moving was associated with a travel cost which was subtracted from participants current earnings as soon as they decided to leave their patch. Crucially, the value in the current patch depleted over time and the value available in the alternative patch replenished. Therefore, participants needed to evaluate whether they are content with the available resources in their current patch, or whether they wanted to switch to the other patch despite associated travel costs. In order to decide advantageously, agents needed to integrate evidence over time *when* they were willing to pay to move to another environment based on the reward history and expected reward trends in both patches. In a second stage of every trial, participants were asked to perform a value-guided choice. Here, the patch value available in the chosen patch was randomly divided to two different choice options and associated with random reward probabilities. The problem to solve during this phase is *what* option to choose based on the potential reward magnitude and the probability of winning.

There are important differences in both classes of decisions since the relevant information is presented on different timescales. The patch-leaving choice crucially depends on the reward history. The agent needs to integrate evidence over time in this sequential decision-problem to decide when she would be better off in an alternative environment and is willing to accept associated switch costs. During value-guided choice, one has to decide based on currently available evidence what is the best available option. Additionally, the current patch choice directly affects the upcoming patch choices whereas the chosen option during value-guided choice does not affect upcoming available options.

2.2 Individual Choice Behaviour

To characterise individual decision-making behaviour, we summarized key choice variables during patch-leaving and value guided choice in two variables, the patch-leaving advantage and the percentage of correct responses during value guided-choice.

The patch-leaving advantage is defined as the average value difference between patches (value available in the alternative patch – value of the current patch – costs) across switch trials. The number of correct responses summarized how many times each participant chose the option with the higher expected value (defined by the product of reward probability and magnitude of each option) during value-guided choice. In addition, we estimated the influence of key reward information, such as the value difference between choice options, on trial-by-trial reaction times in multivariate regression analyses.

2.3 Regression-Based Analyses

A multiple regression-based approach was employed to estimate the relationships of the individual E/I balance and choice behaviour. To limit the number of multiple comparisons, we proceeded along the following hierarchy. First, we only tested those decision-variables of interest for which we had a priori hypotheses. Second, we regressed the E/I balance in all voxels of interest to our main measures of interindividual choice behaviour (patch-leaving advantage and % of correct responses, compare 2.2 Individual Choice Behaviour above). If, and only if, one of the regions exhibited a significant contribution, we further detailed the analysis by assessing whether the effect of E/I balance was contributed to by GABA or glutamate concentrations. Third, only after we found significant relationships of the E/I balance with our broad measures of individual behaviour, we further detailed those analyses on a more fine-grained level. Therefore, the relationships of the E/I balance with the effects of key value parameters on reaction times as well as with model parameters were analysed.

2.4 Computational Modelling

We employed computational modelling to describe participants' decision-making behaviour in a more fine-grained manner. The idea of computational modelling is to capture theoretical accounts on how participants potentially make

decisions using mathematical models. These algorithms can then be compared to actual choice data to determine which theoretical model possibly underlies decision-making behaviour (Wilson and Collins, 2019). The motivation for theoretical modelling was twofold. First, the comparison of a set of different models yields one candidate model which describes the data best and potentially explains observed behaviour in more detail. In the present study, participants needed to integrate reward probabilities and reward magnitudes to decide advantageously during value-guided choice. One could assume that participants weigh reward probabilities differently than reward magnitudes. Additionally, they could, for example, combine the relevant reward information additively or multiplicatively to yield comparable values. If, for example, the choice data is best explained by a model combining reward information multiplicatively this suggests that participants may have employed this theoretically assumed algorithm in their choice process. Importantly, inferring a cognitive algorithm with certainty is impossible since there always is an infinite number of potential competing models that would fit the behaviour equally well but are not considered in the set of theoretical motivated models (Eckstein et al., 2021). Models incorporating different algorithmic assumptions can be used to generate choice probabilities dependent on the reward information presented as well as model-specific variables weighting and combining that information. Those model-specific parameters are *fitted* to individual behaviour, meaning that they are selected in a way that the choice probabilities generated by the model maximize the likelihood of the observed choice data (Wilson and Collins, 2019). These fitted parameters can be used to understand cognitive functions in more detail and to discover potential associations with other variables of interest. One famous example is the association between the firing of dopamine signals and the model-derived reward prediction error (RPE) (Schultz, 1998; Bayer and Glimcher, 2005; Steinberg et al., 2013). The RPE can be thought of as a teaching signal that supports learning when reality does not match predictions and is one key concept in associative learning (Rescorla and Wagner, 1972). Crucially, the firing of dopamine neurons is highly correlated with the strength of the RPE (Schultz et al., 1997; Cohen et al., 2012). This important finding would not be evident from raw behavioural data and highlights the importance of computational modelling and latent variable analysis to understand decision-making processes in more detail. Also in the present work, the second motivation for using

computational modelling was to discover potential relationships of model parameters with neural data, such as the E/I balance.

2.4.1 Choice Models

In the present work, three different classes of models were fitted to explain behaviour during value-guided choice. The key value information presented here are reward probabilities and reward magnitudes. Together, these two properties constitute the overall value of an option. In one class of models, the overall value was derived by combining reward information additively. In the second class, reward information was combined multiplicatively. Finally, we used a hybrid version of both models which included additive and multiplicative components. This approach is based on a recently published fitting procedure (Farashahi et al., 2019). Model fits were compared with the Bayesian information criterion (BIC). The model with the lowest average BIC across all participants was considered as the most likely algorithmic explanation of the observed choice behaviour.

Since it is known that participants do not weigh value information in a statistically optimal way but show systematic distortions (Hsu et al., 2009), we additionally fitted utility functions according to prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Those functions can account for the fact that participants potentially distort the influence of reward magnitudes or probabilities in general by applying non-linear transformations to the objective reward information. We fitted four variants of each model class (additive, multiplicative and hybrid), considering distortions in none of the reward parameters, reward probabilities, reward magnitudes or both.

In order to estimate reliability and validity of computational models and the fitting procedure, it is necessary to generate choices for an artificial agent under different constellations of free parameters (Wilson and Collins, 2019). In this procedure, choice probabilities are generated based on the presented reward information and a random set of free parameters. A choice is then generated based on those probabilities. When those artificial data are used for model fitting, the fitted parameters should match the real parameters (“parameter recovery”) (Wilson and Collins, 2019; Blohm et al., 2020). If the parameters are not recoverable, this is an indication that the experiment is not suited to assess the model of interest which

makes the interpretation of the recovered parameters from real data impossible. We therefore included a model validation step in our analyses and report the correlations between the real parameters used for simulation and recovered parameters.

2.4.2 Drift Diffusion Modelling

We additionally fitted drift diffusion models (DDM) to the data to extract further latent variables, such as indices of evidence accumulation and response caution (Wiecki et al., 2013; Cavanagh et al., 2014), and elucidate their relationship with the E/I balance. In DDMs, each decision is defined by a respective decision threshold. Every choice is modelled as an accumulation of evidence until a threshold is reached at which timepoint the respective choice is executed. Predictions from the model crucially depend on the speed of evidence accumulation (drift rate), the respective response caution (height of the decision threshold) and an initial bias towards either boundary. The individual free parameters, such as drift rate, are fitted to choice and reaction times in a way that those predicted patterns match observed behaviour best.

More specifically, we employed a hierarchical Bayesian DDM (Wiecki et al., 2013). The fitting procedure is named “hierarchical” since individual parameters are constrained by group-level distributions (Wiecki et al., 2013). In addition to that, the modelled hierarchical Bayesian DDMs allows for an estimation of trial-by-trial regression models on choice parameters. As an example, one could assume that drift rate crucially depends on the value difference between options such that the drift towards the decision boundary is higher with greater value differences (easier decisions). We have incorporated the effects of value parameters on free choice variables, such as drift rate, in the decision models. The reader is referred to Publication I (Supplementary material) for an overview of all DDM models fitted.

2.5 Magnetic Resonance Spectroscopy

Magnetic Resonance Spectroscopy (MRS) is a non-invasive technique to measure the concentrations of metabolites (Ding and Lanfermann, 2015). MRS makes use of the fact that certain atomic nuclei within molecules possess their own magnetic fields (Alger, 2009). The respective nuclei are arranged randomly under normal conditions (Haley and Knight-Scott, 2011). However, when nuclei are placed in an external magnetic field, they align along that field (Haley and Knight-Scott, 2011) and their rotation frequency depends upon the strength of the external field

and the nature of the nucleus (Fayed et al., 2006). The external field in the current work was applied with a Magnetic Resonance Imaging (MRI) scanner with a field strength of 7 Tesla (Siemens Healthineers). When a time-varying radio-frequency field is additionally turned on at this particular frequency (Fayed et al., 2006), nuclei are excited (Haley and Knight-Scott, 2011; Gruber et al., 2018). As soon as this second externally applied field is turned off, MRS measures how nuclei return back to their original position in alignment with the external field (Chatham and Blackband, 2001). This process generates distinct signals which can be detected with a physical coil (Alger, 2009). The detected signal is dependent upon the unique chemical environment around the nucleus which is specific for different metabolites allowing their concentrations to be quantified (Chatham and Blackband, 2001; Hwang and Choi, 2015; Tognarelli et al., 2015; Ford and Crewther, 2016).

MRS measures were obtained in predefined localized regions of interest (voxels). In the present work, signals were measured from five voxels of interest: the dorsolateral prefrontal cortex (dlPFC), left M1, right M1, vmPFC and dACC. We used the dACC and vmPFC as regions of interest due to their above-described role in signaling decision variables during patch-leaving and value-guided choice. Since patch-leaving requires maintaining a representation of current patch values across trials, we selected the dlPFC because of its importance for working memory-related processes (Curtis and D'Esposito, 2003). The motor cortex was selected as a control region, where we did not expect a relationship with any decision variable.

MRS signals are usually quite weak and exhibit a low signal to noise-ratio (Alger, 2009; Ligneul et al., 2021; Ip and Bridge, 2022). In the present work, MRS signals are therefore not time resolved but we worked with a single estimate of molecular concentrations within each voxel of interest.

2.6 Theoretical Review

In our second publication we provide a review of a recently published article (Le Heron et al., 2020). A short overview of the article is provided as well as a description of the key results. We discuss the findings by the authors in the light of recent publications and provide a brief discussion on the significance of the paper.

3. Results

3.1 Publication I: Dissociable Roles of Cortical Excitation-Inhibition Balance during Patch-Leaving versus Value-Guided Decisions

The following chapter is based on our manuscript published in Nature Communications (see attachments):

Kaiser, L. F., Gruendler, T. O., Speck, O., Luettgau, L., & Jocham, G. (2021). Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions. *Nature communications*, *12*(1), 1-11.

As described above, previous evidence suggests an influence of the E/I balance in vmPFC on value comparisons during comparative value-guided choice (Hunt et al., 2012; Jocham et al., 2012; Strait et al., 2014). However, even though a similar mechanism was hypothesised to influence patch-leaving behaviour in dACC (Kolling et al., 2016), to the best of our knowledge, this has not been empirically tested before in humans. We therefore sought to investigate the role of the cortical E/I balance in a task combining patch-leaving and value-guided choice. More specifically, interindividual differences in value-guided choice and patch-leaving behaviour were hypothesised to be influenced by E/I balance in vmPFC and dACC, respectively.

In a novel decision-making task, participants (N = 29) first conducted a patch-leaving choice followed by a value-guided choice. Participants were scanned with MEG during the behavioural paradigm (data not shown here). In a separate session, they were measured with MRS at 7 Tesla MRI. During the patch-leaving stage, participants decided whether they want to stay with their current patch or switch to an alternative. Switching to the alternative patch was associated with costs which were subtracted from their current earnings as soon as they decided to leave their current patch. The costs were displayed on screen and remained stable until participants decided to leave. Crucially, the value available in the current patch depleted over time whereas the value in the alternative patch replenished. In every trial, participants therefore needed to decide whether the relative patch-value available in the alternative patch (as well as its expected future dynamics) was worth

paying the costs associated with leaving. To capture individual differences in patch-leaving behaviour, we assessed the average relative value benefits that participants needed to leave their current patch despite current costs (patch-leaving advantage). In the second stage of every trial the value available in the current patch was randomly allocated to two different choice options and associated with random reward probabilities which were displayed on screen. In order to decide advantageously, participants would therefore need to consider the reward probability and reward magnitude in their value-guided choice. Our main measure of interest during value-guided choice was how often participants chose the option with the higher expected value.

We observed dissociable contributions of the E/I balance in different brain regions as a function of decision type. Participants considered the individual cost levels in their patch-leaving choices and left patches at higher relative value in the alternative patch (compared to their current patch value) when leaving was associated with greater costs. As expected, this indicates that participants needed a higher incentive, in the form of alternative patch value, to leave their current patch when leaving was associated with greater costs. With respect to the E/I balance, the patch-leaving advantage was significantly related to the E/I balance in dACC but not any other region investigated. This indicates that participants with a greater E/I balance in dACC left their current patch when the average evidence to leave was higher. Drift Diffusion modelling confirmed these findings by revealing a significant relationship of dACC E/I balance and drift rate indicating that participants with a greater dACC E/I balance show a higher drift towards stay decisions. Additionally, participants showed a trend to slow down their patch-leaving choices when leaving was associated with higher costs. The magnitude of this effect was again related to the E/I balance in dACC.

During value-guided choice, we found that the E/I balance in vmPFC was significantly related to the individual percentage of correct choices, such that participants with a greater vmPFC E/I balance chose the option with the higher expected value less often. There were again no significant contributions of any other regions. We further detailed this analysis by assessing relationships of the E/I balance with effects of key value parameters on reaction times and computational modelling parameters. We found that participants responded slower in trials with low

value difference (harder trials). With respect to E/I balance, the magnitude of this effect was again lower with an increased E/I balance in vmPFC. This points towards an influence of vmPFC E/I balance on the speed of value integration. Computational modelling indicated that a model assuming multiplicative integration of reward magnitudes and probabilities with distortion in both value parameters explained participant's choice data best. The degree of individual distortions was again significantly related to E/I balance in vmPFC, but not any of the other regions investigated.

In sum our findings support our hypotheses. We report significant and dissociable contributions of the E/I balance in dACC and vmPFC to choice behaviour during patch-leaving and value-guided choice, respectively. Our findings further support models that implement competition by mutual inhibition as one candidate mechanism underlying behaviour during value-guided choice and provide evidence for an influence of the E/I balance in dACC on patch-leaving behaviour in humans for the first time.

3.2 Publication II: Neuromodulation of Foraging Decisions: The Role of Dopamine

The following section is based on our work published in Frontiers in Behavioral Neuroscience

Marzecová, A**.**, Kaiser, L.F., & Maddah A. (2021) Neuromodulation of foraging decisions: The role of dopamine. *Frontiers in Behavioral Neuroscience*. 15, 1-4.

In this opinion piece, we discussed recent results on the role of dopamine in patch-leaving decisions (Le Heron et al., 2020). In the task created by the authors, participants were presented with patches that differed in their initial reward rates (low, medium and high yield) and were presented in rich or poor environments. Rich and poor environments differed by the proportions of high, medium and low yield patches such that the availability of high yield patches was greater in the rich background environment. The authors thereby separately manipulated the effects of foreground (patch reward rate) and the background reward rate (average reward in the environment). The reward rate available in the current patch decreased over time and participants could choose to leave their current patch whereas they were more

likely to transition to a high yield patch in the rich environment compared to the poor environment. Leaving one's current patch incurred a cost in the form of a fixed travel time during which no rewards could be obtained.

Evidence on neuromodulatory mechanisms underlying patch-leaving decisions remains scarce. The authors have filled this gap by assessing the role of the D2 agonist cabergoline on behaviour in the above-described decision-making task and reported a modulatory influence of cabergoline on the influence of background reward rates on behaviour. In Publication II we discussed the following background information on this important publication by Le Heron et al. (2020). Participants were administered with a 1 mg dose of cabergoline. The authors hypothesised that this would specifically modulate tonic DA levels, which have previously been suggested to signal the average background reward rate (Niv et al., 2007; Beierholm et al., 2013) and should thereby influence the effects of background reward rate on choice behaviour in the current framework. However, it has been discussed before whether similar doses would more strongly affect phasic DA signalling (Frank and O'Reilly, 2006; Norbury et al., 2013) by influencing higher-affinity presynaptic autoreceptors to a greater degree than postsynaptic receptors (Norbury et al., 2013). In the current design, there was no possibility to assess whether the neurochemical challenge affected pre- or postsynaptic receptors. The discussed study indeed only finds a significant interaction of background, and not foreground reward rate, with drug on behaviour, indicating a specific influence on the perception of environmental richness. However, this effect is due to drug effects in the poor environment only, such that participants left patches there earlier under cabergoline. This speaks against a general influence of the dopaminergic challenge on background reward rate. Disentangling potential contributions of tonic versus phasic dopaminergic signalling on how the value of a current state versus potential other states is estimated could be an interesting question for upcoming research.

Additionally, the perceived interaction effects of drug and background reward rate could be due to two factors. One would be a drug-dependent increased perceived richness of the environment as already discussed (Le Heron et al., 2020). Another one would be a reduced threshold to overcome the inhibition of travel costs. There is extensive literature on the role of dopamine in cost-benefit decisions (Salamone et al., 1994; Beeler and Mourra, 2018). Here, the question is conceptually

similar as in the foraging paradigm: „Is my expected benefit worth the expected costs?" In the discussed study, travelling between patches incurred a fix time cost. During this time, no rewards could be collected. Missing out on rewards is worse in rich environments than in environments with a low reward density, or differently stated, the opportunity costs of time differ (Constantino and Daw, 2015; Constantino et al., 2017) and could thereby affect the weighting of travel times in rich versus poor environments differently. This could be reflected in an effect of background reward rate since costs are only incurred when participants decide to leave their current patch. Parkinson patients, diagnosed with a disease characterized by a depletion of dopamine (Höglinger et al., 2004), have indeed been shown to weigh the opportunity cost of time differently dependent on their status of dopamine replacement therapy (Constantino et al., 2017). Parkinson patients off medication were shown to stay longer with a depleting resource, again consistent with a decreased perceived richness of the environment in patients with dopamine deficiency. However, in this task the environmental richness was varied by changing travel times and not by different distributions of patch types in the environment (Le Heron et al., 2020). It would be of interest for future studies to explicitly vary travel costs (Constantino et al., 2017) and the average value available in the environment (Le Heron et al., 2020) to further the understanding of dopamine`s role in patch-leaving decisions.

4. General Discussion

The present work aimed to investigate neurochemical underpinnings of reward-based decision-making in different frames of reference. Many earlier studies have examined potential anatomical structures involved in value comparisons (FitzGerald et al., 2009; Strait et al., 2015; Shapiro and Grafton, 2020), but research on potential underlying neurochemical mechanisms remains scarce. Here, we provide evidence for dissociable contributions of the E/I balance to patch-leaving and value-guided choice in dACC and vmPFC, respectively.

Specifically, we found evidence for contributions of the E/I balance in dACC to behavioural parameters of patch-leaving behaviour. Participants with a greater E/I balance in dACC left their current patch when the relative value benefit (subtracted by current costs) in the alternative patch was higher. Additionally, the effect of patch-leaving costs on reaction times was significantly related to E/I balance in dACC. These findings extend previous animal literature reporting an influence of inhibitory neurons in the decision on whether to stay or to leave (Kvitsiani et al., 2013). In addition, it has previously been suggested that a mechanism based on the E/I balance in dACC influences patch-leaving behaviour (Kolling et al., 2016). To the best of our knowledge, this is the first study testing that hypothesis empirically.

There are many competing theories about the role of dACC in patch foraging decisions. On the one hand, many studies suggested a direct encoding of choice variables in dACC such as signalling the relative value of leaving (Hayden et al., 2011) or the value of potential alternative options (Kolling et al., 2012). Additionally, recent work provides causal evidence for an active role of the ACC in strategy switching (Tervo et al., 2021) during foraging. One the other hand, the dACC has been suggested to signal more general factors, such as monitoring choice conflict and regulating cognitive control (Shenhav et al., 2013; Shenhav et al., 2014; Shenhav et al., 2016; Kane et al., 2021) or linking different contexts with task-related strategies (Heilbronner and Hayden, 2016), which would mediate the reported role in guiding patch-leaving behaviour. Rats with ACC inactivation, for example, still follow optimal foraging theory (Kane et al., 2021), questioning its direct role in guiding patch-leaving decisions. Both accounts could –in theory– contribute to the observed influence of dACC E/I balance on patch-leaving behaviour. With the current task-design we can only report an influence of E/I balance on overall patch-leaving

behaviour, but future research would be needed to further disentangle the more specific computations underlying this relationship.

With respect to value-guided choice, we found a decreased choice accuracy with a higher E/I balance in vmPFC. Additionally, participants slowed down their reaction times in difficult trials with low value difference. This effect was less pronounced with a greater E/I balance in vmPFC. The presented findings match with model-predictions based on competition by mutual inhibition. Such a model would predict that decision-makers with stronger relative levels of excitation would make more mistakes (Wong and Wang, 2006; Hunt et al., 2012). Our findings are also consistent with earlier work reporting an increase in choice consistency with high levels of GABA and low levels of glutamate in vmPFC (Jocham et al., 2012). Taken together, our findings support dissociable contributions of the E/I balance in the dACC and vmPFC to individual behavioural differences during patch-leaving versus value-guided choice.

Even though we find dissociable contributions of the E/I balance as a function of decision type, our findings do not argue for a modular perspective on choice which prescribes certain functions to specific, well-defined regions or neurotransmitter systems. In other words, the present work does not imply a clear one to one mapping from value-guided choice and patch-leaving to vmPFC and dACC activity, respectively. Earlier research implies no categorically distinct role of cortical brain regions during economic choice (Yoo and Hayden, 2018) and that the involvement of different cortical regions critically depends on task requirements. Signals of value comparisons can, for instance, be found in motor cortex (and not in vmPFC) when options are presented sequentially during value-guided choice and a response-delay is introduced (Hunt et al., 2013). With respect to patch-leaving several regions beyond dACC have been reported to carry important decision variables and contribute to computations during patch-leaving decisions (Barack et al., 2017; Kane et al., 2017; Hall-McMaster and Luyckx, 2019). This evidence suggests that reward-based choice is distributed across many brain regions (Cisek, 2012; Hunt and Hayden, 2017; Yoo and Hayden, 2018; Maisson et al., 2021). Additionally, reward-based decisions are influenced by a complex interplay of many different neurotransmitter systems (Takahashi et al., 2010; Scholl et al., 2014; Burke et al., 2018). For example, in the current work we discuss the results of an important recent

article (Le Heron et al., 2020) describing a modulatory influence of dopamine on patch-leaving behaviour emphasizing the importance of other transmitter systems in understanding the neurochemical underpinnings of reward-based decision-making.

Taken together, making decisions based on their value is a ubiquitous problem in everyday life. Providing mechanistic and biologically sound theories on how options of different value are potentially compared could be an important first step to elucidate when and how decision-makers do not choose the optimal course of action (Leiser and Azar, 2008; Sharp et al., 2012; Constantino and Daw, 2015; Kane et al., 2019) that maximizes their individual gains. Understanding how reward-based decisions are modulated by different neurotransmitters therefore is an important endeavor in ongoing research and this work contributes to this question by highlighting certain neurochemical influences on reward-based decision-making.

5. Future Research

The present work provides evidence for dissociable contributions of the E/I balance in dACC and vmPFC to patch-leaving and value-guided choice respectively. In future research, we will examine how the individual E/I balance affects time-varying value-representations in MEG data, which has been measured during the above-described behavioural choice paradigm. Based on earlier research we expect a faster ramping of signals reflecting value comparisons with greater E/I balance (Jocham et al., 2012). Using MEG, it is possible to capture the temporal evolution of choice signals with greater temporal accuracy, an advantage over fMRI (Hall et al., 2014) that has been used in earlier work (Jocham et al., 2012). We expect the ramping of choice signals to directly depend upon option value and therefore to exhibit temporal trial by trial variations. To capture this dynamic temporal evolution of the value comparison process, we use Hidden Markov Models allowing the detection of hidden states with different dynamics across trials (Vidaurre et al., 2018; Higgins, 2019). We expect a relationship between the distribution of hidden states and the E/I balance.

Our findings in Publication I are in line with theoretical predictions based on mutual inhibition. As in previous research (Hunt et al., 2012), biophysically plausible network models can be used to simulate which activity would be expected from a mechanism based on competition by mutual inhibition. These artificial predictions can then be compared with actual data to estimate whether the observed pattern of results is fundamental to the computational process of value comparisons (Hunt et al., 2012). It would be a promising avenue for future research, to directly take the E/I balance into account during model simulations and estimate whether neural signals match with model predictions.

Additionally, neuronal gamma oscillations are assumed to be influenced by the E/I balance (Muthukumaraswamy et al., 2009; Buzsáki and Wang, 2012; Bitzenhofer et al., 2021). Gamma-frequency oscillations have been shown previously to support evidence accumulation during value-based choice, whereas fronto-parietal synchronization was discovered as predictive for choice accuracy (Polanía et al., 2014). In ongoing research, we therefore investigate how the measured E/I balance affects frequency-specific neuronal value correlates as well as their temporal dynamics. Furthermore, earlier research suggests frequency-specific interactions

between activity in the vmPFC and dorsomedial prefrontal cortex during the decision to switch from an ongoing strategy (Domenech et al., 2020). By analyzing the MEG data, we hope to uncover how the E/I balance influences evidence accumulation and interregional functional connectivity guiding the decision to switch away from a current patch.

Recent work (Kaanders et al., 2021b; Kaanders et al., 2021a) provides a potential alternative explanation for the observed results. It has been suggested that pre-supplementary motor/dACC activity is reflective of how much information is sampled before committing to a choice. It could be of interest for future studies to disentangle whether the observed differences in switching behaviour are due to actual different leaving thresholds or to differential information sampling, since those factors are intertwined in the current task design.

In the present work, we have discussed the findings of a recent paper (Le Heron et al., 2020) presenting a modulatory influence of dopamine on patch-leaving decisions. The firing of dopamine neurons has been extensively researched in decisions requiring to learn a link between stimulus and reward (Frank et al., 2004; Ersche et al., 2008; Cools et al., 2009; Jocham et al., 2011; Maes et al., 2020). However, the dopaminergic system emerges as a neural substrate influencing reward-based choice in many different ways. Dopamine has not only been shown to influence patch-leaving decisions (Le Heron et al., 2020) but also other factors affecting value-guided choice. Dopamine neurons encode reward probabilities and magnitudes (Tobler et al., 2005) and play a central role in the computation of subjective value computations during value-based choice (Burke et al., 2018). Dopamine is not only implicated in value coding but also a vast array of other functions such as, for example, movement control, action selection (Barter et al., 2015; Howard et al., 2017; Da Silva et al., 2018) and in signaling the proximity to distant rewards (Howe et al., 2013). Current animal research deals with the question how these diverse functions are organized and encoded by dopaminergic neurons (Engelhard et al., 2019; Schwerdt et al., 2020). In human research, it has been shown that subtle manipulations in the timing of option presentation and the response scheme during value-guided choice affects whether option values are mainly encoded in prefrontal regions or motoric structures (Hunt et al., 2013). Building on this previous evidence, we are currently investigating whether a

dopaminergic challenge affects the encoding of value or motor related task features (or both) during value-guided choice. More specifically, we used a sequential value-guided choice task to track the encoding of value and movement-preparation during the decision process. Participants performed this task under the dopamine precursor L-Dopa, the D2/D3 receptor dopamine antagonist amisulprid and a placebo. We measured the encoding of task related variables, such as option value and movement preparation, with MEG in high temporal resolution and are investigating whether they differ as a function of the dopaminergic challenge (in preparation).

6. Conclusion

The present work focused on the neurochemical underpinnings of reward-based decision-making providing evidence for dissociable contributions of the E/I balance in dACC and vmPFC to patch-leaving versus value-guided choice. To the best of our knowledge, this is the first study assessing potential influences of GABA and glutamate concentrations on behaviour in a task combining both decision-problems. We additionally highlighted a potential role of dopamine in governing patch-leaving decisions, suggesting that the dopaminergic state affects the influence of potential alternative options on the decision to leave a current patch.

Understanding the role of different neurotransmitter systems in guiding reward-based choice is an important endeavor. Only if we learn how certain transmitter systems affect behaviour, we might help treating and understanding diseases which are associated with altered decision-making (Blanco et al., 2013; Gillan et al., 2014; Pushkarskaya et al., 2015; Albrecht et al., 2016; Constantino et al., 2017; Verharen et al., 2018; Cavanagh et al., 2020) and signaling in discussed neurotransmitter systems (Scatton et al., 1982; Pittenger et al., 2011; Belujon and Grace, 2017; Selten et al., 2018; McCutcheon et al., 2020). Apathy, a disorder of diminished motivation is, for example, associated with a disruption in dopaminergic signaling (Chong and Husain, 2016). The above discussed relationship between dopamine and background reward rate potentially leads to a state where the increased perceived richness of the environment is never worth the effort of switching (Le Heron et al., 2018; Le Heron et al., 2020). Therefore, research on basal decision-making mechanisms, such as patch-leaving, can provide potential mechanistic explanations underlying the link how pathophysiology relates to symptomatic cognitive deficits in psychiatric conditions. Researching potential consequences of neuropharmacological manipulations on different types of everyday choices may additionally help to estimate potential treatment risks and could facilitate clinical monitoring. Dopaminergic medication, used to treat Parkinson's disease, has, for instance, been reported to influence learning about rewards (Voon et al., 2010) and the results discussed in this work point towards additional expected modulatory influences on patch-leaving.

Overall, our work contributes to understand the various cognitive computations carried by certain neurotransmitter systems to elucidate how they potentially affect reward-based decision-making. Examining potential mechanisms underlying value comparisons is of high practical relevance given the pervasive importance to identify high valuable options and adapt behaviour accordingly in everyday life.

References

Adams, Watson, Pearson, Platt (2012) Neuroethology of decision-making. Curr Opin Neurobiol 22:982–989.

Albrecht MA, Waltz JA, Frank MJ, Gold JM (2016) Probability and magnitude evaluation in schizophrenia. Schizophrenia Research: Cognition 5:41–46.

Alger JR (2009) Magnetic Resonance Spectroscopy. In: Encyclopedia of Neuroscience (Squire LR, ed), pp 601–607. Oxford: Academic Press.

Barack DL, Chang SWC, Platt ML (2017) Posterior Cingulate Neurons Dynamically Signal Decisions to Disengage during Foraging. Neuron 96:339-347.

Barter JW, Li S, Lu D, Bartholomew RA, Rossi MA, Shoemaker CT, Salas-Meza D, Gaidis E, Yin HH (2015) Beyond reward prediction errors: the role of dopamine in movement kinematics. Front Integr Neurosci 9:39.

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129–141.

Beeler JA, Mourra D (2018) To Do or Not to Do: Dopamine, Affordability and the Economics of Opportunity. Front Integr Neurosci 12:6.

Beierholm U, Guitart-Masip M, Economides M, Chowdhury R, Düzel E, Dolan R, Dayan P (2013) Dopamine Modulates Reward-Related Vigor. Neuropsychopharmacology 38:1495–1503.

Belujon P, Grace AA (2017) Dopamine System Dysregulation in Major Depressive Disorders. Int J Neuropsychopharmacol 20:1036–1046.

Bernacchia A, Seo H, Lee D, Wang X-J (2011) A reservoir of time constants for memory traces in cortical neurons. Nat Neurosci 14:366–372.

Bitzenhofer SH, Pöpplau JA, Chini M, Marquardt A, Hanganu-Opatz IL (2021) A transient developmental increase in prefrontal activity alters network maturation and causes cognitive dysfunction in adult mice. Neuron 109:1350-1364.

Blanco NJ, Otto AR, Maddox WT, Beevers CG, Love BC (2013) The influence of depression symptoms on exploratory decision-making. Cognition 129:563–568.

Blohm G, Kording KP, Schrater PR (2020) A How-to-Model Guide for Neuroscience. eNeuro 7:Eneuro.0352–19.2019.

Boorman ED, Behrens TEJ, Woolrich MW, Rushworth MFS (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron 62:733–743.

Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O (2010) Multiple timescales of memory in lateral habenula and dopamine neurons. Neuron 67:499–510.

Burke CJ, Soutschek A, Weber S, Raja Beharelle A, Fehr E, Haker H, Tobler PN (2018) Dopamine Receptor-Specific Contributions to the Computation of Value. Neuropsychopharmacol 43:1415–1424.

Buzsáki, Wang (2012) Mechanisms of Gamma Oscillations. Annu Rev Neurosci 35:203–225.

Cavanagh JF, Wiecki TV, Kochar A, Frank MJ (2014) Eye tracking and pupillometry are indicators of dissociable latent decision processes. J Exp Psychol Gen 143:1476–1488.

Cavanagh SE, Lam NH, Murray JD, Hunt LT, Kennerley SW (2020) A circuit mechanism for decision-making biases and NMDA receptor hypofunction. eLife 9:e53664.

Charnov EL (1976) Optimal foraging, the marginal value theorem. Theoretical population biology 9:129–136.

Chatham JC, Blackband SJ (2001) Nuclear magnetic resonance spectroscopy and imaging in animal research. IlAR Journal 42:189–208.

Cho KKA, Davidson TJ, Bouvier G, Marshall JD, Schnitzer MJ, Sohal VS (2020) Cross-hemispheric gamma synchrony between prefrontal parvalbumin interneurons supports behavioral adaptation during rule shift learning. Nat Neurosci 23:892–902.

Chong T-J, Husain M (2016) Chapter 17 - The role of dopamine in the pathophysiology and treatment of apathy. In: Progress in Brain Research : Motivation (Studer B, Knecht S, eds), pp 389–426. Elsevier.

Cisek P (2012) Making decisions through a distributed consensus. Curr Opin Neurobiol 22:927–936.

Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482:85–88.

Constantino SM, Dalrymple J, Gilbert RW, Varanese S, Di Rocco A, Daw ND (2017) A Neural Mechanism for the Opportunity Cost of Time. bioRxiv. doi:10.1101/173443.

Constantino SM, Daw ND (2015) Learning the opportunity cost of time in a patch-foraging task. Cogn Affect Behav Neurosci 15:837–853.

Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M (2009) Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J Neurosci 29:1538–1543.

Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7:415–423.

Da Silva JA, Tecuapetla F, Paixão V, Costa RM (2018) Dopamine neuron activity before action initiation gates and invigorates future movements. Nature 554:244–248.

Day JJ, Jones JL, Wightman RM, Carelli RM (2010) Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. Biological Psychiatry 68:306–309.

Ding X-Q, Lanfermann H (2015) Whole Brain 1H-Spectroscopy: A Developing Technique for Advanced Analysis of Cerebral Metabolism. Clinical Neuroradiology 25:245–250.

Domenech P, Rheims S, Koechlin E (2020) Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. Science 369:eabb0184.

Eckstein MK, Wilbrecht L, Collins AGE (2021) What do reinforcement learning models measure? Interpreting model parameters in cognition and neuroscience. Current Opinion in Behavioral Sciences 41:128–137.

Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, Witten IB (2019) Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature 570:509–513.

Ersche KD, Roiser JP, Robbins TW, Sahakian BJ (2008) Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology 197:421–431.

Farashahi S, Donahue CH, Hayden BY, Lee D, Soltani A (2019) Flexible combination of reward information across primates. Nat Hum Behav 3:1215–1224.

Fayed N, Olmos S, Morales H, J Modrego P (2006) Physical Basis of Magnetic Resonance Spectroscopy and its Application to Central Nervous System Diseases. American J. of Applied Sciences 3:1836–1845.

FitzGerald THB, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. Journal of Neuroscience 29:8388–8395.

Ford TC, Crewther DP (2016) A comprehensive review of the 1H-MRS metabolite spectrum in autism spectrum disorder. Frontiers in Molecular Neuroscience 9:14.

Fouragnan EF, Chau BKH, Folloni D, Kolling N, Verhagen L, Klein-Flügge M, Tankelevitch L, Papageorgiou GK, Aubry J-F, Sallet J (2019) The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. Nat Neurosci 22:797–808.

Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. Behavioral Neuroscience 120:497–517.

Frank MJ, Seeberger LC, Reilly RC (2004) By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. Science 306:1940–1943.

Gillan CM, Morein-Zamir S, Kaser M, Fineberg NA, Sule A, Sahakian BJ, Cardinal RN, Robbins TW (2014) Counterfactual processing of economic action-outcome alternatives in obsessive-compulsive disorder: further evidence of impaired goal-directed behavior. Biological Psychiatry 75:639–646.

Glimcher PW, Fehr E (2013) Neuroeconomics: Decision making and the brain. Academic Press.

Gorelova N, Seamans JK, Yang CR (2002) Mechanisms of dopamine activation of fast-spiking interneurons that exert inhibition in rat prefrontal cortex. J Neurophysiol 88:3150–3166.

Gruber B, Froeling M, Leiner T, Klomp DWJ (2018) RF coils: A practical guide for nonphysicists. J Magn Reson Imaging 48:590.

Haley AP, Knight-Scott J (2011) Proton Magnetic Resonance Spectroscopy (H MRS): A Practical Guide for the Clinical Neuroscientist. In: Brain Imaging in Behavioral Medicine and Clinical Neuroscience, pp 83–91. Springer, New York, NY.

Hall EL, Robson SE, Morris PG, Brookes MJ (2014) The relationship between MEG and fMRI. Neuroimage 102:80–91.

Hall-McMaster S, Luyckx F (2019) Revisiting foraging approaches in neuroscience. Cogn Affect Behav Neurosci 19:225–230.

Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD (2016) Mesolimbic dopamine signals the value of work. Nat Neurosci 19:117–126.

Hayden BY, Pearson JM, Platt ML (2011) Neuronal basis of sequential foraging decisions in a patchy environment. Nat Neurosci 14:933–939.

Heilbronner SR, Hayden BY (2016) Dorsal Anterior Cingulate Cortex: A Bottom-Up View. Annu Rev Neurosci 39:149–170.

Higgins C (2019) Uncovering temporal structure in neural data with statistical machine learning models (Doctoral dissertation, University of Oxford).

Hikosaka O, Kim HF, Yasuda M, Yamamoto S (2014) Basal ganglia circuits for reward value-guided behavior. Annu Rev Neurosci 37:289–306.

Höglinger GU, Rizk P, Muriel MP, Duyckaerts C, Oertel WH, Caille I, Hirsch EC (2004) Dopamine depletion impairs precursor cell proliferation in Parkinson disease. Nat Neurosci 7:726–735.

Holroyd CB, Coles MGH (2008) Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. Cortex 44:548–559.

Howard CD, Li H, Geddes CE, Jin X (2017) Dynamic nigrostriatal dopamine biases action selection. Neuron 93:1436-1450.

Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature 500:575–579.

Hsu M, Krajbich I, Zhao C, Camerer CF (2009) Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities. Journal of Neuroscience 29:2231–2237.

Hunt LT, Hayden BY (2017) A distributed, hierarchical and recurrent framework for reward-based choice. Nat Rev Neurosci 18:172–182.

Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MFS, Behrens TEJ (2012) Mechanisms underlying cortical activity during value-guided choice. Nat Neurosci 15:470-S3.

Hunt LT, Woolrich MW, Rushworth MFS, Behrens TEJ (2013) Trial-type dependent frames of reference for value comparison. PLoS Comput Biol 9:e1003225.

Hwang J-H, Choi CS (2015) Use of in vivo magnetic resonance spectroscopy for studying metabolic diseases. Experimental & molecular medicine 47:e139-e139.

Ip IB, Bridge H (2022) Investigating the neurochemistry of the human visual system using magnetic resonance spectroscopy. Brain Struct Funct 227:1491–1505.

Jocham G, Boorman E, Behrens T (2016) Neuroscience of Value-Guided Choice. In: The Wiley Handbook on the Cognitive Neuroscience of Learning (Murphy RA, Honey RC, eds), pp 554–591. Chichester, UK: John Wiley & Sons, Ltd.

Jocham G, Furlong PM, Kröger IL, Kahn MC, Hunt LT, Behrens TE (2014) Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. Neuroimage 100:498–506.

Jocham G, Hunt LT, Near J, Behrens TEJ (2012) A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. Nat Neurosci 15:960–961.

Jocham G, Klein TA, Ullsperger M (2011) Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. Journal of Neuroscience 31:1606–1613.

Kaanders P, Juechems K, O'Reilly J, Hunt L (2021a) Dissociable mechanisms of information sampling in prefrontal cortex and the dopaminergic system. Current Opinion in Behavioral Sciences 41:63–70.

Kaanders P, Nili H, O'Reilly JX, Hunt L (2021b) Medial Frontal Cortex Activity Predicts Information Sampling in Economic Choice. J Neurosci 41:8403–8413.

Kahneman D, Tversky A (1979) Prospect Theory: An Analysis of Decision under Risk. Econometrica 47:263–291.

Kaiser LF, Gruendler TOJ, Speck O, Luettgau L, Jocham G (2021) Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions. Nat Commun 12:1–11.

Kane GA, Bornstein AM, Shenhav A, Wilson RC, Daw ND, Cohen JD (2019) Rats exhibit similar biases in foraging and intertemporal choice tasks. eLife 8:e48429.

Kane GA, James MH, Shenhav A, Daw ND, Cohen JD, Aston-Jones G (2021) Rat anterior cingulate cortex continuously signals decision variables in a patch foraging task. bioRxiv. doi:10.1101/2021.06.07.447464.

Kane GA, Vazey EM, Wilson RC, Shenhav A, Daw ND, Aston-Jones G, Cohen JD (2017) Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task. Cogn Affect Behav Neurosci 17:1073–1083.

Kolling N, Behrens TEJ, Mars RB, Rushworth MFS (2012) Neural mechanisms of foraging. Science 336:95–98.

Kolling N, Wittmann MK, Behrens TEJ, Boorman ED, Mars RB, Rushworth MFS (2016) Value, search, persistence and model updating in anterior cingulate cortex. Nat Neurosci 19:1280–1285.

Kunar MA, Watson DG, Tsetsos K, Chater N (2017) The influence of attention on value integration. Attention, Perception, & Psychophysics 79:1615–1627.

Kvitsiani D, Ranade S, Hangya B, Taniguchi H, Huang JZ, Kepecs A (2013) Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature 498:363–366.

Le Heron C, Apps MA, Husain M (2018) The anatomy of apathy: a neurocognitive framework for amotivated behaviour. Neuropsychologia 118:54–67.

Le Heron C, Kolling N, Plant O, Kienast A, Janska R, Ang Y-S, Fallon S, Husain M, Apps MAJ (2020) Dopamine Modulates Dynamic Decision-Making during Foraging. J Neurosci 40:5273–5282.

Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. Neuron 64:431–439.

Leiser D, Azar OH (2008) Behavioral economics and decision making: Applying insights from psychology to understand how people make economic decisions. Journal of Economic Psychology 29:613–618.

Levy I, Lazzaro SC, Rutledge RB, Glimcher PW (2011) Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. J Neurosci 31:118–125.

Ligneul C, Fernandes FF, Shemesh N (2021) High temporal resolution functional magnetic resonance spectroscopy in the mouse upon visual stimulation. Neuroimage 234:117973.

Lim S-L, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. J Neurosci 31:13214–13223.

Maes EJP, Sharpe MJ, Usypchuk AA, Lozzi M, Chang CY, Gardner MPH, Schoenbaum G, Iordanova MD (2020) Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. Nat Neurosci 23:176–178.

Maisson DJ-N, Cash-Padgett TV, Wang MZ, Hayden BY, Heilbronner SR, Zimmermann J (2021) Choice-relevant information transformation along a ventrodorsal axis in the medial prefrontal cortex. Nat Commun 12:1–14.

Marzecová A, Kaiser LF, Maddah A (2021) Neuromodulation of Foraging Decisions: The Role of Dopamine. Frontiers in Behavioral Neuroscience 15:1–4.

Mattson MP (2019) An Evolutionary Perspective on Why Food Overconsumption Impairs Cognition. Trends Cogn Sci 23:200–212.

Maunsell JHR (2004) Neuronal representations of cognitive state: reward or attention? Trends Cogn Sci 8:261–265.

McCutcheon RA, Krystal JH, Howes OD (2020) Dopamine and glutamate in schizophrenia: biology, symptoms and treatment. World Psychiatry 19:15–33.

Muthukumaraswamy SD, Edden RAE, Jones DK, Swettenham JB, Singh KD (2009) Resting GABA concentration predicts peak gamma frequency and fMRI amplitude in response to visual stimulation in humans. Proceedings of the National Academy of Sciences 106:8356–8361.

Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology 191:507–520.

Norbury A, Manohar S, Rogers RD, Husain M (2013) Dopamine Modulates Risk-Taking as a Function of Baseline Sensation-Seeking Trait. Journal of Neuroscience 33:12982–12986.

O'Doherty JP (2014) The problem with value. Neurosci Biobehav Rev 43:259–268.

Papageorgiou GK, Sallet J, Wittmann MK, Chau BKH, Schüffelgen U, Buckley MJ, Rushworth MFS (2017) Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. Nat Commun 8:1–14.

Pearson JM, Watson KK, Platt ML (2014) Decision making: the neuroethological turn. Neuron 82:950–965.

Philiastides MG, Biele G, Heekeren HR (2010) A mechanistic account of value computation in the human brain. Proc Natl Acad Sci U S A 107:9430–9435.

Pittenger C, Bloch MH, Williams K (2011) Glutamate abnormalities in obsessive compulsive disorder: neurobiology, pathophysiology, and treatment. Pharmacology & therapeutics 132:314–332.

Polanía R, Krajbich I, Grueschow M, Ruff CC (2014) Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. Neuron 82:709–720.

Pushkarskaya H, Tolin D, Ruderman L, Kirshenbaum A, Kelly JM, Pittenger C, Levy I (2015) Decision-making under uncertainty in obsessive-compulsive disorder. J Psychiatr Res 69:166–173.

Rescorla RA, Wagner A (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Current research and theory:64–99.

Salamone JD, Cousins MS, Bucher S (1994) Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. Behavioural Brain Research 65:221–229.

Santesso DL, Evins AE, Frank MJ, Schetter EC, Bogdan R, Pizzagalli DA (2009) Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. Hum. Brain Mapp. 30:1963–1976.

Scatton B, Rouquier L, Javoy-Agid F, Agid Y (1982) Dopamine deficiency in the cerebral cortex in Parkinson disease. Neurology 32:1039.

Scholl J, Günthner J, Kolling N, Favaron E, Rushworth MF, Harmer CJ, Reinecke A (2014) A role beyond learning for NMDA receptors in reward-based decision-making-a pharmacological study using d-cycloserine. Neuropsychopharmacol 39:2900–2909.

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599.

Schwerdt HN, Amemori K, Gibson DJ, Stanwicks LL, Yoshida T, Bichot NP, Amemori S, Desimone R, Langer R, Cima MJ, Graybiel AM (2020) Dopamine and beta-band oscillations differentially link to striatal value and motor control. Sci Adv 6:eabb9226.

Seamans JK, Gorelova N, Durstewitz D, Yang CR (2001) Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. Journal of Neuroscience 21:3628–3638.

Selten M, van Bokhoven H, Nadif Kasri N (2018) Inhibitory control of the excitatory/inhibitory balance in psychiatric disorders. F1000Res 7:23.

Shapiro AD, Grafton ST (2020) Subjective value then confidence in human ventromedial prefrontal cortex. PLoS One 15:e0225617.

Sharp ME, Viswanathan J, Lanyon LJ, Barton JJS (2012) Sensitivity and bias in decision-making under risk: evaluating the perception of reward, its probability and value. PLoS One 7:e33460.

Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79:217–240.

Shenhav A, Straccia MA, Botvinick MM, Cohen JD (2016) Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. Cogn Affect Behav Neurosci 16:1127–1139.

Shenhav A, Straccia MA, Cohen JD, Botvinick MM (2014) Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat Neurosci 17:1249–1254.

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci 16:966–973.

Stephens DW (2008) Decision ecology: Foraging and the ecology of animal decision making. Cogn Affect Behav Neurosci 8:475–484.

Strait CE, Blanchard TC, Hayden BY (2014) Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron 82:1357–1366.

Strait CE, Sleezer BJ, Hayden BY (2015) Signatures of Value Comparison in Ventral Striatum Neurons. PLoS Biol 13:e1002173.

Sutton RS, Barto AG (1998) Reinforcement Learning: An Introduction. Cambridge: The MIT Press.

Takahashi H, Matsui H, Camerer C, Takano H, Kodaka F, Ideno T, Okubo S, Takemura K, Arakawa R, Eguchi Y, Murai T, Okubo Y, Kato M, Ito H, Suhara T (2010) Dopamine D_1 receptors and nonlinear probability weighting in risky choice. J Neurosci 30:16567–16572.

Tervo DGR, Kuleshova E, Manakov M, Proskurin M, Karlsson M, Lustig A, Behnam R, Karpova AY (2021) The anterior cingulate cortex directs exploration of alternative strategies. Neuron 109:1876-1887.e6.

Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. Science 307:1642–1645.

Tognarelli JM, Dawood M, Shariff MIF, Grover VPB, Crossey MME, Cox IJ, Taylor-Robinson SD, McPhail MJW (2015) Magnetic resonance spectroscopy: principles and techniques: lessons for clinicians. Journal of clinical and experimental hepatology 5:320–328.

Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertainty 5:297–323.

Verharen JPH, Jong JW de, Roelofs TJM, Huffels CFM, van Zessen R, Luijendijk MCM, Hamelink R, Willuhn I, Ouden HEM den, van der Plasse G, Adan RAH, Vanderschuren, Louk J. M. J. (2018) A neuronal mechanism underlying decision-making deficits during hyperdopaminergic states. Nat Commun 9:1–15.

Vidaurre D, Myers NE, Stokes M, Nobre AC, Woolrich MW (2018) Temporally Unconstrained Decoding Reveals Consistent but Time-Varying Stages of Stimulus Processing. Cerebral Cortex 29:863–874.

Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, Hallett M (2010) Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. Neuron 65:135–142.

Wang X-J (2002) Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. Neuron 36:955–968.

Wang X-J (2008) Decision making in recurrent neuronal circuits. Neuron 60:215–234.

Wiecki TV, Sofer I, Frank MJ (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. Front. Neuroinform.:14.

Wilson RC, Collins AGE (2019) Ten simple rules for the computational modeling of behavioral data. eLife 8:e49547.

Winterer G, Weinberger DR (2004) Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. Trends Neurosci 27:683–690.

Wittmann MK, Kolling N, Akaishi R, Chau BKH, Brown JW, Nelissen N, Rushworth MFS (2016) Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. Nat Commun 7:12327.

Wolfe JM (2013) When is it time to move to the next raspberry bush? Foraging rules in human visual search. J Vis 13:10.

Wong K-F, Wang X-J (2006) A recurrent network mechanism of time integration in perceptual decisions. Journal of Neuroscience 26:1314–1328.

Wunderlich K, Rangel A, O'Doherty JP (2009) Neural computations underlying action-based decision making in the human brain. Proceedings of the National Academy of Sciences 106:17199–17204.

Wunderlich K, Rangel A, O'Doherty JP (2010) Economic choices can be made using only stimulus values. Proceedings of the National Academy of Sciences 107:15005–15010.

Yoo SBM, Hayden BY (2018) Economic Choice as an Untangling of Options into Actions. Neuron 99:434–447.

Eidesstattliche Erklärung

Ich versichere an Eides statt, dass die Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.
Die Dissertation wurde in der vorliegenden oder in ähnlicher Form noch bei keiner anderen Fakultät eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 12.05.2022

Luca Franziska Kaiser

List of Publications

The following section lists the publications in journals with peer-review process that this thesis is based on. It further details the contributions of each author to the respective articles

I.) Kaiser, L. F., Gruendler, T. O., Speck, O., Luettgau, L., & Jocham, G. (2021). Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions. *Nature communications*, *12*(1), 1-11.

Author contributions: Luca Franziska Kaiser analyzed the data, drafted the manuscript, read and edited versions of the manuscript and approved the final version of the manuscript. Theo Gruendler designed the study and conceptualized research, acquired the data, analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript. Oliver Speck set up the MRI acquisition protocol and MRS routine, read and edited versions of the manuscript and approved the final version of the manuscript. Lennart Luettgau read and edited versions of the manuscript and approved the final version of the manuscript. Gerhard Jocham designed the study and conceptualized research, analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript.

II.) Marzecová, A., Kaiser, L. F., & Maddah, A. (2021). Neuromodulation of Foraging Decisions: The Role of Dopamine. *Frontiers in Behavioral Neuroscience*, *15*, 1-4.

Anna Marzecová conceptualized the research, drafted the manuscript, participated in review-writing and editing. Luca Franziska Kaiser conceptualized the research, drafted the manuscript, participated in review-writing and editing. Armin Maddah conceptualized the research and supported investigation for the manuscript as well as editing and review-writing. All authors contributed to the article and approved the submitted version.

Attachments

Published Manuscripts

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-20875-w OPEN

Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions

Luca F. Kaiser[1,2,7✉], Theo O. J. Gruendler[2,3,7], Oliver Speck[2,4,5,6], Lennart Luettgau[1,2] & Gerhard Jocham[1,2]

In a dynamic world, it is essential to decide when to leave an exploited resource. Such patch-leaving decisions involve balancing the cost of moving against the gain expected from the alternative patch. This contrasts with value-guided decisions that typically involve maximizing reward by selecting the current best option. Patterns of neuronal activity pertaining to patch-leaving decisions have been reported in dorsal anterior cingulate cortex (dACC), whereas competition via mutual inhibition in ventromedial prefrontal cortex (vmPFC) is thought to underlie value-guided choice. Here, we show that the balance between cortical excitation and inhibition (E/I balance), measured by the ratio of GABA and glutamate concentrations, plays a dissociable role for the two kinds of decisions. Patch-leaving decision behaviour relates to E/I balance in dACC. In contrast, value-guided decision-making relates to E/I balance in vmPFC. These results support mechanistic accounts of value-guided choice and provide evidence for a role of dACC E/I balance in patch-leaving decisions.

[1] Biological Psychology of Decision Making, Institute of Experimental Psychology, Heinrich Heine University, Düsseldorf, Germany. [2] Center for Behavioral Brain Sciences, Otto von Guericke University, Magdeburg, Germany. [3] Center for Military Mental Health, Military Hospital Berlin, Berlin, Germany. [4] Leibniz Institute for Neurobiology, Magdeburg, Germany. [5] German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany. [6] Department of Biomedical Magnetic Resonance, Institute for Physics, Otto von Guericke University, Magdeburg, Germany. [7] These authors contributed equally: Luca F. Kaiser, Theo O. J. Gruendler. ✉email: kaiserl@hhu.de

In an ever-changing world with non-uniformly distributed goods, organisms have to decide whether they want to accept the resources provided by their current environment or switch to an alternative course of action. These patch-leaving decisions require balancing potential benefits in alternative environments against costs associated with abandoning the current patch (or current course of action). Patch-leaving decisions can be contrasted with value-guided choices, where agents often need to integrate multiple attributes to select the option currently most valuable. Consider for example a researcher working at a university in a small town who considers moving to Munich (famous in Germany for high cost of living). Initially, there is no benefit in leaving: moving costs money and the cost for living is higher in Munich. However, better career prospects and a higher salary might, in the long run, overcompensate the financial and social costs incurred. This constitutes a patch-leaving decision. In Munich, our researcher may face a decision on where to live—and attributes like the quality of different apartments, the rent, and the distance from the office may contribute to how valuable each flat is judged. Based on these attributes, our researcher would simply select the one they judge more valuable altogether, thus maximizing immediate reward. This kind of decision is commonly referred to as a value-based decision.

Studies in animals and humans suggest a role of the dorsal anterior cingulate cortex (dACC) in patch-leaving decisions[1–4], as well as signalling potential costs entailed by behavioural adjustments[2,3]. Activity in the dACC has been reported to reflect diminished rewards within the current environment[2,5,6] as well as the average value of potential alternatives[3] suggesting an important role in guiding behavioural adjustments[7]. In contrast, value-guided decision-making has been linked to the ventromedial prefrontal cortex (vmPFC)[8–16]. Activity in this region covaries with the values of the available options, positively with the value of the chosen option, and negatively with the value of the unchosen option[8–10,14]. Both theoretical and experimental results strongly suggest that a mechanism based on competition via mutual inhibition in vmPFC supports value-guided choice[9,15,17]. This competition is driven by the balance between GABAergic inhibition and recurrent glutamatergic excitation. Concentrations of the major excitatory and inhibitory neurotransmitters, glutamate and GABA, have been shown to be related to both choice performance and a vmPFC value comparison signal in a manner that is consistent with biophysical models[9,17]. Animal studies[2,4] suggest a similar role for the balance between glutamate and GABA in patch-leaving decisions but in dACC rather than in vmPFC[18].

We hypothesized that patch-leaving behaviour is guided by the balance between cortical excitation and inhibition (E/I balance) in dACC. In contrast, we expected that value-guided decision-making is governed by E/I balance in vmPFC. Healthy human participants performed a decision-making task (Fig. 1a) combining patch-leaving and value-based decision-making. We measured GABA and glutamate concentrations using magnetic resonance spectroscopy (MRS) at 7 T in five cortical areas of interest: vmPFC, dACC, dorsolateral prefrontal cortex (dlPFC), and bilateral primary motor cortex (M1). Specifically, we predicted that interindividual differences in how costs and patch values influence behaviour relates to variations in E/I balance in dACC[2,4,18] over and above the effects of all other voxels of interest. Further, we predicted that decision performance during value-guided choice would depend on vmPFC E/I balance[9,19]. Additionally, models based on competition by mutual inhibition predict that the speed at which a decision unfolds is driven by the available evidence, and the rate of this evidence accumulation is again crucially dependent upon E/I balance[9]. We therefore further predicted that the effect of the key decision variables on response times would also be related to E/I balance in dACC and vmPFC, respectively[20].

We report contributions of E/I balance that are dissociable as a function of decision type and cortical area. Patch-leaving behaviour is related to E/I balance in dACC but not in any of the other regions investigated. In contrast, value-guided decision-making is related to E/I balance in vmPFC but not in any of the other cortical areas.

Results

Participants ($N = 29$) performed 320 trials of a behavioural task combining patch-leaving and value-guided choice. Each trial of the behavioural task consisted of a patch-leaving decision followed by a value-guided choice (Fig. 1a). Importantly, the task was designed such that the value-guided choice was explicitly temporally separated from the choice to leave or stay in the current patch. At the first stage, participants indicated by button press whether they wanted to stay in their current patch or leave for the alternative patch. Leaving was associated with a cost (randomly drawn from the set {5, 10, 15, 20}), which was subtracted from the participant's current total earnings. Over trials, the reward available in the current patch stochastically depleted according to a decaying Gaussian Random Walk, whereas the reward in the alternative patch replenished. The cost level was displayed to participants and remained constant until a decision to leave the patch was made, at which time a new cost level was randomly selected. Thus participants needed to monitor, over trials, the relative advantage of leaving for the alternative patch and to compare this against the cost for leaving. No money could be won at this stage. Following the patch decision stage, participants entered the value-guided choice. Here the reward available in the patch chosen by the participant was randomly divided and allocated to two choice options. Additionally, a probability with which this reward could be obtained was randomly assigned to each of these two options. This design feature ensured that the value-guided choice was temporally decorrelated from the choice to leave or stay in the current patch. While being in a rich patch will, on average, lead to better choice options at the value-guided choice stage, the exact options to choose from are not known to participants when they make their patch choice. After choice, participants received a feedback on whether their choice had been rewarded. This was followed by the next trial. In a separate session, 24–48 h after volunteers completed the behavioural task, we obtained estimates of GABA and glutamate concentrations in five cortical areas of interest (Fig. 1b and Supplementary Fig. 3) using single-voxel MRS at 7 T (see Fig. 1c for an example spectrum from one participant). We recorded from the dACC, the vmPFC, the right dlPFC, and the bilateral primary motor cortices (M1). Note that the vmPFC voxel is located in a rather dorsal position, covering parts of pregenual ACC. This location, which is also in line with previous work[9], was chosen based on methodological considerations since obtaining MRS measurements in more ventral positions is difficult due to field inhomogeneities. However, please also note that value signals in vmPFC, while not centred on this location, often extend to cover this region across a large swath of the ventral to dorsal extent of the mesial prefrontal cortex[9]. In addition to vmPFC and dACC, we selected the dlPFC because of its importance for working memory-related processes[21,22]. Since patch leaving requires carrying a representation of patch-leaving value across trials, we reasoned that dlPFC E/I balance might play a role in patch-leaving but not value-guided choice behaviour. The motor cortex was selected as a control region, where we expected a relationship with motor, but not task-specific parameters, neither value nor patch leaving related.

We used multiple linear regression to test our hypotheses. In order to limit the number of statistical comparisons, we proceeded

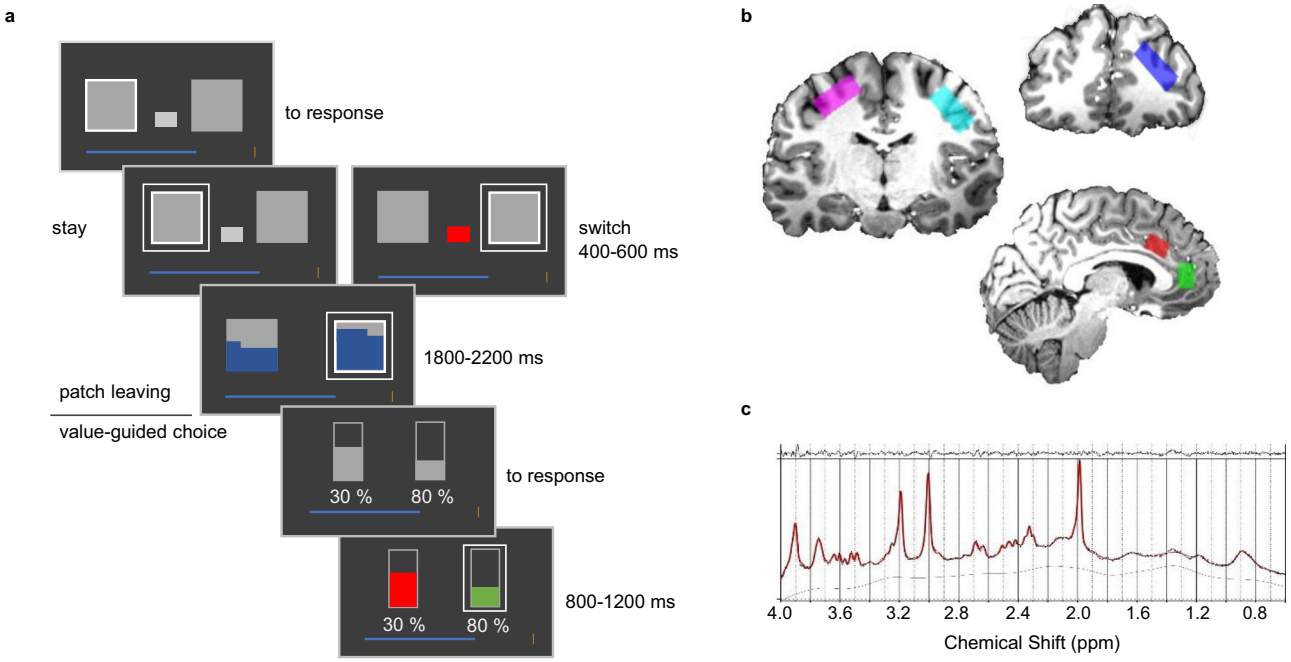


Fig. 1 Behavioural task and MRS recordings. a Schematic of the task structure. Participants made a patch decision in the first stage of each trial. A white outline indicated the location of the participant's current patch. If they chose to switch, they had to pay a cost indicated by the size of a grey rectangle. Following the choice, the values of both patches were revealed, indicated by the blue fillings. In the value-guided choice phase, the chosen patch value was randomly divided between the two different options and reward probabilities were randomly assigned to them. Participants selected an option by pressing a button, which was followed by feedback on both options. If they obtained a reward, the blue progress bar at the bottom of the screen grew in proportion to this reward. **b** Example MRS voxel placement for one participant (Supplementary Fig. 3 for overlay of all participants). Spectroscopy voxels were placed in right dlPFC (blue), bilateral primary motor cortex (pink and cyan), dACC (red) and vmPFC (green). **c** Spectrum obtained from dACC of one exemplary participant.

along the following hierarchy: First, we only tested behavioural variables, which we hypothesized to relate to E/I balance (see above). Second, using a general linear model (GLM), we first projected E/I balance (ratio between glutamate and GABA) from all five regions of interest onto main behavioural parameters from the patch-leaving and value-guided choice phase. All of these analyses were performed exclusively using the design matrix containing E/I balance from all five regions. Third, if, and only if, this GLM yielded a significant effect for one brain region, we followed this up by asking whether glutamate, GABA, or both within that specific region contributed to this effect of E/I balance. To this end, we then computed partial correlations, regressing out the effects of all other factors than the one currently of interest (see 'Methods'). These partial correlations can therefore be thought of as post hoc test, further investigating the individual contributions of GABA and glutamate to a main effect of E/I balance (if present).

Patch-leaving behaviour. Participants took costs and patch value differences into account in guiding their patch-leaving choices. On average, participants left their current patch on 20.55 ± 1.40 (mean ± SEM) out of 320 trials. We found that participants stayed longer in their current patch when they had to pay higher switch costs. The average (across-participants mean of the median per cost level) patch value differences (alternative − current patch) at which participants left their current patch increased with cost level (repeated-measures analysis of variance (RM-ANOVA): $F_{3,81} = 5.941$, $p = 0.001$, $\eta^2 = 0.063$, significant positive linear trend: $t_{27} = 3.961$, $p < 0.001$, confidence interval $(CI)_{95} = [1.341–4.223]$, Cohen's $U3_1$ for one sample = 0.179; Fig. 2a). To quantify how participants balanced patch values against cost, we computed a patch-leaving advantage by subtracting, for every patch-leaving trial, the switch costs from the relative benefit of leaving (alternative − current patch value). These patch-leaving advantages were then averaged across switch trials. The average patch-leaving advantage across subjects was 19.39 ± 2.07 (mean ± SEM, see Fig. 2b for an evolution of patch-leaving advantages across all trials for one example participant).

To investigate the factors governing the speed of responding, we set up a multiple linear regression model. Patch value difference, patch-leaving trials, cost levels, trial number, switch (left/right) of patch presentation (relative to the previous trial), and wins in the previous trials were entered as independent variables to predict (logarithmic) response times. Participants' responses showed a trend of being slower when switching entailed greater costs ($t_{28} = 1.844$, $p = 0.076$, $CI_{95} = [-0.003$ to $0.056]$, $U3_1 = 0.345$). Furthermore, they responded slower in switch trials ($t_{28} = 3.683$, $p = 0.001$, $CI_{95} = [0.040$ to $0.138]$, $U3_1 = 0.276$), when they had received reward at the value-guided choice stage of the previous trial ($t_{28} = 3.733$, $p = 0.001$, $CI_{95} = [0.021$ to $0.071]$, $U3_1 = 0.207$) and when there was a change in presentation sides of patch values ($t_{28} = 3.952$, $p = 0.001$, $CI_{95} = [0.026$ to $0.082]$, $U3_1 = 0.276$). Further to this, participants' responding became significantly faster over the course of the experiment ($t_{28} = -7.985$, $p < 0.001$, $CI_{95} = [-0.317$ to $-0.188]$, $U3_1 = 0.897$). There was no significant effect of patch value difference ($t_{28} = 0.290$, $p = 0.774$, $CI_{95} = [-0.034$ to $0.045]$, $U3_1 = 0.517$) on reaction times in the patch-leaving phase. Similarly, trial-wise patch-leaving advantages had no significant effect on reaction times either (see Supplementary Notes 1).

Cortical E/I balance and patch-leaving behaviour. We computed a patch-leaving advantage that indicates how participants balance the relative benefit expected from leaving against the cost. Regressing E/I balance against patch-leaving advantage revealed a

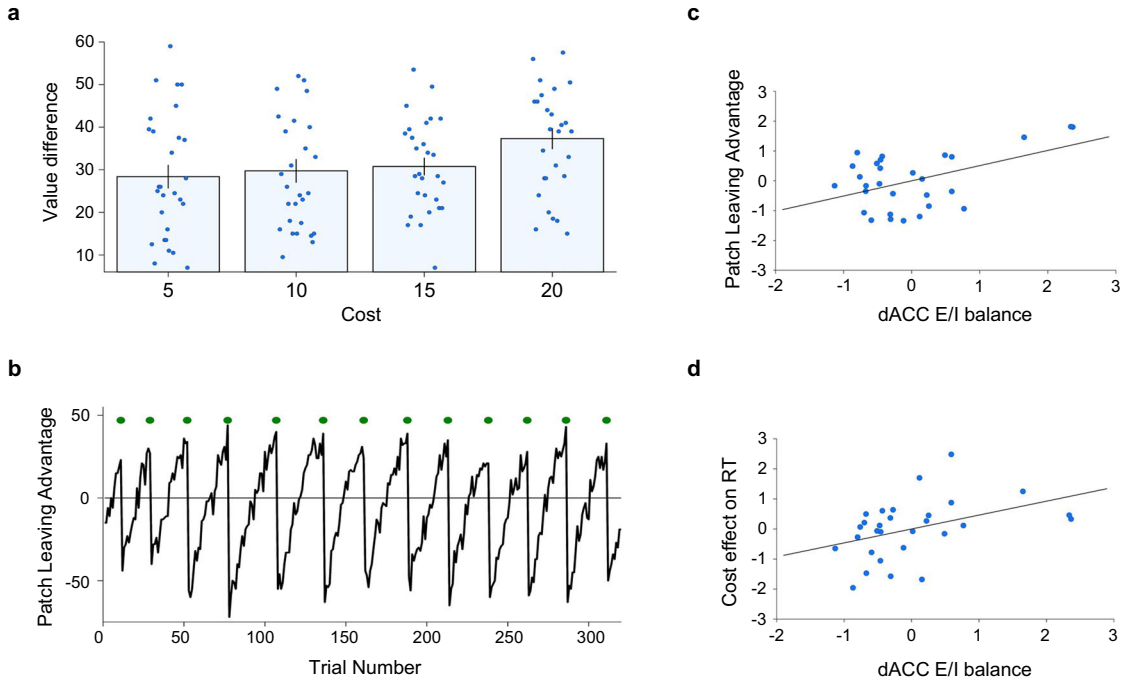


Fig. 2 Patch-leaving behaviour and cortical E/I balance. a Participants left their current patch at higher value differences (alternative − current patch value) when leaving was associated with higher costs (RM-ANOVA: $F_{3,81} = 5.941$, $p = 0.001$, $\eta^2 = 0.063$, $N = 28$). Individual data points are overlaid as dot plots. Bars represent across-participants mean of the median per cost level. Error bars indicate standard error of the mean. **b** Example timecourse of patch-leaving advantages (PLA) for one example participant. PLA = [value alternative patch − value current patch − cost]. Green circles indicate patch-leaving trials. **c** Participants with higher dACC E/I balance leave at higher average PLA (Pearson correlation on residuals (compare main text and methods): $r = 0.483$, $p = 0.008$, $CI_{95} = [0.141\text{–}0.722]$, $N = 29$). **d** Participants' patch-leaving decisions are slowed down with increasing cost levels, and this effect is most pronounced in participants with high levels of dACC E/I balance (Pearson correlation on residuals: $r = 0.415$, $p = 0.025$, $CI_{95} = [0.057\text{–}0.678]$, $N = 29$). Source data are provided as a Source data file.

significant influence of E/I balance in dACC ($t_{23} = 2.643$, $p = 0.015$, $CI_{95} = [0.111$ to $0.908]$; $r = 0.483$, $p = 0.008$, $CI_{95} = [0.141$ to $0.722]$, Fig. 2c) but not in any other region of interest (all $p > 0.199$). This effect was, by trend, driven by GABA in dACC ($r = -0.323$, $p = 0.087$, $CI_{95} = [-0.617$ to $0.049]$; Supplementary Fig. 1) but not glutamate in dACC ($r = -0.152$, $p = 0.431$, $CI_{95} = [-0.491$ to $0.227]$). In addition to this, a direct contrast showed that the relationship between patch-leaving advantage and E/I balance was stronger in dACC compared to both vmPFC ($t_{23} = 2.613$, $p = 0.016$) and dlPFC E/I ($t_{23} = 2.087$, $p = 0.048$). These findings suggest that the manner in which the relative benefit of leaving is balanced against travel costs is uniquely related to E/I balance in dACC but not in any of the other areas investigated.

We next assessed how cortical E/I balance was related to patch response speed and to how key task parameters affected response speed. We did not find any significant effects on overall response speed (all $p > 0.198$) but a specific effect of E/I balance in dACC on the degree to which patch decision choices were slowed by costs. Participants showed slowing of patch choices with higher cost levels, and the magnitude of this effect was related to E/I balance in dACC ($t_{23} = 2.187$, $p = 0.039$, $CI_{95} = [0.025$ to $0.900]$; $r = 0.415$, $p = 0.025$, $CI_{95} = [0.057$ to $0.678]$; Fig. 2d).

Please note that the results displayed in Fig. 2c, d primarily serve to illustrate the effects obtained in our main GLM. Nevertheless, inspection of these panels reveals three data points that appear further away from the remainder of the data. Therefore, we have additionally analysed the same data (the residuals of dACC E/I balance and behaviour) with a robust regression analysis. This confirmed the pattern of results reported above (Fig. 2c: $t_{27} = 2.665$, $p = 0.013$, $CI_{95} = [0.120$ to $0.924]$; Fig. 2d: $t_{27} = 2.034$, $p = 0.052$, $CI_{95} = [-0.004$ to $0.810]$). Thus far, our results are consistent with our hypothesis. E/I balance in dACC, but not in any of the other regions investigated, is related both to how participants balance expected benefits against travel costs and to how costs affect the speed at which patch decisions are made.

Value-guided choice behaviour. Participants selected the objectively correct option (higher expected value) in 81.85 % ± 1.52 (mean ± SEM) of all trials. We set up a logistic regression model to investigate the factors that affected participants' decisions (right vs left option). Participants choices were strongly guided by the differences (right minus left) in expected values between options ($t_{28} = 6.422$, $p < 0.001$, $CI_{95} = [3.184$ to $6.166]$, $U3_1 = 0$). As expected, value sum had no significant effect on choice ($t_{28} = -1.380$, $p = 0.179$, $CI_{95} = [-0.268$ to $0.052]$, $U3_1 = 0.621$). Additionally, there was a significant effect of no-brainer trials (trials in which both probability and magnitude favoured one option) on choice ($t_{28} = 5.983$, $p < 0.001$, $CI_{95} = [0.196$ to $0.400]$, $U3_1 = 0.103$), which is likely due to an increased occurrence of no-brainer trials favouring the right option (see 'Methods'). There was no significant influence of either patch value difference, switch costs, whether the current trial was a switch trial, the current trial's patch choice (left/right), the value-guided choice from the previous trial, whether this choice had been rewarded, and of trial number (all $p > 0.486$). This indicates that participants' value-guided choices were guided by the key value-related parameters, not by other aspects, such as whether a choice had been rewarded on the previous trial.

To accumulate maximal returns, participants need to compute the Pascalian expected values by multiplying reward probabilities

and reward magnitudes and then choose the option with the higher expected value. However, humans do not weigh probabilities and magnitudes in a statistically optimal way and show systematic distortions[8,23]. We fitted several different models to explain how participants combine reward probabilities and magnitudes[23]. The models incorporated different utility functions to represent distortions in the weighting of reward information and assumed either multiplicative or additive strategies to combine reward probabilities and magnitudes. We found that choices were best explained by a model assuming multiplicative value integration with non-linear probability and magnitude weighting[23].

$$V = \omega_{\mathrm{mult}} * (u(m_{\mathrm{O}}) * w(p_{\mathrm{O}})) \quad (1)$$

$$w(p_{\mathrm{O}}) = \frac{{p_{\mathrm{O}}}^{\gamma}}{({p_{\mathrm{O}}}^{\gamma} + (1 - p_{\mathrm{O}})^{\gamma})^{1/\gamma}} \quad (2)$$

$$u(m_{\mathrm{O}}) = {m_{\mathrm{O}}}^{\alpha} \quad (3)$$

where m_{O} and p_{O} are the objective reward magnitudes and probabilities that are transformed into subjective reward magnitudes and probabilities, respectively, with the shape of the functions governed by the free parameters α and γ. The parameter ω_{mult} scales the effect of value difference and thus corresponds to a softmax inverse temperature. For model fitting, we fixed ω_{mult} at 6.62 (see 'Methods', Supplementary Table 6, and Supplementary Information 4 for parameter recovery).

Finally, we investigated whether the same variables used to predict choices have a significant effect on normalized (log) response times. The only difference from the model used to predict binary choice (of right option) is that we used absolute expected value differences here (rather than right minus left values), since we did not expect any effect conditional on side of presentation. Participants exhibited faster responding with greater value difference between options ($t_{28} = -5.928$, $p < 0.001$, $CI_{95} = [-0.273$ to $-0.133]$, $U3_1 = 0.862$). Value sum had no significant effect ($t_{28} = -1.543$, $p = 0.134$). Furthermore, participants showed faster responding on trials with high patch value difference ($t_{28} = -7.105$, $p < 0.001$, $CI_{95} = [-0.156$ to $-0.086]$, $U3_1 = 1$), in no-brainer trials ($t_{28} = -14.344$, $p < 0.001$, $CI_{95} = [-0.369$ to $-0.277]$, $U3_1 = 1$), and with increasing trial number ($t_{28} = -8.119$, $p < 0.001$, $CI_{95} = [-0.261$ to $-0.156]$, $U3_1 = 0.931$). Finally, we found significantly slower responses in patch-leaving trials ($t_{28} = 4.590$, $p < 0.001$, $CI_{95} = [0.026$ to $0.067]$, $U3_1 = 0.276$, Fig. 3a). Neither cost levels, the previous trial's value-guided choice, nor whether reward had been received in theprevious trial had an effect on reaction times in the value-guided choice phase (all $p > 0.126$). However, participants responded more slowly during value-guided choice when they had chosen the right patch during patch leaving ($t_{28} = 2.538$, $p = 0.017$, $CI_{95} = [0.005$ to $0.043]$, $U3_1 = 0.345$).

Cortical E/I balance and value-guided choice behaviour. To relate cortical neurochemistry to value-guided choice behaviour, we used the same approach as above for the patch-leaving phase. Similar to our previous work[9], we found that E/I balance in vmPFC was related to value-guided choice performance. Decision accuracy (percentage of choices of the higher value option) was negatively related to E/I balance in vmPFC ($t_{22} = -2.437$, $p = 0.023$, $CI_{95} = [-0.947$ to $-0.076]$; $r = -0.461$, $p = 0.012$, $CI_{95} = [-0.708$ to $-0.114]$; Fig. 3b) but not in any of the other regions investigated ($p > 0.406$), indicating that subjects with higher concentrations of GABA relative to glutamate were better at selecting the higher value option. When we followed this up with partial correlations, neither GABA ($r = 0.234$, $p = 0.221$) nor glutamate alone ($r = -0.225$, $p = 0.241$) was significantly correlated with decision accuracy. These findings indicate that participants with higher concentrations of glutamate relative to GABA in vmPFC indeed tend to exhibit less accuracy in their choice behaviour. We found the same relationship with vmPFC E/I balance when we used the regression coefficients for expected value differences (that is, the degree to which participants' choices were guided by the value difference between options, see Supplementary Notes 2) instead of percentage of correct choices.

To further investigate the relationship between E/I balance and value-guided choice behaviour, we fitted a behavioural model accounting for systematic deviations in the weighting of reward information. We found that the extent to which participants distort reward magnitudes in guiding their choices (model parameter α) was significantly related to vmPFC E/I balance ($t_{22} = -2.409$, $p = 0.025$, $CI_{95} = [-0.945$ to $-0.071]$; $r = -0.457$, $p = 0.013$, $CI_{95} = [-0.705$ to $-0.109]$; Fig. 3c), again without any effect of the other four regions ($p > 0.267$). This effect showed a trend of being influenced by GABA in vmPFC ($r = 0.339$, $p = 0.073$, $CI_{95} = [-0.032$ to $0.627]$). Additionally, we found a significant relationship between γ and E/I balance in vmPFC ($t_{22} = 2.144$, $p = 0.043$, $CI_{95} = [0.015$ to $0.878]$; $r = 0.416$, $p = 0.025$, $CI_{95} = [0.058$ to $0.679]$) but not in any other region of interest (all $p > 0.327$). The effects of α and γ potentially mediate the influence of vmPFC E/I balance on choice accuracy. After adding α and γ to the E/I design matrix, we did not find any significant effect of vmPFC E/I balance on choice accuracy anymore ($t_{20} = 0.361$, $p = 0.722$). We found a similar effect instead when using coefficients from our logistic regression analysis, showing that a greater reliance on reward probabilities compared to magnitudes was related to E/I balance in vmPFC (Supplementary Notes 2).

Taken together, value-guided choice performance was related to E/I balance in vmPFC but not in any of the other cortical regions. Participants with high levels of GABA relative to glutamate were most reliable at selecting the higher value options, possibly due to a more optimal weighting of reward magnitudes. Finally, we asked how value-guided response speed was related to cortical E/I balance. Theoretical models predict that higher levels of inhibition will lead to more pronounced slowing on difficult decisions (trials with low value difference[9]). We first observed that overall response times in the value-guided choice phase were specifically related to dACC E/I balance ($t_{22} = -2.423$, $p = 0.024$, $CI_{95} = [-0.975$ to $-0.076]$; $r = -0.459$, $p = 0.012$, $CI_{95} = [-0.707$ to $-0.111]$, Supplementary Fig. 2B). This effect was contributed by a positive effect of GABA ($r = 0.452$, $p = 0.014$, $CI_{95} = [0.102$ to $0.702]$, Supplementary Fig. 2C), with no significant effect of glutamate ($r = 0.052$, $p = 0.789$). In contrast to these general effects of dACC E/I balance on overall response speed, we found a specific effect of vmPFC E/I balance on the degree to which responses were speeded up by high value difference. The effect of value difference on response times in the value-guided choice phase was lowest in individuals with high vmPFC E/I balance ($t_{22} = 2.877$, $p = 0.009$, $CI_{95} = [0.158$ to $0.972]$; $r = 0.523$, $p = 0.004$, $CI_{95} = [0.193$ to $0.746]$, Fig. 3d). Specifically, GABA levels correlated negatively with this effect ($r = -0.357$, $p = 0.057$, $CI_{95} = [-0.640$ to $0.011]$; Supplementary Fig. 2A). These findings indicate that participants' responses slowed down on difficult trials with low value difference and that this slowing was most pronounced in individuals with relatively higher levels of GABA compared to glutamate. This pattern is consistent with our previous findings showing that vmPFC decision signals emerged more rapidly with higher concentrations of glutamate and low levels of GABA[9].

In conclusion, this pattern of results mirrors the findings from the patch-leaving phase. Whereas patch-leaving behaviour was

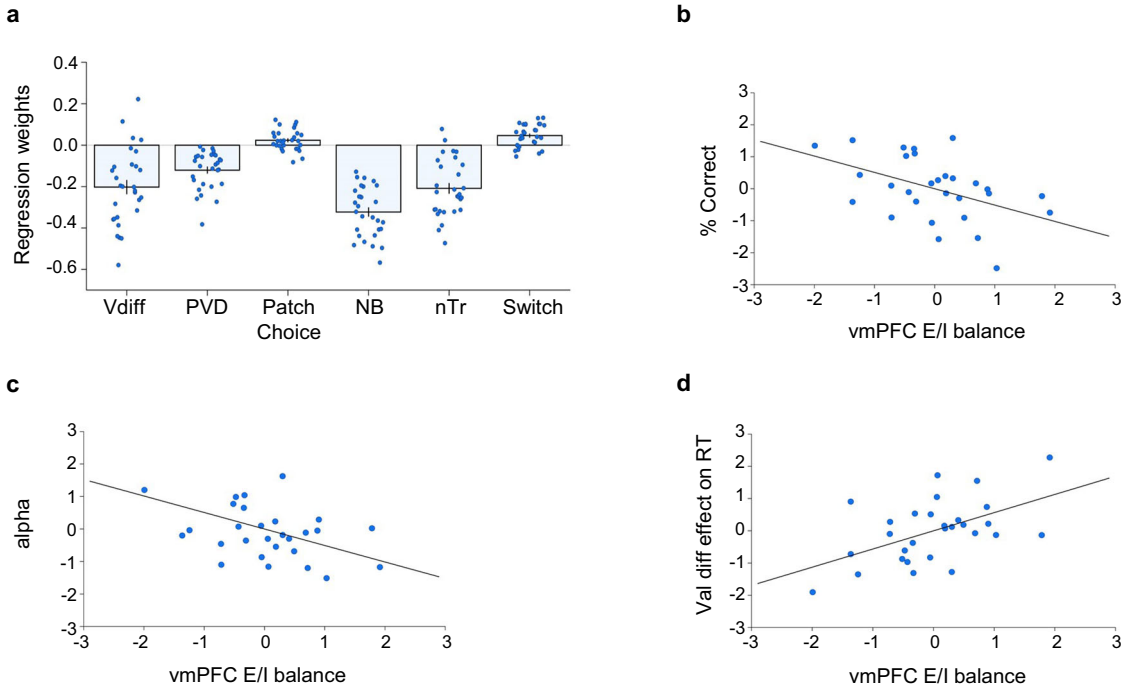


Fig. 3 Value-guided choice behaviour and cortical E/I balance. a Regression showing that participant's reaction times are guided by value differences between choice options (Vdiff; two-sided one-sample *t* test against zero: $t_{28} = -5.928$, $p < 0.001$, $CI_{95} = [-0.273$ to $-0.133]$, $U3_1 = 0.862$), patch value differences (PVD; two-sided one-sample *t* test against zero: ($t_{28} = -7.105$, $p < 0.001$, $CI_{95} = [-0.156$ to $-0.086]$, $U3_1 = 1$), choices in the patch phase (two-sided one-sample *t* test against zero: $t_{28} = 2.538$, $p = 0.017$, $CI_{95} = [0.005$ to $0.043]$, $U3_1 = 0.345$), whether each trial is a no-brainer trial (NB; two-sided one-sample *t* test against zero: $t_{28} = -14.344$, $p < 0.001$, $CI_{95} = [-0.369$ to $-0.277]$, $U3_1 = 1$), trial number (nTr; two-sided one-sample *t* test against zero: $t_{28} = -8.119$, $p < 0.001$, $CI_{95} = [-0.261$ to $-0.156]$, $U3_1 = 0.931$) and whether each trial was a patch-leaving trial (Switch; two-sided one-sample *t* test against zero: $t_{28} = 4.590$, $p < 0.001$, $CI_{95} = [0.026$ to $0.067]$, $U3_1 = 0.276$). Individual data points are overlaid as dot plots. Bars represent mean values across participants. Error bars indicate standard error of the mean. **b** Accuracy of value-guided choice is highest in participants with low levels of E/I balance in vmPFC (Pearson correlation on residuals: $r = -0.461$, $p = 0.012$, $CI_{95} = [-0.708$ to $-0.114]$). **c** Distortions in reward magnitude weighting relate to E/I balance in vmPFC (Pearson correlation on residuals: $r = -0.457$, $p = 0.013$, $CI_{95} = [-0.705$ to $-0.109]$). **d** Participants with high levels of E/I balance in vmPFC exhibit less slowing in difficult trials (Val diff = value difference) during value-guided choice (Pearson correlation on residuals: $r = 0.523$, $p = 0.004$, $CI_{95} = [0.193$ to $0.746]$). $N = 29$ in all figures. Source data are provided as a Source data file.

specifically related to dACC E/I balance, key parameters of value-guided choice behaviour were related to vmPFC E/I balance in a consistent and mechanistically plausible manner.

Discussion

Knowing when to leave a depleting resource is a central problem for decision makers in naturalistic environments. It requires the agent to track the value of current resources, compare it to potential alternatives, and balance the potential benefits of moving against the cost incurred by moving. Within a given environment, it is crucial to consider the various attributes that jointly determine an option's value—and then to select the most valuable option in order to maximize rewards. Thus both patch-leaving and value-guided decisions are key elements of adaptive behaviour. Using a behavioural task and assessment of cortical E/I balance by MRS quantification of GABA and glutamate concentrations, we have provided evidence for a double dissociation: E/I balance in dACC, but not in any of the other regions investigated, was related to the manner in which participants balanced potential benefits of leaving against costs during patch-leaving decisions. In contrast, E/I balance in vmPFC was related to various aspects of value-guided choice.

Participants took costs into account in guiding their patch-leaving choices, as evident from the finding that they waited for higher advantages (higher value difference between current and alternative patch) as cost levels increased, and participants who required higher advantages compared to travel costs were characterized by high dACC E/I balance. Similarly, these participants also showed stronger slowing of patch response times with increasing cost levels. An extensive literature has implicated neural activity in dACC in behavioural adjustments[24–28]. Recently, these patterns of activity have been recast in light of new evidence suggesting that dACC may encode the evidence in favour of switching away from a current default option[11]. Specifically, dACC activity contained information about the value of searching the environment for better alternatives compared to the currently available options[3]. In both primates and rodents, firing of neurons in ACC ramps up just before the animal is about to abandon its current patch and move elsewhere[2,4] or when rats abandoned current beliefs and explored alternative strategies[29]. Similarly, ACC local field potentials in the gamma range have been related to switching between exploratory and exploitative modes of behaviour[29–31]. Since gamma oscillations are driven by a balance between glutamatergic E/I by GABAergic interneurons[32–34], it was plausible for us to assume a role for cortical E/I balance in patch-leaving decisions. We found that participants with higher E/I balance (higher levels of glutamate relative to GABA) required a higher patch-leaving advantage (a higher difference between the benefits and costs expected from leaving) and showed more pronounced slowing of patch response times when costs were high. The effect of E/I balance on patch-leaving advantages was, as a trend, contributed to by GABA, but not by glutamate levels. Our findings are in line with previous reports

showing a relationship between inhibitory neurotransmission in the dACC and patch-leaving behaviour. Parvalbumin-positive GABA interneurons in rodent anterior cingulate cortex have been shown to ramp up their firing prior to the animal leaving its current patch, and firing rates of these neurons represented the animal's stay duration in the patch[4]. Furthermore, interhemispheric gamma synchronization driven by the same class of GABAergic interneurons in medial prefrontal cortex has recently been shown to enable mice to adaptively respond to changing environments[35]. Together with our results, these findings suggest that GABAergic activity in dACC may provide a signal for leaving one's current patch. Another study in primates found a similar ramping pattern in dACC neurons[2]. While the cell type from which these recordings were obtained is not known, it is likely that they were predominantly obtained from glutamatergic pyramidal cells[36,37]. This may appear contradictory at first glance but could be easily reconciled when assuming that there is an asymmetry in the proportion of neuronal pools whose activity represents the value of leaving the current patch vs those that represent the value of staying. Such an assumption is plausible given that dACC has been shown to dominantly represent value of switching away from a current strategy[3,7,38]. Under such a scenario, in a recurrently connected network with GABAergic feedback inhibition, a ramping of pyramidal cell firing would recruit feedback inhibition, which would then further increase the asymmetry between the neuronal pools, gradually favouring the pools representing the value of switching. Thus increased levels of GABAergic feedback inhibition would amplify the network transition towards favouring the alternative option and consequently bias the agent towards leaving the patch earlier. However, unlike for value-guided choice (see below), while a hypothetical model has been postulated[18], to date there exist no biophysically plausible mechanistic models for patch-leaving behaviour.

In contrast to patch-leaving decisions, value-guided choice was specifically related to E/I balance in vmPFC: high vmPFC concentrations of GABA relative to glutamate were related to an increased decision accuracy (selection of the higher value option) and a more optimal weighting of reward magnitudes. Furthermore, vmPFC GABA concentrations were also related to how participants slowed on difficult trials (choices with low value difference), with participants with high GABA concentrations again showing more pronounced slowing. These results are in line both with mechanistic models of decision-making[17,39] and our own previous findings[9]. It is thought that decisions may be generated by a mechanism that is based on competition via mutual inhibition in recurrent cortical networks that exhibit attractor dynamics[39,40]. In these models, a winner-take-all competition is implemented, where (in the binary case) activity in only one of the two pools representing the two options remains (the chosen option), whereas activity in the other pool is suppressed. One key prediction of these models is that increased GABAergic feedback inhibition slows down the attractor dynamics, allowing for more evidence to be accumulated[9,17,20]. Thus increased GABAergic tone makes decisions slower but more accurate. In our previous work, we showed that higher concentrations of GABA and low concentrations of glutamate were related to increased decision accuracy. Neurally, this was accompanied by a slower but more stable ramping of a value difference correlate in vmPFC, a neural signature of a decision[9]. Our present results match with this pattern. Choice performance was highest in participants with high vmPFC concentrations of GABA relative to glutamate, and these participants also showed the most pronounced slowing on difficult trials. Previously, we had reported a relationship between vmPFC E/I balance and choice stochasticity[9,19] whereas in the current study we find a relationship with choice accuracy. This discrepancy is likely due to differences in task structure. Previously[9], reward magnitudes had been independent between the two options and occupied a fixed range across participants. In contrast, here, they are the result of the current patch value (with the two options' magnitudes summing up to the patch value). This has two consequences. First, trials with low magnitude difference are less likely to occur. Second, since the range of magnitudes covered is dependent on each participant's pattern of patch leaving, estimates of choice stochasticity (softmax inverse temperature) are poorly comparable across participants. Note that we also set the inverse temperature to a fixed value for fitting the models.

We found task-specific effects for patch-leaving and value-guided choice in dACC and vmPFC, respectively. However, while participants were significantly influenced by value differences between choice options during value-guided choice, we did not find any significant influence of patch value difference on reaction times during patch-leaving decisions. This discrepancy between the two stages may appear surprising at first glance. However, response times likely indicate rather different factors in the two stages. In the patch stage, participants can already make up their mind whether to switch or stay on the next trial immediately after they observe the outcome of their patch choice. In contrast, at the value-guided choice, participants cannot anticipate the options they will encounter and instead have to compute option values on the fly.

A notable aspect of our findings is that response times in each phase were modulated by events from the respective other phase. Participants' value-guided choices speeded up as the value difference between the alternative and current patch increased on trials leading up to a switch, but when participants chose to leave their patch, the immediately subsequent value-guided choice was slowed down. Conversely, patch-leaving decisions were slowed when the previous trial's value-guided choice had been rewarded. The functional significance of these effects is not clear, but the former might indicate that participants switch to a more cautious, evaluative mode of value-guided choice upon entering the alternative patch. A recent modelling account[38] suggests an interplay between dorsomedial prefrontal cortex (including dACC) and vmPFC in deciding when to switch away from ongoing behaviour, based on reliability ratings of the current strategy[41]. In our task, rewards are only obtained during value-guided choice. These potentially serve as a feedback on the current strategy, which in turn might mediate a switch from ongoing behaviour. This explanation potentially also relates to the effects of dACC E/I balance on reaction times during value-guided choice. It has been suggested previously[42] that increased glutamate levels in dACC lead participants to exploit underlying task structure, whereas increased GABA concentrations allow for learning a new model of that task.

In summary, we have shown that cortical E/I balance, as assessed by MRS quantification of baseline GABA and glutamate concentrations, is related to both patch-leaving and value-guided decision-making. We found a double dissociation, where E/I balance in dACC is related to patch-leaving, but E/I balance in vmPFC is related to value-guided choice. The pattern of results further supports models that implement a competition via mutual inhibition in recurrent cortical networks as a candidate mechanism for value-guided choice. Importantly, we provide evidence that relates dACC E/I balance to patch-leaving decisions. The pattern of results suggests that elevated GABAergic relative to glutamatergic tone in dACC may increase the propensity to switch away from a current policy. Understanding the neurochemical mechanisms underlying different types of decision-making is of potential clinical relevance, since alterations in E/I balance have been described in a number of

neuropsychiatric disorders[43,44], which are characterized by impaired decision-making behaviour[45,46].

Methods

Participants. Thirty-three right-handed (Oldfield-Score[47]: 91.95 ± 1.89, mean ± SEM) male participants (age: 26.18 ± 0.65 years, mean ± SEM, range: 22–36 years) with normal ($N = 16$) or corrected to normal ($N = 17$) vision participated in this experiment. Exclusion criteria comprised a history of neurological or psychiatric illness, drug abuse, and use of psychoactive drugs or medication 24 h prior to participation. Four subjects were excluded due to excessive noise in at least one of the five spectroscopic measurements (see MR imaging (MRI) for criteria). All reported results are from the remaining $N = 29$ subjects (mean age: 26.48 ± 0.72 years, range: 22–36 years; normal vision: $N = 14$; non-smoker: $N = 22$). Written informed consent to the procedure was obtained from all subjects prior to the experiment, which was approved by the local ethics committee of the Medical Faculty of the Otto-von-Guericke-University, Magdeburg. Participants were compensated for each session and received a bonus that depended on their performance in the decision-making task.

General study procedure. Each participant took part in two sessions (average time between sessions: 1.52 days). The first session always involved acquisition of the decision-making task during scanning with magnetoencephalography (data not presented here); the second session involved MRS acquisition. Practical limitations prevented us from acquiring both behavioural and MRS data on the same day. However, note that MRS measures of GABA and glutamate have been reported to be stable over extended periods (weeks to months) and to be non-responsive to current task demands. Therefore, they may reflect relatively stable, trait-like properties[48–51].

Decision-making task. Participants were asked to maximize their rewards in a two-stage decision-making task consisting of 320 trials (Fig. 1a). They first completed 15 practice trials to familiarize themselves with the task before commencing the experiment. Stimulus presentation was controlled by Psychtoolbox 3[52,53] running on Matlab 2012b (The Mathworks Company, Natick, MA). Each trial started with presentation of the two patches (two grey squares). The patch in which the participant currently resided was indicated by a grey frame around the patch. At this stage, participants simply had to indicate by button press (with the index finger of the left or right hand, respectively) whether they wanted to stay in their current patch or switch to the alternative patch. If participants chose to leave their patch, they had to pay a travel cost indicated by the size of a grey bar presented centrally between the two patches. Travel costs were randomly drawn from the set {5, 10, 15, 20 points} and remained constant until a participant chose to leave their patch, at which stage a new cost was selected. The participant's patch choice was highlighted by a frame around the selected patch (400–600 ms, jittered). In trials where participants chose to leave, the rectangular bar signalling travel costs turned red and the costs were subtracted from a progress bar displayed below the patches that indicated the participant's total earnings. Presentation sides (left/right) of the two patches were randomly selected on each trial. Therefore, while participants could decide in advance whether they wanted to stay or leave their patch, this prevented them from preparing the actual motor response before trial onset. Afterwards, the values of the two patches (the reward available in each of them, as indicated by the blue filling) were revealed (1800–2200 ms, jittered). Importantly, the value of the participant's current patch stochastically depleted over time, whereas the alternative patch replenished. Therefore, participants were required to continually accumulate evidence in favour of abandoning their current patch. On each trial, values of the two patches were drawn from Gaussian distributions with non-stationary means and variance = 3.5. The means μ of both patches were set to 50 points initially and then diffused according to a decaying Gaussian random walk on each trial:

$$\mu_{t+1} = \lambda\mu_t + (1-\lambda)\kappa + \varepsilon \quad (4)$$

where λ is the decay rate that was set to 0.96, κ is the decay centre (1 for the chosen patch, 100 for the unchosen patch), and ε is zero-mean Gaussian random noise with a standard deviation = 1.2. Patch values were controlled to fall within an interval of 10–90 points. After the value of the two patches was revealed, participants entered the second stage value-guided choice. The reward available in the chosen patch was allocated to two choice options at a random ratio (ensuring that none of the two options received <10% of the total patch value and excluding a 50% split between options). Furthermore, both options were assigned a probability with which this reward could be obtained, randomly sampled from the set {0.1, 0.2,…, 0.9}. Reward probabilities were independent of each other, such that, in any given trial, either of the two options, both options, or neither of them could be rewarded. Importantly, while higher patch values will, on average, lead to better choice options, this procedure ensures that participants do not know the two choice options when they make their patch decision, thereby explicitly decoupling the patch choice from the value-guided choice. In trials where both the reward probability and magnitude of one option was higher than that of the other option (a 'no-brainer trial'), we randomly flipped the reward magnitudes of the two options in 50% of cases to control for task difficulty. Due to an error in our code, however, this change was only applied to no-brainer trials in which the left option had higher values than the right one. Therefore, no-brainer trials were more likely to be presented on the right side of the screen and the average expected value for the right option was also higher than for the left one ($t_{18558} = -20.915$, $p < 0.001$). Reward magnitudes were indicated by the height of a grey bar and reward probabilities were presented as numbers (in percent) below each bar (Fig. 1a). Reward magnitudes were displayed relative to an outline that corresponded to the overall reward available in the current patch. Participants selected an option by pressing a button with the right or left index finger, respectively. The chosen option was highlighted by a rectangular frame and the outcome of both options was revealed (800–1200 ms, jittered). The bar representing the reward magnitude turned green if an option was rewarded in the current trial, or red otherwise. Even though participants could not benefit from knowing whether the unchosen option would have been rewarded, this procedure has proven useful to remind participants that even low probability options are occasionally rewarded and that it is beneficial to consider each option's reward probability and magnitude. Every time participants were rewarded, the progress bar grew (in proportion to the obtained magnitude) towards a goal indicated by a gold target line to the right of the screen. Outcome presentation was followed by an intertrial interval (1800–2200 ms, jittered) before participants entered the patch decision stage of the next trial. A centrally located fixation cross was present throughout the entire trial. See Supplementary Information for further details on stimulus presentation.

Analysis of behavioural parameters. For the patch-leaving phase, we computed, for each participant and each cost level separately, the median patch value difference (alternative − current patch) at which participants left their current patch. These average patch value differences were compared across cost levels using an RM-ANOVA. Additionally, linear trends along with a constant term were regressed against cost-level-dependent switching behaviour to estimate whether the median patch value difference increased linearly with cost level. In all cost-level-dependent analyses, the data of $N = 28$ participants were analysed since one subject was never presented with the highest cost level (cost levels were randomly assigned after each switch). Furthermore, we defined a patch-leaving advantage by subtracting travel costs from patch value differences at each patch-leaving decision. These values were then averaged across switch trials per participant. We used patch value differences from the previous trial for all analyses pertaining to the patch-leaving phase since the updated patch values are only revealed following the patch choice and hence are informative for the next trial.

For the value-guided choice phase, we computed the percentage of correct responses as the percentage of trials in which subjects chose the option with higher expected value divided by the number of trials with unequal expected value. Regression coefficients were tested against zero with a t test for one sample (two sided). The MEST toolbox was used to provide estimates for effect size measures (η^2 for RM-ANOVA and Cohen's $U3_1$ for one sample to test regression coefficients against 0)[54].

Regression analyses. For each regression analysis, all variables were normalized (z-scored) and a constant term was added to each design matrix. All regressions were performed for each participant separately. Regression coefficients were tested against zero using one-sample t tests (two tailed) and the MEST toolbox was used to provide estimates for effect size measures (Cohen's $U3_1$ for one sample)[54,55]. Additionally, we report the 95% CI for the mean of each distribution of regression coefficients across participants.

To analyse how key value parameters influence value-guided choice, we set up a multiple logistic regression model with choice of the left vs right (0/1) option as dependent variable. Differences and sums of expected values (the product of probabilities and magnitudes for each option) as well as patch value differences and travel costs were entered as independent variables. Here we used patch value differences from the current trial since they are already known by the participant at the time of their value-guided choice. Furthermore, we added the previous trial's value-guided choice, the current trial's patch-leaving choice, a regressor coding for whether the participant had been rewarded in the previous trial, a regressor coding whether each trial was a no-brainer trial (where both probabilities and magnitudes favour the same option), the trial number, and a regressor coding whether each trial was a switch trial or not to the design matrix. We did not run a regression model to explain choices during patch-leaving decisions, since, by design of the task, the maximal patch value difference is always reached when the participants decide to leave their current patch. The regression weights for the effects of value differences would therefore merely reflect a participant's consistency in their patch-leaving behaviour.

To analyse how various task parameters influenced response speed both in the patch-leaving and value-guided choice stage, we used multiple linear regressions with normalized log response time as the dependent variable. As above for the logistic regression, all variables were normalized, a constant was added to the design matrix, and regressions were run for each participant separately. For the patch stage, the design matrix included patch value differences from the previous trial, a binary regressor indicating whether the current trial was a switch trial (one in which the participant left their current patch), the travel cost, a regressor coding for the linear effect of trial number, and two binary regressors indicating whether the side (left/right) of patch presentation had changed with respect to the previous

trial and whether the previous trial's value-guided choice had been rewarded. We used the patch value difference from the previous trial, because the outcome of the (updated) patch values are only revealed after participants' patch choice. For the value-guided choice stage, the design matrix was the same as for the multiple logistic regression model. The only differences between models was that we used a linear link function and absolute value differences between choice options. For the patch-leaving phase (Fig. 2c, d), we have additionally analysed the residuals of dACC E/I balance and behaviour with a robust regression analyses with a bisquare weight function (tuning constant = 4.685).

Behavioural modelling of value-guided decisions. To formally characterize choice behaviour, we fitted several models that combined reward probabilities and magnitudes multiplicatively, additively, or as a combination of both, similar to a recently published approach[23]. All reward magnitudes were rescaled between 1 and 10 before fitting. Since it is known that humans do not weigh magnitudes and probabilities in a statistically optimal way, we considered systematic distortions in the weighting of reward information in our models ($u(m)$ and $w(p)$, for reward magnitudes and probabilities, Eqs. (5) and (6), respectively)[56].

$$w(p_O) = \frac{{p_O}^{\gamma}}{({p_O}^{\gamma} + (1 - p_O)^{\gamma})^{1/\gamma}} \quad (5)$$

where p_O are the objective reward probabilities and γ is a free parameter used to fit subjective reward probabilities. Subjective magnitudes were estimated by:

$$u(m_O) = {m_O}^{\alpha} \quad (6)$$

where m_O is the objective reward magnitude and α is a free parameter used to fit the subjective magnitude. We tested not only models with distorted values but also models where objective reward information is used. In models with objective reward information α and $\gamma = 1$. In all additive models, values were computed according to:

$$V = \omega_m * u(m_O) + \omega_p * w(p_O) \quad (7)$$

where ω_m is a weighting factor for reward magnitudes, and ω_p for reward probabilities. In the multiplicative models, values were computed according to:

$$V = \omega_{mult} * (u(m_O) * w(p_O)) \quad (8)$$

We also estimated a model where we fixed ω_{mult} to the median parameter across all previously recovered parameters, since we found no sufficient recovery for ω_{mult} as well as a better model fit (Supplementary Table 6) with a fixed ω_{mult} parameter.

Finally, we also estimated hybrid models as proposed earlier[23]. Here value is computed according to:

$$V = \omega_{sum} * \left(\left(1 - \left(\frac{\omega_{mult}}{\omega_p + \omega_m + \omega_{mult}}\right)\right) * \left(\left(\frac{\omega_m}{\omega_{p+}\omega_m}\right) * u(m_O) + \left(\frac{\omega_p}{\omega_{p+}\omega_m}\right) * w(p_O)\right) + \left(\frac{\omega_{mult}}{\omega_p + \omega_m + \omega_{mult}}\right) * (u(m_O) * w(p_O))\right) \quad (9)$$

where $\omega_{sum} = \omega_m + \omega_p + \omega_{mult}$. We fitted each of the three different model families (hybrid, additive, and multiplicative) with distorted values for probabilities and magnitudes (SU models), with distorted values only for magnitudes but objective reward probabilities (EU models), and with objective reward magnitudes but subjective reward probabilities (EVPW models) and objective reward probabilities and magnitudes (EV models)[23]. Choice probabilites were modelled with a softmax rule based on option values. Parameters were optimized using custom-written scripts in MATLAB R2019a (The Mathworks Company, Natick, MA) and constrained non-linear optimization using MATLAB's function fmincon was used to minimize the negative log likelihood of the data given the parameters. In order to decrease the probability of fitting local minima, we used 1000 random starting points and report the combination of parameters with the lowest negative log likelihood. The Bayesian information criterion was used to compare between models. For the winning model, we simulated choices for a random set of 500 parameters for each participant and recovered parameters from these artificial data[57]. Correlations between true and recovered parameters across participants can be found in Supplementary Notes 4.

MRS data acquisition. MR data were acquired on a 7 T system (Siemens Healthineers) equipped with a 32-channel array head coil (Nova Medical). First, a high-resolution T1-weighted scan was acquired using an MPRAGE sequence (echo time (TE) = 2.73 ms, repetition time (TR) = 2300 ms, inversion time = 1050 ms, flip angle = 5°, bandwidth = 150 Hz/pixel, acquisition matrix = 320 × 320 × 224, voxel size = 0.8 mm^3 isotropic) aligned with the anterior–posterior commissure (AC–PC). This scan was used not only for the placement of MRS voxels but also for tissue segmentation. We positioned voxels in five regions of interest, including right dlPFC, bilateral primary motor cortices (rM1 and lM1), perigenual anterior cingulate cortex within vmPFC (vmPFC/pgACC), and dACC. The dlPFC voxel was placed on the right hemisphere within the middle frontal gyrus by using the superior frontal sulcus and the inferior frontal sulcus as anatomical landmarks. We positioned the voxel as far dorsally as possible when excluding the calvaria and all extracalvarial structures. The average dlPFC voxel centroid across participants was estimated at MNI $x = 29.79 \pm 0.85$, $y = 37.72 \pm 1.38$, $z = 24.21 \pm 1.51$ (mean ± SEM). Primary motor cortex voxels were placed on the hand knob structures, identified by their omega-like shape on the central sulcus in axial slices. Average M1 voxel centroids in standard space were estimated at MNI $x = -28.97 \pm 0.82$, $y = -18.48 \pm 0.92$, $z = 51.86 \pm 0.59$ and MNI $x = 31.90 \pm 0.71$, $y = -14.76 \pm 1.06$, $z = 49.76 \pm 0.88$ for lM1 and rM1, respectively. The vmPFC voxel was mediolaterally centred on the midline and dorsoventrally on the genu of the corpus callosum, with its posterior boundary just rostral to the genu. The average voxel centroid position across subjects was estimated at MNI $x = -0.17 \pm 0.15$, $y = 41.41 \pm 1.29$, $z = 7.00 \pm 0.44$. The dACC voxel was placed with reference to the corpus callosum, the cingulate as well as surrounding sulci. We used the posterior border of the genu of the corpus callosum perpendicular to AC–PC orientation to centre the voxel (Fig. 1b)[58]. The average centroid voxel position across subjects was MNI $x = -0.07 \pm 0.19$, $y = 24.14 \pm 0.43$, $z = 29.69 \pm 0.47$ (Fig. 1b). For the MRS measurements, region-specific shimming was performed. Voxel sizes were 10 × 20 × 15 mm^3 for the vmPFC voxel and 10 × 25 × 15 mm^3 for all other voxels of interest. Afterwards, MR spectra were acquired using a stimulated echo acquisition mode (STEAM VERSE) sequence (128 averages, TR = 3000 ms, TE = 20 ms, mixing time = 10 ms, data size = 2048, bandwidth = 2800 Hz) from each voxel of interest[58].

MR data analysis. Spectral data were analysed using the LCModel[59]. Only metabolite measurements with a Cramér–Rao lower bound <20%, full-width half-maximum <25 Hz, and signal-to-noise ratio >8 were included. We analysed the quality of each voxel measurement using LCModel immediately after acquisition of the voxel. If one voxel did not meet the quality criteria, we repeated the acquisition of this specific voxel. We had to repeat measurement of 1 of the 5 voxels in 13 of our 29 subjects to obtain valid measurements for all 5 voxels of interest. SPM 12 (Wellcome Trust Centre for Neuroimaging, London, UK) was used to segment participants' T1-weighted anatomical images into grey matter (GM), white matter (WM), cerebrospinal fluid, soft tissue, and air/background. Each voxel's GABA and glutamate concentrations were corrected for relative GM concentrations[60] by dividing their absolute concentrations by relative GM, based on the assumption that GABA and glutamate are predominantly present in GM. As SPM 12 provides tissue probability maps, we summed across probabilities for GM for each voxel in the mask and divided by the total number of voxels within each mask to approximate relative GM. Total creatine concentrations (creatine + phosphocreatine) were normalized by the relative amount of GM and WM within each voxel ((GM+WM)/number of $voxels_{mask}$) based on the assumption that creatine is predominantly present in GM and WM. All GABA and glutamate concentrations we report are normalized by total creatine concentrations. We defined E/I balance as the ratio of (normalized) glutamate to GABA levels. Voxel masks were then interpolated to individual MRI volumes with FieldTrip[61]. To estimate average voxel centroid positions, we normalized individual volume data. More specifically, data were registered to MNI space by using tissue probability maps (TPM.nii template from SPM 12). Estimated average centroid positions were extracted from each mask in MNI space.

Behavioural parameters and E/I balance. For each behavioural analysis, we obtained a measure describing the individual influence of key value parameters on behaviour (decision variable). To relate decision variables to cortical E/I balance, we used multiple linear regression. In order to limit the number of comparisons, we used the following hierarchy of testing: First, we only tested those decision variables for their relationship with cortical neurochemistry (a) for which we had an a priori hypothesis and (b) that had a significant effect on behaviour. Second, using a GLM, we first projected E/I balance (ratio between glutamate and GABA) from all five regions of interest against main behavioural parameters from the patch-leaving and value-guided choice phase. All of these analyses were performed exclusively using the design matrix containing E/I balance from all five regions. Third, if, and only if, this GLM yielded a significant effect for one brain region, we followed this up by asking whether this effect of E/I balance was contributed to by glutamate, GABA, or both within that specific region. To this end, we then computed partial correlations, regressing out the effects of all other factors than the one currently of interest (see below). Fourth, when there was an effect of E/I balance on a decision variable of interest, we followed this up by analysing in more detail what aspects of a decision variable were related to E/I balance. Because of our clearly defined a priori hypotheses regarding the role of dACC vs vmPFC, we did not apply correction for multiple comparison based on the number of brain regions tested in our regression models.

We analysed the following decision variables: for the patch-leaving phase, we considered the patch-leaving advantage as the key measure of interest that describes how participants balance the expected advantage of leaving against the travel cost. For the value-guided choice phase, overall decision performance, measured as percentage of correct choices (choices of the option with higher expected value) was the primary measure of interest. The latter relationship was investigated in further detail by testing to what extent the relationship between E/I balance and % correct was driven by the individual-specific distortions of reward information as captured by the model-derived parameters α and γ. Finally, we assessed how E/I balance was related to the effect of relevant reward information

on reaction times. Here, for the patch-leaving phase, the effect of cost on response speed was our key parameter of interest, whereas for the value-guide choice phase, we focussed on the effect of value difference on reaction times. As a control, we also included overall response times, independent of any task parameters (Supplementary Tables 4 and 5 for an overview).

All of the E/I analyses were performed exclusively using the design matrix containing balances from all five regions. The ratios of glutamate to GABA in all voxels of interest were entered as regressors in a design matrix (along with a constant term) to predict the contributions of E/I balance onto each decision variable (one linear model per decision variable, all variables normalized). Only if a significant influence of E/I balance in a specific region was identified, we computed partial correlations. These partial correlations were first computed for E/I balance in the target region (the one showing a significant effect in the main GLM) by orthogonalizing (removing the effect of all other E/I balances) both from the decision variable and from E/I balance in the target region. These partial correlations (Pearson correlation between the residuals of decision variable and E/I balance) are what is shown in Figs. 2c, d and 3b–d. To further investigate whether a main effect of E/I balance in one region was driven by GABA or glutamate (or both), we computed further partial correlations on the orthogonal contribution of GABA and glutamate. To do so, we orthogonalized both the decision variable and the neurotransmitter of interest (GABA or glutamate, respectively) in the target area with respect to both the respective other neurotransmitter in the same region and both GABA and glutamate in all other voxels. As an example, if the GLM detected a main effect of E/I in dACC on patch-leaving advantage, and we wanted to compute the orthogonal contribution of dACC GABA to this effect, then we removed the effect of dACC glutamate and the effects of both GABA and glutamate in the other four voxels from both the patch-leaving advantage and from dACC GABA. Again, we computed Pearson correlations between the residuals of the decision variable and the residuals of GABA or glutamate, respectively. In the analysis of value-guided decisions, we also controlled for the amount of no-brainer trials (mean ± SEM: 38.52% ± 0.01) as a predictor variable of no interest. We report t and p values for each significant regression coefficient ($p < 0.05$), testing for differences from zero, as well as r and p values for the subsequent partial correlations (if applicable), both with their respective 95% CIs.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw MRS data that support the findings of this study are available from the corresponding author upon reasonable request. Raw MRS data are not publicly available due to them containing information that could compromise research participant privacy/consent. The behavioural data and tables summarizing all MRS results are available under www.github.com/luckyluc25/ei_exp. A reporting summary for this article is available as a Supplementary Information file. Source data are provided with this paper.

Code availability

Custom written code used to analyse the behavioural data of the current study is available under www.github.com/luckyluc25/ei_exp.

Received: 16 December 2019; Accepted: 11 December 2020;
Published online: 10 February 2021

References

1. Wittmann, M. K. et al. Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat. Commun.* **7**, 12327 (2016).
2. Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011).
3. Kolling, N., Behrens, T. E. J., Mars, R. B. & Rushworth, M. F. S. Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
4. Kvitsiani, D. et al. Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* **498**, 363–366 (2013).
5. Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R. & Eskandar, E. N. Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat. Neurosci.* **7**, 1370–1375 (2004).
6. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335–1338 (1998).
7. Fouragnan, E. F. et al. The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. *Nat. Neurosci.* **22**, 797–808 (2019).
8. Jocham, G. et al. Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. *NeuroImage* **100**, 498–506 (2014).
9. Jocham, G., Hunt, L. T., Near, J. & Behrens, T. E. J. A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat. Neurosci.* **15**, 960–961 (2012).
10. Papageorgiou, G. K. et al. Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. *Nat. Commun.* **8**, 1886 (2017).
11. Boorman, E. D., Rushworth, M. F. & Behrens, T. E. Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *J. Neurosci.* **33**, 2242–2253 (2013).
12. Wunderlich, K., Rangel, A. & O' Doherty, J. P. Neural computations underlying action-based decision making in the human brain. *Proc. Natl Acad. Sci. USA* **106**, 17199–17204 (2009).
13. Bechara, A., Tranel, D. & Damasio, H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* **123**, 2189–2202 (2000).
14. Boorman, E. D., Behrens, T. E. J., Woolrich, M. W. & Rushworth, M. F. S. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* **62**, 733–743 (2009).
15. Strait, C. E., Blanchard, T. C. & Hayden, B. Y. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* **82**, 1357–1366 (2014).
16. FitzGerald, T. H. B., Seymour, B. & Dolan, R. J. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J. Neurosci.* **29**, 8388–8395 (2009).
17. Hunt, L. T. et al. Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* **15**, 470–476 (2012).
18. Kolling, N. et al. Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.* **19**, 1280 (2016).
19. Hämmerer, D., Bonaiuto, J., Klein-Flügge, M., Bikson, M. & Bestmann, S. Selective alteration of human value decisions with medial frontal tDCS is predicted by changes in attractor dynamics. *Sci. Rep.* **6**, 25160 (2016).
20. Standage, D. & Paré, M. Persistent storage capability impairs decision making in a biophysical network model. *Neural Netw.* **24**, 1062–1073 (2011).
21. Curtis, C. E. & D'Esposito, M. Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* **7**, 415–423 (2003).
22. Barbey, A. K., Koenigs, M. & Grafman, J. Dorsolateral prefrontal contributions to human working memory. *Cortex* **49**, 1195–1205 (2013).
23. Farashahi, S., Donahue, C. H., Hayden, B. Y., Lee, D. & Soltani, A. Flexible combination of reward information across primates. *Nat. Hum. Behav.* **3**, 1215–1224 (2019).
24. Hayden, B. Y., Heilbronner, S. R., Pearson, J. M. & Platt, M. L. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.* **31**, 4178–4187 (2011).
25. Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J. & Rushworth, M. F. S. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
26. Fischer, A. G. & Ullsperger, M. When is the time for a change? Decomposing dynamic learning rates. *Neuron* **84**, 662–664 (2014).
27. Sheth, S. A. et al. Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* **488**, 218–221 (2012).
28. Wessel, J. R., Danielmeier, C., Morton, J. B. & Ullsperger, M. Surprise and error: common neuronal architecture for the processing of errors and novelty. *J. Neurosci.* **32**, 7528–7537 (2012).
29. Karlsson, M. P., Tervo, D. G. R. & Karpova, A. Y. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* **338**, 135–139 (2012).
30. Quilodran, R., Rothé, M. & Procyk, E. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* **57**, 314–325 (2008).
31. Rothé, M., Quilodran, R., Sallet, J. & Procyk, E. Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. *J. Neurosci.* **31**, 11110–11117 (2011).
32. Fuchs, E. C. et al. Recruitment of parvalbumin-positive interneurons determines hippocampal function and associated behavior. *Neuron* **53**, 591–604 (2007).
33. Wang, X.-J. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* **90**, 1195–1268 (2010).
34. Buzsáki, G. & Wang, X.-J. Mechanisms of gamma oscillations. *Annu. Rev. Neurosci.* **35**, 203–225 (2012).
35. Kathleen, K. A. Cho et al. Cross-hemispheric gamma synchrony between prefrontal parvalbumin interneurons supports behavioral adaptation during rule shift learning. *Nat. Neurosci.* **23**, 892–902 (2020).
36. Vogt, B. A. In *Cingulate Neurobiology and Disease* (ed. Vogt, B. A.) 65–93 (Oxford University Press, Oxford, 2009).
37. Vogt, B. A. Midcingulate cortex: Structure, connections, homologies, functions and diseases. *J. Chem. Neuroanat.* **74**, 28–46 (2016).

38. Donoso, M., Collins, A. G. E. & Koechlin, E. Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science* **344**, 1481–1486 (2014).
39. Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).
40. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328 (2006).
41. Domenech, P. & Koechlin, E. Executive control and decision-making in the prefrontal cortex. *Curr. Opin. Behav. Sci.* **1**, 101–106 (2015).
42. Scholl, J. et al. Excitation and inhibition in anterior cingulate predict use of past experiences. *eLife* **6**, e20365 (2017).
43. Kehrer, C., Maziashvili, N., Dugladze, T. & Gloveli, T. Altered excitatory-inhibitory balance in the NMDA-hypofunction model of schizophrenia. *Front. Mol. Neurosci.* **1**, 6 (2008).
44. Godfrey, K. E. M., Gardner, A. C., Kwon, S., Chea, W. & Muthukumaraswamy, S. D. Differences in excitatory and inhibitory neurotransmitter levels between depressed patients and healthy controls: a systematic review and meta-analysis. *J. Psychiatr. Res.* **105**, 33–44 (2018).
45. Collins, A. G. E., Brown, J. K., Gold, J. M., Waltz, J. A. & Frank, M. J. Working memory contributions to reinforcement learning impairments in schizophrenia. *J. Neurosci.* **34**, 13747–13756 (2014).
46. Admon, R. & Pizzagalli, D. A. Dysfunctional reward processing in depression. *Curr. Opin. Psychol.* **4**, 114–118 (2015).
47. Oldfield, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
48. Greenhouse, I., King, M., Noah, S., Maddock, R. J. & Ivry, R. B. Individual differences in resting corticospinal excitability are correlated with reaction time and GABA content in motor cortex. *J. Neurosci.* **37**, 2686–2696 (2017).
49. Houtepen, L. C. et al. Acute stress effects on GABA and glutamate levels in the prefrontal cortex: a 7T 1H magnetic resonance spectroscopy study. *Neuroimage. Clin.* **14**, 195–200 (2017).
50. Near, J., Ho, Y.-C. L., Sandberg, K., Kumaragamage, C. & Blicher, J. U. Long-term reproducibility of GABA magnetic resonance spectroscopy. *Neuroimage* **99**, 191–196 (2014).
51. Talsma, L., van Loon, A., Scholte, H. S. & Slagter, H. A. State or trait? MRS-measured GABA and Glutamate concentrations are not modulated by task demand and do not robustly predict task performance. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/543140v1 (2019).
52. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
53. Kleiner, M. et al. What's new in Psychtoolbox-3. *Perception* **36**, 1–16 (2007).
54. Hentschke, H. & Stüttgen, M. C. Computation of measures of effect size for neuroscience data sets. *Eur. J. Neurosci.* **34**, 1887–1894 (2011).
55. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (Academic Press, New York, 1988).
56. Kahneman, D. & Tversky, A. Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 363–391 (1979).
57. Wilson, R. C. & Collins, A. G. Ten simple rules for the computational modeling of behavioral data. *eLife* **8**, e49547 (2019).
58. Dou, W. et al. Systematic regional variations of GABA, glutamine, and glutamate concentrations follow receptor fingerprints of human cingulate cortex. *J. Neurosci.* **33**, 12698–12704 (2013).
59. Provencher, S. W. Automatic quantitation of localized in vivo 1H spectra with LCModel. *NMR Biomed.* **14**, 260–264 (2001).
60. Bachtiar, V., Near, J., Johansen-Berg, H. & Stagg, C. J. Modulation of GABA and resting state functional connectivity by transcranial direct current stimulation. *eLife* **4**, e08789 (2015).
61. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).

Acknowledgements

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (JO-787/6-1) to G.J. and by the federal state of Saxony-Anhalt and the European Regional Development Fund (ERDF 2014-2020), Vorhaben: Center for Behavioral Brain Sciences (CBBS), FKZ: ZS/2016/04/78113. The authors thank Renate Blobel-Lüer for her help with MRS recordings and Alex Waite for technical support. Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf.

Author contributions

T.O.J.G. and G.J. designed the research. T.O.J.G. and O.S. recorded the data. L.F.K., T.O.J.G. and G.J. analysed the data. L.F.K. and G.J. wrote the manuscript. All authors discussed the results at all stages of the experiment and read and edited versions and approved the final version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20875-w.

Correspondence and requests for materials should be addressed to L.F.K.

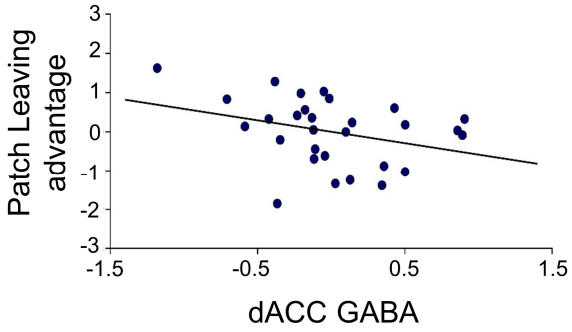
Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

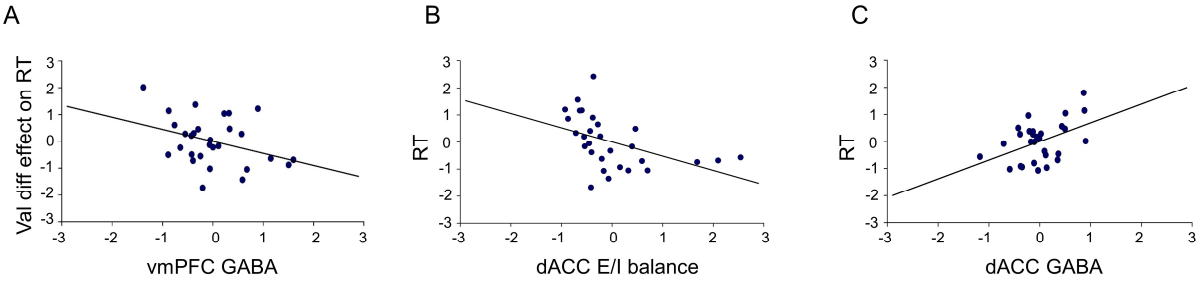
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions

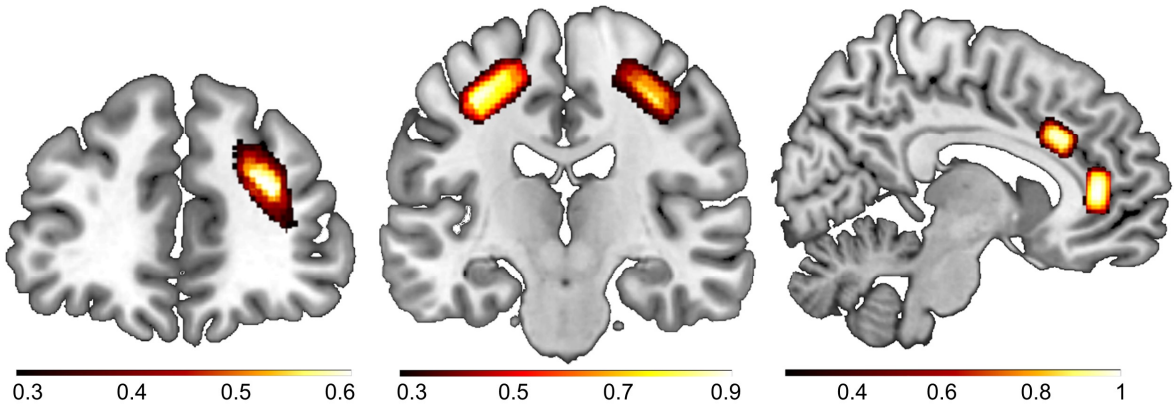
Supplementary Information



Supplementary Figure S1. Relationship between dACC GABA and patch-leaving. GABA contribution to the effect of E/I balance on patch-leaving shown in main figure 2C. Higher dACC concentrations of GABA are, by trend, associated with earlier patch leaving (lower average patch leaving advantages) (Pearson correlation on residuals (compare main text and methods): r = -0.323, p = 0.087, CI_{95} = [-0.617 – 0.049]; N = 29). Source data are provided as a Source Data file.



Supplementary Figure S2. Additional analysis for the relationship between value-guided choice and cortical neurochemistry. A) GABA contribution to the effect of E/I balance shown in main figure 3D. Participants' responses slowed down on difficult trials (trials with low value difference). This effect was related to vmPFC GABA concentrations (Pearson correlation on residuals: r = -0.357, p = 0.057, CI_{95} = [-0.640 – 0.011]). Val diff = value difference. B) dACC E/I balance relates to overall response speed during value guided choice (Pearson correlation on residuals: r = -0.459, p = 0.012, CI_{95} = [-0.707 – -0.111]). RT = Reaction Time. C) dACC GABA relates to overall response speed during value guided choice (Pearson correlation on residuals: r = 0.452, p = 0.014, CI_{95} = [0.102 – 0.702]). RT = Reaction Time. N = 29 in all figures. Source data are provided as a Source Data file.



Supplementary Figure S3. Overlay of voxel placements across all participants: Average locations of all regions of interest. Brighter colors indicate a greater overlap across participants. Left: Average placement of dlPFC voxel. Middle: Average location of M1 voxels. Right: Average location of vmPFC and dACC MRS voxel. *N* = 29.

Supplementary Notes

1) Reaction times in the patch leaving phase are not influenced by trial-wise patch leaving advantages

We reran a regression model to analyze reaction times during patch-leaving decisions. Here, we included trial-wise patch leaving advantages instead of using costs and patch value differences as separate regressors. This analysis revealed no significant influence of PLA (t_{28} = -0.118, p = 0.907, CI_{95} = [-0.042 – 0.038], $U3_1$ = 0.483). We again find a significant influence of whether each trial was a switch trial or not (t_{28} = 3.776, p = 0.001, CI_{95} = [0.042 – 0.141], $U3_1$ = 0.310), of trial number (t_{28} = -8.039, p < 0.001, CI_{95} = [-0.310 – -0.184], $U3_1$ = 0.897), of whether the presentation side of patch values changed with respect to the last trial (t_{28} = 3.844, p = 0.001, CI_{95} = [0.025 – 0.083], $U3_1$ = 0.276) and of whether the value-guided choice in the last trial had been rewarded (t_{28} = 3.689, p = 0.001, CI_{95} = [0.020 – 0.071], $U3_1$ = 0.207). The finding that costs did influence reaction times whereas neither PLA nor patch value differences (analysis in main text) had an effect is likely related to the structure of the task. The cost of leaving is displayed on screen at the outset of each trial, whereas patch values have to be held in memory from the outcome of the last trial's patch choice.

2) vmPFC E/I balance relates to weighting of reward information during value-guided choice

In the main text, we report a negative relationship of vmPFC E/I balance and choice accuracy. To further investigate this, we wanted to quantify the degree to which participants' choices were guided by the options' expected values. To this end, we used a logistic regression which is already reported in the main text. Similar to our relationship between E/I balance and choice accuracy reported in the main manuscript, we found that E/I balance in vmPFC was related to the degree to which participants' choices were governed by expected values. There was a significant negative relationship between vmPFC E/I balance and the effect of value difference on choice (t_{22} = -2.593, p = 0.017, CI_{95} = [-0.959 – 0.107]; r = -0.484, p = 0.008, CI_{95}= [-0.722 – -0.143]). Thus, mirroring the effects on % correct choices, participants with higher levels of GABA relative to glutamate in vmPFC based their choices more strongly on the options' expected values. We further detailed this effect by re-running the regression with separate regressors for the differences in reward probabilities and magnitudes instead of one coding for difference in expected value. Participants used both reward probabilities (t_{28} = 11.708, p < 0.001, CI_{95} = [3.560 – 5.070], $U3_1$ = 0) and magnitudes (t_{28} = 14.431, p <0.001, CI_{95} = [2.806 – 3.734], $U3_1$ = 0) to guide their choices. Choices of participants with increased vmPFC E/I balance were more strongly influenced by reward probabilities compared to magnitudes (t_{22} = 2.736, p = 0.012, CI_{95} = [0.134 – 0.971]; r = 0.504, p = 0.005, CI_{95}= [0.169 – 0.735]). This pattern of results matches our findings reported in the main text and again indicates that participants with a greater E/I balance in vmPFC based their decisions less on objective differences in expected values. This effect is potentially mediated by a stronger reliance on reward probabilities than magnitudes.

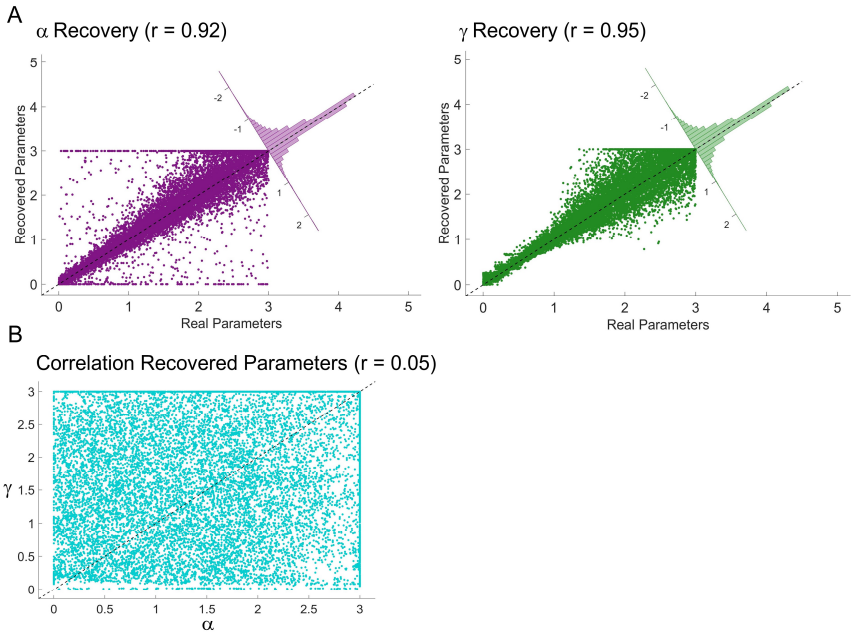
3) Simultaneous regression of all behavioural parameters of interest against E/I balance in dACC and vmPFC

Some of our dependent variables may be correlated with each other across participants. This is expected since some of the tests investigate parameters that we assume to be driven by a

shared underlying mechanism[1]. For instance, consider the case for % correct choices on one hand and the effect of value difference on RT on the other. As can be seen from Supplementary Table 2, there is a negative correlation between these two variables, indicating that the (negative) effect of value difference on RT is most pronounced in participants with high percentage of correct choices. This, however, is exactly what would be mechanistically predicted from models using competition via mutual inhibition: Slowing the decision in the face of a lot of noise (a difficult trial with low value difference) allows for the choice to be dominated by the available evidence, while averaging out (neural) noise over time. To assess the orthogonal contributions of all the different behavioural parameters across both the patch-leaving and value-guided choice phase, we therefore included all of the parameters of interest from both phases (Supplementary Table 1 and 2) into one single regression model and now used either dACC or vmPFC E/I balance as the dependent variable. We still find a significant effect of patch leaving advantage on dACC E/I balance (t_{20} = 3.013, p = 0.007, CI_{95} = [0.175 – 0.961]) but no significant effect of any other variable of interest (all p > 0.119). When regressing the same design matrix against vmPFC E/I balance, we find no significant effect of any behavioural parameter (all p > 0.151).

4) Model Validation: Simulate and Recover

To validate our model fitting routines [2,3], we generated and recovered data for the model with the lowest BIC (prospect model with α and γ as free parameters). We generated 500 artificial data sets by randomly selecting α and γ parameters in the range between 0 and 3. We then recovered these parameters from the artificial data with the same procedure as used for our real participants. We used 1000 random starting points to find the combination of free parameters yielding the minimal negative log likelihood across iterations. All fittings were done for each participant separately. The distance and correlations between recovered parameters and the ground truth parameters across subjects were estimated (Supplementary Figure S4) as well as the correlations between recovered parameters.



Supplementary Figure S4. Overview of Simulated and Recovered Model Parameters: A) Correlation between true and recovered parameters and a histogram of the difference between true and recovered parameters for our winning model (see methods for model details). B) Correlations between recovered parameters. Source data are provided as a Source Data file.

Supplementary Table 1: **Correlations between behavioral variables of interest during the patch–leaving phase.** Correlations > 0.4 are marked in red. PLA = Patch Leaving Advantage, RT = Reaction Time, Costs on RT = β regression weight of costs on reaction times. Source data are provided as a Source Data file.

	PLA	RT	Cost effect on RT
PLA	1	-0.08	0.13
RT	-0.08	1	-0.35
Cost effect on RT	0.13	-0.35	1

Supplementary Table 2: **Correlations between behavioral variables of interest during the value-guided choice phase.** Correlations > 0.4 are marked in red. RT = Reaction Time, val diff RT = β regression weight of value difference on RT, α = parameter transforming objective to subjective magnitudes, γ = parameter transforming objective probabilities to subjective probabilities. Source data are provided as a Source Data file.

	% correct	RT	val diff RT	α	γ
% correct	1	0.42	-0.65	0.87	-0.64
RT	0.42	1	-0.09	0.49	-0.03
val diff RT	-0.65	-0.09	1	-0.63	0.36
α	0.87	0.49	-0.63	1	-0.38
γ	-0.64	-0.03	0.36	-0.38	1

Supplementary Table 3: **Correlation of E/I balance across cortical areas.** Overview of correlations between E/I balances between all regions of interest. Correlations > 0.4 are marked in red. dlPFC = dorsolateral prefrontal cortex, vmPFC = ventromedial prefrontal cortex, dACC = dorsal anterior cingulate cortex. Source data are provided as a Source Data file.

	dlPFC	M1 left	M1 right	vmPFC	dACC
dlPFC	1	0.15	0.22	-0.09	0.11
M1 left	0.15	1	0.18	-0.09	-0.22
M1 right	0.22	0.18	1	-0.19	0.04
vmPFC	-0.09	-0.09	-0.19	1	-0.38
dACC	0.11	-0.22	0.04	-0.38	1

Supplementary Table 4: Overview of neurochemical effects in all regression models. Neurochemical effects for all regression models conducted to analyze patch-leaving behavior. All regression models include all E/I balances (and an intercept) as dependent variables and the behavioural variable of interest as independent variable. *T*- and *p*- values indicate the test statistic for each coefficient in the regression model to test the null hypothesis that the coefficient is zero. Significant effects (p <= 0.05) are marked in red. Source data are provided as a Source Data file.

	dlPFC E/I [t-stat, p-value]	M1 left E/I [t-stat, p-value]	M1 right E/I [t-stat, p-value]	vmPFC [t-stat, p-value]	dACC [t-stat, p-value]
PLA	-0.39, 0.70	0.82, 0.42	1.32, 0.20	-0.20, 0.85	2.64, 0.02
RT	0.15, 0.88	0.87, 0.39	0.00,1	-0.16, 0.87	-1.33, 0.20
Effects of Cost on RT	-0.53,0.60	1.00, 0.33	0.57, 0.57	0.46, 0.65	2.19, 0.04

Supplementary Table 5: Overview of neurochemical effects in all regression models: Neurochemical effects for all regression models conducted to analyze value-guided choice behavior. All regression models include all E/I balances (and an intercept) as dependent variables and the behavioural variable of interest as independent variable. *T*- and *p*- values indicate the test statistic for each coefficient in the regression model to test the null hypothesis that the coefficient is zero. Significant effects (p <= 0.05) are marked in red. Source data are provided as a Source Data file.

	dlPFC E/I [t-stat, p-value]	M1 left E/I [t-stat, p-value]	M1 right E/I [t-stat, p-value]	vmPFC [t-stat, p-value]	dACC [t-stat, p-value]
% correct	-0.44, 0.67	-0.62, 0.54	-0.53, 0.60	-2.44, 0.02	-0.85, 0.41
RT	-0.24, 0.81	0.22, 0.83	-0.10, 0.92	-1.58, 0.13	-2.42, 0.02
Effect of Val Diff on RT	-0.41, 0.68	1.68, 0.11	0.53, 0.60	2.88, 0.01	1.28, 0.21
α	-0.29, 0.77	-0.55, 0.59	-0.47, 0.64	-2.41, 0.02	-1.14, 0.27
γ	0.49, 0.63	1.00, 0.33	0.14, 0.89	2.14, 0.04	0.02, 0.98

Supplementary Table 6: Parameter values and model fits for behavioural models. Overview of all recovered model parameters as well as their model fit (BIC = Bayesian Information Criterion). The model with the lowest BIC is bold. See Supplementary Analysis 4 for model validation. EV models assume no distortions in value weighting. EU models assume distortions in reward magnitude weighting, EVPW in reward probabilities and SU in both reward probabilities and magnitudes. Additive (Add) models assume additive value integration, multiplicative (multi) models assume multiplicative value integration, and hybrid models a combination of both. Source data are provided as a Source Data file.

Model	$\frac{\omega_{mult}}{\omega_m+\omega_p+\omega_{mult}}$	$\frac{\omega_m}{\omega_m+\omega_p}$	$\frac{\omega_p}{\omega_m+\omega_p}$	ω_{mult}	α	γ	BIC
EV							
Add		0.08 ± 0.01	0.92 ± 0.01				182.78 ± 6.47
Multi				1.93 ± 0.18			223.44 ± 11.75
Hybrid	0.27 ± 0.05	0.17 ± 0.06	0.83 ± 0.06				166.45 ± 7.16
EU							
Add		0.66 ± 0.06	0.34 ± 0.06		0.22 ± 0.06		174.23 ± 6.92
Multi				5.54 ± 0.73	0.71 ± 0.07		164.40 ± 7.92
Hybrid	0.54 ± 0.06	0.65 ± 0.08	0.35 ± 0.08		0.59 ± 0.08		165.61 ± 7.51
Fixω_{mult} Multi					0.69 ± 0.06		172.06 ± 7.91
EVPW							
Add		0.08 ± 0.01	0.92 ± 0.01			1.03 ± 0.04	187.46 ± 6.46
Multi				3.11 ± 0.27		1.79 ± 0.15	180.78 ± 8.49
Hybrid	0.38 ± 0.07	0.23 ± 0.06	0.77 ± 0.06			1.23 ± 0.09	168.51 ± 7.21
Fixω_{mult} Multi						1.84 ± 0.15	182.23 ± 8.41
SU							
Adi		0.65 ± 0.06	0.35 ± 0.06		0.23 ± 0.06	1.03 ± 0.04	178.88 ± 6.89
Multi				10.04 ± 1.80	0.59 ± 0.06	0.83 ± 0.07	162.96 ± 7.78
Hybrid	0.58 ± 0.06	0.64 ± 0.08	0.36 ± 0.08		0.65 ± 0.08	1.14 ± 0.07	168.99 ± 7.50
Fixω_{mult} Multi					0.58 ± 0.05	0.88 ± 0.05	**161.99 ± 7.06**

Note: All values are mean values across participants ± standard error of the mean. *N* = 29.

Supplementary Note: Exploratory Findings - Relating current findings to own previous work

In an earlier study, we have already reported relationships between vmPFC E/I balance and optimal choice behaviour[4]. In particular, we had reported that high levels of GABA, and low levels of glutamate, respectively, were related to participants' performance on difficult trials (those with low value difference), as measured by the softmax inverse temperature. This finding is exactly predicted by mechanistic models based on competition by mutual inhibition[4]. However, a recent study found that choices were more strongly guided by multiplicative as opposed to additive value computation after administration of the NMDA receptor agonist d-cycloserine to healthy volunteers[5]. Combining values multiplicatively is considered more optimal whereas an additive value integration is potentially less complex. In our own previous data, we found an effect of vmPFC E/I balance on softmax inverse temperature[4]. In this work, however, we had not compared between different models featuring multiplicative versus additive value construction, or a mixture of both. We have therefore reanalyzed our previous data with the same set of models as used in the current study. All magnitudes have been rescaled between 1 and 10 prior to model fitting. We find that a hybrid model with no distortions in value weighting fits the data best. Since the EV hybrid model fits our previous data best, we assessed the relationship between this model's free parameters and E/I balance. One participant had to be excluded because GABA and glutamate could not be successfully detected[4]. We don't find any significant relationship between vmPFC GABA (t_{21} = -1.559, p = 0.134) or glutamate (t_{21} = 0.421, p = 0.678) on the reliance of integrative versus additive value integration. When we compared the reliance on multiplicative versus additive value updating in our current data set (EV hybrid), we find, as expected, a greater reliance on multiplicative value integration with a lower vmPFC E/I balance (t_{22} = -2.423, p = 0.024) as well as a greater reliance on magnitude compared to probability values within the additive module (t_{22} = -2.711, p = 0.013). Neither the previous nor this study was primarily designed to study whether E/I balance measured with MRS relates to a multiplicative or additive value integration. It would be interesting for further

studies to analyze this question with a set of options where choices would explicitly dissociate multiplicative from additive value integration.

As reported in the main text, for our present study, we find that a multiplicative SU model fits the data best. However, we did not obtain sufficient model recovery for the choice stochasticity parameter and therefore decided to fix it at the median recovered vale. There are a number of possible reasons for this. First, in the 2012 data, the trials' combination of reward attributes had been specifically optimized (offline) for the value-guided choice task to allow a certain level of difficulty, to control for correlation between chosen and unchosen value, and to incorporate a certain range of no-brainer trials. In contrast, in the current task, reward magnitudes are generated from the chosen patch, a random fraction of which is allocated to the two patches. Small magnitude differences are therefore less likely to occur, which potentially prevents a reliable estimation of the choice stochasticity parameter. Secondly, in the current task the distortion of reward magnitudes becomes more important since magnitudes can potentially cover a wider range of values that depends on the current patch value, as opposed to a fixed minimum and maximum in the 2012 study.

Supplementary Table 7: Overview of model fitting to the data presented in Jocham et al. (2012)[4]. EV models assume no distortions in value weighting. EU models in reward magnitude weighting, EVPW in reward probabilities and SU in reward probabilities and magnitudes. Add models assume additive value integration, multi models multiplicative value integration and hybrid models a combination of both. Source data are provided as a Source Data file.

Model	$\frac{\omega_{mult}}{\omega_m+\omega_p+\omega_{mult}}$	$\frac{\omega_m}{\omega_m+\omega_p}$	$\frac{\omega_p}{\omega_m+\omega_p}$	ω_{mult}	α	γ	BIC
EV							
Add		0.07 ± 0.00	0.93 ± 0.00				237.05 ± 8.16
Multi				1.80 ± 0.17			257.36 ± 15.11
Hybrid	0.19 ± 0.04	0.10 ± 0.04	0.90 ± 0.04				**224.69 ± 9.56**
EU							
Add		0.44 ± 0.06	0.56 ± 0.06		0.40 ± 0.06		229.43 ± 8.50
Multi				3.70 ± 0.47	0.77 ± 0.06		231.97 ± 10.46
Hybrid	0.27 ± 0.05	0.37 ± 0.08	0.63 ± 0.08		0.69 ± 0.06		227.20 ± 9.43
EVPW							
Add		0.07 ± 0.00	0.93 ± 0.00			1.02 ± 0.04	241.89 ± 8.02
Multi				2.07 ± 0.15		1.47 ± 0.16	243.44 ± 11.62
Hybrid	0.22 ± 0.04	0.06 ± 0.01	0.94 ± 0.01			1.09 ± 0.08	228.48 ± 9.43
SU							
Add		0.44 ± 0.06	0.56 ± 0.06		0.40 ± 0.06	1.00 ± 0.04	234.20 ± 8.30
Multi				8.74 ± 1.06	0.55 ± 0.04	0.67 ± 0.03	228.09 ± 9.38
Hybrid	0.29 ± 0.06	0.35 ± 0.08	0.65 ± 0.08		0.70 ± 0.05	1.03 ± 0.06	231.60 ± 9.29

Note: All values are mean values across participants ± standard error of the mean. $N = 25$.

Supplementary Note: Exploratory Analysis - Drift Diffusion Modelling of Choice Data

To obtain a formal characterization of the process of evidence accumulation across trials, we fitted a hierarchical drift diffusion model (DDM)[6]. In brief, DDM assume that choices between two alternatives depend upon accumulation of noisy evidence until a decision threshold is reached. The model thereby not only predicts choice probabilities but also response time (RT) distributions. The predicted choice probabilities and RTs critically depend upon three free parameters. First, the decision boundary *a* determines how much evidence needs to be accumulated. Second, the drift rate *v* captures the speed at which the evidence accumulation process approaches either boundary[6]. Third, RT is assumed not to solely depend on the choice process itself, but also on other non-decisional processes like stimulus perception and the execution of a motor response, which is reflected in the non-decision time (*ndT*)[7]. To account for across-trial variations[8], we also tested the effects of variability in *ndt* (*st*) and *v (sv*).

We used the Bayesian hierarchical drift diffusion modeling toolbox with default priors[6] in Python 2 to infer latent variables underlying response time distributions of patch leaving trials and correct vs. incorrect value - guided choices. The estimation of individual parameters is hierarchical since they are not assumed to be independent of one another but drawn from an underlying group distribution[6]. We estimated drift rate, boundary separation and non-decision time individually, but across-trial variability in drift rate and non-decision time on a group level[9,10]. For model comparisons, we included additional effects of bias towards one decision boundary (*z*) and variations in bias (*sz*) in patch leaving trials as well as linear regression models assessing the effect of reward information onto free DDM parameters. For regressions, all continuous variables were z-scored per participant before estimating regression coefficients. Cost levels were z-scored on a group level. Since our task does not involve a maximum response time, we excluded all trials with response times below 0.3 or above 4 seconds before model fitting. Additionally, we specified 5 % of responses to be contaminants. The toolbox uses Markov-Chain Monte Carlo sampling for a Bayesian approximation of the posterior distribution of each model parameter. For every model, we ran

thirty separate Markov chains and report parameter estimates and posterior distributions of a concatenated model across all chains[9]. We generated 5000 samples for every chain and discarded one half of all samples as burn-in[9]. Every third sample was discarded for thinning, thereby reducing autocorrelations in the chains. To assess model convergence, we inspected the sampled posterior traces, their autocorrelation and the Gelman-Rubin $\hat{R}$ statistics, which compares between and within chain variance[6,11]. $\hat{R}$ for a group level parameter with a distance of > 0.02 from one were defined as non-converged models. To compare between models, we used the Deviance Information Criterion (DIC) where a lower DIC points towards a better fit. Based on previous findings[12–16], we predicted a relationship between E/I balance and the drift rate *v* and decision boundary *a*.

For the patch leaving phase we find that a drift diffusion model with *a, v, ndt*, *z*, *st* and *sz* fit the data best. When we assess the effects for individual model parameters for their relationship with E/I balance, we find a significant effect of dACC E/I balance on drift rate (t_{23} = 2.011, p = 0.056, CI_{95} = [-0.012 – 0.837], r = 0.387, p = 0.038, CI_{95}= [0.023 – 0.660]). This indicates that participants with a greater dACC E/I balance show a higher drift towards stay decisions and confirms our model free findings. There were no significant effects in any other region of interest (all p > 0.597) nor with decision boundary (all p > 0.269). We included E/I balances directly in the model rather than correlating E/I balances with individual slopes (after model fitting) since the latter might be biased towards the group mean. However, since the standard DDM without incorporating E/I balance fitted the patch-leaving data best, we ran the exploratory analysis reported above.

For the value guided choice phase, we find that a DDM incorporating a regression model on drift rate fits the accuracy-coded data best. While we do not find any main effect of vmPFC E/I on drift rate (highest posterior density interval (HPDI): [-0.279 – 0.052]), we find an interaction effect between vmPFC E/I and the effects of value difference on drift rate (HPDI: [0.032 – 0.124], <0.001 % of the posterior distribution below zero). Additionally, we find an overall greater drift rate with higher value difference between options (HPDI: [0.515 – 0.627], <0.001% of distribution below zero) and in no brainer trials (HPDI: [1.843 – 2.117],

<0.001% of distribution below zero). This pattern of results matches our findings obtained in regression analyses and again points towards an influence of vmPFC E/I balance onto the speed of value integration.

Supplementary Table 8: Overview of DDM models: Overview of HDDM model specifications.

Free Parameters	**Linear Model**	**DIC**	**Gelman - Rubin**
Patch Leaving Phase			
a, v, ndt, sv, st		3145.12	yes
a, v, ndt, z, sz, st		3104.46	yes
a, v, ndt, z, st, sz	*v*~1+costs+dacc+costs:dacc	4205.75	yes
a, v, ndt, z, st, sz	*a*~1+costs+dacc+costs:dacc	4189.73	no
Value guided choice phase			
a, v, ndt, sv, st		15112.98	yes
a, v, ndt, sv, st	*v*~1+valdiff+NB+vmpfc+valdiff:vmpfc	11818.15	yes
a, v, ndt, sv, st	*a*~1+valdiff+NB+vmpfc+valdiff:vmpfc	14290.62	no

Supplementary Methods

Details of the behavioural task

All stimuli were presented on a grey (RGB: 60, 60, 60) background with a contrast optimized for the MEG recording chamber on a screen in a distance of one meter from the sitting participants. Stimuli were displayed via a projector with a refresh rate of 75 Hz located outside the MEG recording chamber. During patch-leaving, participants were presented with two patches (RGB: 80, 80, 80) framed with a white outline indicating in which patch participants are currently staying. If participants chose to switch they had to pay a travelling cost indicated by the size of a grey bar (RGB: 160, 160, 160) presented between both

patches. In trials where participants chose to switch, the rectangular bar signaling switch costs turned red (RGB: 178, 70 70) and the respective costs were subtracted from the subjects total earnings up to this trial. Afterwards the current patch values were revealed. Patch Values were presented in blue (RGB: 69,102,174). In both stages of the experiment a blue progress bar (RGB: 65,105, 204) was shown at the bottom of the screen indicating subjects current score. Participants selected an option by means of a button press with the right or left index finger, respectively. After value-guided choice, participants received a feedback on both options. If an option was rewarded in the current trial, the bar presenting the reward magnitude turned green (RGB: 46, 139, 60) or red (RGB: 178, 70, 70) otherwise. Every time participants were rewarded, the progress bar grew proportional to the obtained magnitude towards a goal state indicated by a golden rectangle (RGB: 184, 134, 11). The goal in the experiment was to reach the goal state as often as possible.

References

1. Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).

2. Palminteri, S., Wyart, V. & Koechlin, E. The Importance of Falsification in Computational Cognitive Modeling. *Trends in cognitive sciences* **21**, 425–433 (2017).

3. Wilson, R.C. & Collins, A.G. Ten simple rules for the computational modeling of behavioral data. *eLife* **8** (2019).

4. Jocham, G., Hunt, L.T., Near, J. & Behrens, T.E.J. A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat. Neurosci.* **15**, 960–961 (2012).

5. Scholl, J. *et al.* A Role Beyond Learning for NMDA Receptors in Reward-Based Decision-Making—a Pharmacological Study Using d-Cycloserine. *Neuropsychopharmacology* **39**, 2900–2909 (2014).

6. Wiecki, T.V., Sofer, I. & Frank, M.J. HDDM: hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in neuroinformatics* **7**, 14 (2013).

7. Palmer, J., Huk, A.C. & Shadlen, M.N. The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of vision* **5**, 1 (2005).

8. Boehm, U. *et al.* Estimating across-trial variability parameters of the Diffusion Decision Model: Expert advice and recommendations. *Journal of Mathematical Psychology* **87**, 46–75 (2018).

9. Urai, A.E., Gee, J.W. de, Tsetsos, K. & Donner, T.H. Choice history biases subsequent evidence accumulation. *eLife* **8** (2019).

10. Ratcliff, R. & Childers, R. Individual Differences and Fitting Methods for the Two-Choice Diffusion Model of Decision Making. *Decision* **2 (4)**, 237 (2015).

11. Gelman, A. & Rubin, D.B. Inference from iterative simulation using multiple sequences. *Statistical science* **7**, 457–472 (1992).

12. Fouragnan, E.F. *et al.* The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. *Nat. Neurosci.* **22** (2019).

13. Khalighinejad, N. *et al.* A Basal Forebrain-Cingulate Circuit in Macaques Decides It Is Time to Act. *Neuron* **105**, 370-384.e8 (2020).

14. Brockett, A.T., Tennyson, S.S., deBettencourt, C.A., Gaye, F. & Roesch, M.R. Anterior cingulate cortex is necessary for adaptation of action plans. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 6196–6204 (2020).

15. Standage, D. & Paré, M. Persistent storage capability impairs decision making in a biophysical network model. *Neural networks : the official journal of the International Neural Network Society* **24**, 1062–1073 (2011).

16. Wang, X.-J. Neural dynamics and circuit mechanisms of decision-making. *Current opinion in neurobiology* **22**, 1039–1046 (2012).

frontiers in Behavioral Neuroscience

OPINION
published: 13 April 2021
doi: 10.3389/fnbeh.2021.660667

Neuromodulation of Foraging Decisions: The Role of Dopamine

Anna Marzecová[1,2]*, *Luca F. Kaiser*[2] *and Armin Maddah*[2]

[1] *Department of Experimental Psychology, Ghent University, Ghent, Belgium,* [2] *Institute of Experimental Psychology, Heinrich-Heine University, Düsseldorf, Germany*

Keywords: decision making, dopamine, foraging, neuromodulation, patch-leaving

When searching for food, animals need to decide whether they can maximize rewards by harvesting at a current resource, or whether they should instead leave for another foraging site. Humans face similar types of problems when deciding whether to stay with their current job, or to move to a new one with a prospect of better career opportunities. Such decisions to leave, often referred to as patch-leaving decisions, require dynamically weighing the time and energy costs of leaving, as well as the benefits of encountering more rewarding resources at new locations. How neuromodulators are involved in patch-leaving decisions, especially in humans, is, at present, scarcely researched. In their recent study, Le Heron et al. (2020) fill this gap by investigating how these decisions are causally affected by dopaminergic state in an ecologically valid foraging scenario. In their study, participants could choose between collecting reward (milk filling a bucket) at one location (patch) or leaving for another patch which incurred a cost in the form of a fixed travel time. As soon as participants started harvesting (collecting milk) from one patch, the reward per time in that patch decreased exponentially, emulating a depleting resource. To maximize their reward rate, participants were thus faced with the task of continuously comparing the rewards at current location against potential rewards at other locations, whilst taking into account the time cost for leaving.

The optimum solution to this foraging problem is given by the Marginal Value Theorem (MVT, Charnov, 1976; Stephens and Krebs, 1986), which has been shown to predict foraging behavior in many species (Cassini et al., 1993; Hayden et al., 2011). MVT states that the optimal time to leave the current patch is when its marginal reward rate ("foreground") drops below the average reward rate in the environment ("background"). To separately manipulate background and foreground reward rates, the authors created patches that differed in their (initial) reward rates (low, medium, and high yield). These could be encountered in either a rich or poor environment. In the rich environment, participants were most likely to transition to a high yield patch upon leaving the current patch, whereas in the poor environment, encountering a low yield patch was most likely. The reward obtained in the current patch thus constituted the foreground, whereas the proportion of the different patch types determined the background reward rate. MVT predicts that optimally behaving agents will.

H1: leave patches within an environment (i.e., equal background) at the same reward rate for all patch types; therefore leave patches with lower initial foreground reward earlier than patches with higher foreground reward.

H2: leave earlier in general when in rich compared to poor environments (high vs. low background reward rate).

A main effect of background reward rate on patch leaving times was observed, supporting H2. In contrast, pertaining to H1, participants left patches with lower foreground reward rate earlier, but they seemed to exhibit a tendency to stay longer in high yield patches, in contrast with the prediction that at leaving, the foreground rate is the same for all patch types. Additionally, participants stayed in patches longer than optimal ("overharvested") across all patch types, leading to less reward obtained than predicted by MVT. Overharvesting is a phenomenon reported

OPEN ACCESS

Edited by:
Satoshi Ikemoto, National Institute on Drug Abuse (NIDA), United States

Reviewed by:
Martin Zack, Centre for Addiction and Mental Health (CAMH), Canada
Francois Cinotti, University of Oxford, United Kingdom

***Correspondence:**
Anna Marzecová
anna.marzecova@ugent.be

Specialty section:
This article was submitted to Motivation and Reward, a section of the journal Frontiers in Behavioral Neuroscience

Received: *01 February 2021*
Accepted: *15 March 2021*
Published: *13 April 2021*

Citation:
Marzecová A, Kaiser LF and Maddah A (2021) Neuromodulation of Foraging Decisions: The Role of Dopamine. Front. Behav. Neurosci. 15:660667. doi: 10.3389/fnbeh.2021.660667

ubiquitously in the foraging literature (see e.g., Hayden et al., 2011; Kane et al., 2019) and has been related to different factors, including time preferences (Kane et al., 2019), and behavioral variability (Cash-Padgett and Hayden, 2020).

Evidence on neuromodulatory mechanisms underlying value comparisons in foraging environments remains scant. Tonic dopamine (DA) levels have been previously suggested to scale with the average background reward rate (average of prediction errors) in the environment (Niv et al., 2007; Beierholm et al., 2013) and could therefore be considered a key element in signaling decisions to leave a patch (Constantino et al., 2017). Le Heron et al. (2020) thus hypothesized that tonic DA levels would modulate the impact of the background, but not the foreground reward rate on patch-leaving decisions. To test this hypothesis, a group of elderly participants was tested twice on the foraging task under the influence of either placebo or the D2 receptor agonist cabergoline. When "on" cabergoline, participants left patches in the poor environment earlier. In contrast, cabergoline did not modulate the effect of the foreground reward rate on patch leaving. This pattern resonates well with the hypothesized role of tonic DA in encoding the average background reward rate. Since participants generally overharvested, this may also imply a shift toward more optimal behavior.

A 1 mg dose of cabergoline was hypothesized to specifically influence the perceived background reward by increasing tonic DA levels, acting via postsynaptic mechanisms (Brooks et al., 1998). However, there have been discussions of whether similar doses of D2 agonists would instead impact phasic rather than tonic DA signaling (Santesso et al., 2009; Norbury et al., 2013) through a modulation of presynaptic autoreceptors (Frank and O'Reilly, 2006). Given that there has been no possibility to assess pre- vs. post-synaptic medication effects in the current study, one may not exclude the possibility that the cabergoline dose resulted in a reduction of the phasic tone (Frank and O'Reilly, 2006). A recent study has shown that a reduction of phasic DA may lead to an increase in (random) exploration (Cinotti et al., 2019), and could thus promote patch-leaving behavior. Contributions of both the phasic and the tonic mode in modulating perceived background reward rate may be considered, bearing in mind it has recently been suggested that the distinction between tonic and phasic DA release and its relation to behavior may not be as clear-cut as previously thought (Berke, 2018).

In another recent study, DA depletion associated with Parkinson's disease (PD), has been linked to a lower estimate of background reward rates in a previous study. PD patients overharvested to a larger extent than control participants when "off" DA medication, while their performance was comparable to controls when "on" medication (Constantino et al., 2017). In that study, the richness of the environment varied due to long and short travel costs. Notably, the difference in leaving time between control and PD participants was more pronounced in the richer (short travel) environment. This may imply multiplicative effects on the perceived richness of the environment, but contrasts with Le Heron et al. (2020) finding of effects in poorer environments only. Since participants in both studies discussed above can be assumed to differ with respect to their baseline DA levels, and potential compensatory changes to DA systems, different ceiling effects may have brought about differing patterns of results. Noteworthy, Le Heron et al. (2020) increased DA levels by targeting D2 receptors, while the depletion of DA in PD is likely to affect both D1 and D2 type receptors (Seeman and Niznik, 1990). However, D2 receptors, owing to their higher affinity for DA, may be still sensitive to (subtle) variations in DA concentration in PD patients. Additionally, whether the effects of DA manipulation extend to younger healthy populations (with likely higher baseline DA levels) is an open question. Future work should seek to delineate under which specific circumstances DA modulates the influence of perceived environmental richness on behavior.

Importantly, the specific drug effects might potentially be considered in relation to different manipulations of environmental richness in the two studies. According to MVT, the background reward rate is determined by the value of potential alternatives as well as by costs of accessing these options. During traveling, the net reward intake is zero, therefore the agent needs to consider whether the potential benefits in alternative patches are worth the invested cost of time (i.e., the foregone reward while traveling). As in Constantino et al. (2017) study, decreasing travel time costs should lead to earlier patch leaving, since it translates into an increased background reward rate. Travel times have been previously found to influence the leaving threshold in patch leaving tasks (Hayden et al., 2011; Wolfe, 2013; Ramakrishnan et al., 2019). In the study by Le Heron et al. (2020), travel costs were kept constant in both environments. However, since the average expected reward rate is different in both environments, the opportunity costs of time differ. The equal travel times therefore potentially have a distinct effect in poor and rich environments. While the relationship between DA modulations and subjective travel cost estimates has been scarcely addressed so far (Constantino et al., 2017), there is a rich literature about the effects of DA on cost-benefit decisions (Salamone et al., 1994; Beeler and Mourra, 2018). In these paradigms, subjects usually decide whether a potential outcome is worth a certain effort, which is a conceptually similar question as in the reported foraging scenario: "Is my investment worth the expected payoff?". A potential route to an increase in the subjective estimate of environmental richness may be a decrease in the subjective estimate of the opportunity costs of time. It would be interesting for further research to explicitly vary travel time costs to assess the contribution of costs to estimates of environmental richness. Combining the experimental manipulations of travel time costs (Constantino et al., 2017) and patch reward yield proportions determining environmental richness (Le Heron et al., 2020) could thus prove useful to further a comprehensive framework on how DA modulates patch-leaving. To build a full picture of dopaminergic control of patch-leaving behavior, future research should systematically consider pharmacological effects of particular drug manipulations, behavioral consequences of experimental manipulations, and the extent to which learning takes place in the task.

Research on the role of other neuromodulators implicated in patch-leaving decisions has started to emerge. The locus-coeruleus (LC) noradrenaline system may be involved in patch-leaving, as it promotes behavioral flexibility (Aston-Jones and Cohen, 2005). A recent study reported that tonic LC stimulation in rats led to an earlier patch leaving, which was related to an increased decision noise (Kane et al., 2017). Conversely, optogenetic stimulation of serotonergic cells in the dorsal raphe nucleus led to later leaving times in a patch leaving task (Lottem et al., 2018). Additionally, a recent whole-brain imaging evidence showing that persistent serotonergic activity correlates with a state of exploitation (Marques et al., 2020). Furthermore, GABA and glutamate concentrations in the anterior cingulate cortex have been shown to predict patch-leaving behavior in healthy participants (Kaiser et al., 2021).

Understanding patch leaving decisions and their underlying neurochemical mechanisms is of fundamental relevance to understanding many neuropsychiatric disorders (Addicott et al., 2015). Le Heron et al. (2020) and Constantino et al. (2017) therefore provide new evidence of high practical importance by exploring a modulatory role of DA in the encoding of background reward rates in patch leaving decisions.

AUTHOR CONTRIBUTIONS

AMar: conceptualization (equal), investigation (equal), writing-original draft (lead), writing-review, and editing (equal). LK: conceptualization (equal), investigation (equal), writing-original draft (supporting), writing-review, and editing (equal). AMad: conceptualization (equal), investigation (supporting), writing-review, and editing (equal). All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Research Foundation Flanders (FWO) postdoctoral fellowship (12V5620N) and a grant from the Deutsche Forschungsgemeinschaft (JO-787/6-1).

ACKNOWLEDGMENTS

We thank Monja Froböse, Tom Verguts, Bahador Bahrami, Gerhard Jocham, and two reviewers for helpful discussions on the manuscript.

REFERENCES

Addicott, M. A., Pearson, J. M., Kaiser, N., Platt, M. L., and McClernon, F. J. (2015). Suboptimal foraging behavior: a new perspective on gambling. *Behav. Neurosci.* 129:656. doi: 10.1037/bne0000082

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709

Beeler, J. A., and Mourra, D. (2018). To do or not to do: dopamine, affordability and the economics of opportunity. *Front. Integr. Neurosci.* 12:6. doi: 10.3389/fnint.2018.00006

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., et al. (2013). Dopamine modulates reward-related vigor. *Neuropsychopharmacology* 38, 1495–1503. doi: 10.1038/npp.2013.48

Berke, J. D. (2018). What does dopamine mean? *Nat. Neurosci.* 21, 787–793. doi: 10.1038/s41593-018-0152-y

Brooks, D., Abbott, R., Lees, A., Martignoni, E., Philcox, D., Rascol, O., et al. (1998). A placebo-controlled evaluation of ropinirole, a novel D2 agonist, as sole dopaminergic therapy in Parkinson's disease. *Clin. Neuropharmacol.* 21, 101–107.

Cash-Padgett, T., and Hayden, B. (2020). Behavioural variability contributes to over-staying in patchy foraging. *Biol. Lett.* 16:20190915. doi: 10.1098/rsbl.2019.0915

Cassini, M. H., Lichtenstein, G., Ongay, J. P., and Kacelnik, A. (1993). Foraging behaviour in guinea pigs: further tests of the marginal value theorem. *Behav. Process.* 29, 99–112. doi: 10.1016/0376-6357(93)90030-U

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* 9, 129–136. doi: 10.1016/0040-5809(76)90040-X

Cinotti, F., Fresno, V., Aklil, N., Coutureau, E., Girard, B., Marchand, A. R., et al. (2019). Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Sci. Rep.* 9:6770. doi: 10.1038/s41598-019-43245-z

Constantino, S. M., Dalrymple, J., Gilbert, R. W., Varanese, S., Di Rocco, A., and Daw, N. D. (2017). A neural mechanism for the opportunity cost of time. *bioRxiv*. doi: 10.1101/173443

Frank, M. J., and 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

Hayden, B. Y., Pearson, J. M., and Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* 14, 933–939. doi: 10.1038/nn.2856

Kaiser, L. F., Gruendler, T. O. J., Speck, O., Luettgau, L., and Jocham, G. (2021). Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions. *Nat. Commun.* 12:904. doi: 10.1038/s41467-020-20875-w

Kane, G. A., Bornstein, A. M., Shenhav, A., Wilson, R. C., Daw, N. D., and Cohen, J. D. (2019). Rats exhibit similar biases in foraging and intertemporal choice tasks. *eLife* 8:e48429. doi: 10.7554/eLife.48429

Kane, G. A., Vazey, E. M., Wilson, R. C., Shenhav, A., Daw, N. D., Aston-Jones, G., et al. (2017). Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task. *Cogn. Affect Behav. Neurosci.* 17, 1073–1083. doi: 10.3758/s13415-017-0531-y

Le Heron, C., Kolling, N., Plant, O., Kienast, A., Janska, R., Ang, Y.-S., et al. (2020). Dopamine modulates dynamic decision-making during foraging. *J. Neurosci.* 40, 5273–5282. doi: 10.1523/JNEUROSCI.2586-19.2020

Lottem, E., Banerjee, D., Vertechi, P., Sarra, D., oude Lohuis, M., and Mainen, Z. F. (2018). Activation of serotonin neurons promotes active persistence in a probabilistic foraging task. *Nat. Commun.* 9, 1–12. doi: 10.1038/s41467-018-03438-y

Marques, J. C., Li, M., Schaak, D., Robson, D. N., and Li, J. M. (2020). Internal state dynamics shape brainwide activity and foraging behaviour. *Nature* 577, 239–243. doi: 10.1038/s41586-019-1858-z

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

Norbury, A., Manohar, S., Rogers, R. D., and Husain, M. (2013). Dopamine modulates risk-taking as a function of baseline sensation-seeking trait. *J. Neurosci.* 33, 12982–12986. doi: 10.1523/JNEUROSCI.5587-12.2013

Ramakrishnan, A., Hayden, B. Y., and Platt, M. L. (2019). Local field potentials in dorsal anterior cingulate sulcus reflect rewards but not travel time costs during foraging. *Brain Neurosci. Adv.* 3:239821281881793. doi: 10.1177/2398212818817932

Salamone, J. D., Cousins, M. S., and Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine

depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behav. Brain Res.* 65, 221–229. doi: 10.1016/0166-4328(94)90108-2

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., and Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. *Hum. Brain Mapp.* 30, 1963–1976. doi: 10.1002/hbm.20642

Seeman, P., and Niznik, H. B. (1990). Dopamine receptors and transporters in Parkinson's disease and schizophrenia. *FASEB J.* 4, 2737–2744. doi: 10.1096/fasebj.4.10.2197154

Stephens, D. W., and Krebs, J. R. (1986). *Foraging Theory*. Princeton, NJ: Princeton University Press.

Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules in human visual search. *J. Vis.* 13, 1–17. doi: 10.1167/13.3.10

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.