Behavioral demands gate reward integration in prefrontal cortex and ventral hippocampus

Eshaan S. Iyer

Integrated Program in Neuroscience
Faculty of Science
McGill University, Montréal, Québec, Canada

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

© Eshaan Iyer, 2025

Table of Contents

Abstract

The ability to adapt behavior is critical to survival in an ever-changing environment. To maintain and support flexible, adaptive behavior, information about outcomes must be integrated across time. The nucleus accumbens (NAc) is implicated in learning and decision-making by integrating glutamatergic inputs with dopaminergic input from the ventral tegmental area. However, the role of NAc glutamatergic inputs in reward processing remains unclear. Theories of NAc integration suggest distinct functions for inputs from the medial prefrontal cortex (mPFC) and ventral hippocampus (vHip), but there is limited data to support this. To probe this, we utilized dual-site in vivo fiber photometry to simultaneously interrogate pathway-specific neural encoding of outcomes in these two input regions. We first identify a novel mechanism of outcome integration in NAc-projecting cells common to both the mPFC and vHip. Reward drives suppression of neural activity and unrewarded outcomes gradually restore activity, resulting in a moving baseline for subsequent outcomes that tracks the reward statistics of the environment. Despite similar encoding across regions, we identify a level of specialization for each input: while the mPFC invariantly encodes reward, vHip encoding is uniquely anchored to unrewarded outcomes. By correlating activity in these inputs to behavior, and then by leveraging targeted optogenetic manipulations, we first demonstrate that both inputs can modulate task engagement and work co-operatively to do so. We then hypothesized that inhibitory interneurons could be driving the reward-mediated suppression observed in NAc inputs. Using dual-site in vivo fiber photometry to simultaneously record from GABAergic interneuron populations in the mPFC and

vHip, we unexpectedly find a pattern of outcome integration, similar to that identified in glutamatergic neurons that project to the NAc. Using a hidden Markov model to decompose behavior into bouts of exploration and exploitation, we show that the richness of encoding is gated by behavioral state, with more granular encoding of outcome history during exploration. Correlating activity in these populations to behavior, we reveal that while, similar to NAc inputs, they modulate engagement, these regions also modulate choice behavior in a state-dependent manner that is distributed between the mPFC and vHip. Taken together, these results identify a novel mechanism of outcome integration, mediated by a suppression of neural activity to reward. We further demonstrate that neural encoding both depends upon and is modulated by behavioral state and task demands. Finally, this work highlights the cooperative roles of the mPFC and vHip in supporting outcome integration and reward-motivated behavior, underlining the importance of moving towards multi-region and circuit-wide descriptions of the neural bases of behavior.

Résumé

La capacité d'adaptation du comportement est essentielle à la survie dans un environnement en constante évolution. Pour maintenir et soutenir un comportement flexible et adaptatif, l'information sur les conséquences d'une action doit être intégrée dans le temps. Le noyau accumbens (NAc) est impliqué dans l'apprentissage et la prise de décision en intégrant des entrées glutamatergiques avec des entrées dopaminergiques de l'aire tegmentale ventrale. Cependant, le rôle des entrées glutamatergiques au NAc dans le traitement de la récompense demeure flou. Les théories de l'intégration de la récompense dans le NAc suggèrent des fonctions distinctes en ce qui concerne les entrées venant du cortex préfrontal médian (mPFC) et de l'hippocampe ventral (vHip), mais les données étayant cette hypothèse sont limitées. Afin d'étudier ce phénomène, nous avons utilisé la photométrie à fibre à double site *in vivo* pour interroger simultanément l'encodage neuronal spécifique des conséquences d'une action dans ces deux voies glutamatergiques. Nous identifions un nouveau mécanisme d'intégration des conséquences d'une action dans les cellules projetant vers le NAc, commun au mPFC et au vHip. La récompense d'une action entraîne une suppression d'activité neuronale, tandis que la non-récompense restaure progressivement l'activité neuronale. Ceci crée une base de référence dynamique pour les futures réponses comportementales, qui suit les statistiques de récompense de l'environnement. Malgré un encodage similaire entre les régions, nous identifions un niveau de spécialisation pour chaque entrée : alors que le mPFC encode invariablement la récompense, l'encodage du vHip est ancré spécifiquement à la non-récompense. En corrélant l'activité neuronale de ces entrées au comportement, puis en

exploitant des manipulations optogénétiques ciblées, nous démontrons que les deux entrées peuvent moduler l'engagement de l'animal dans la tâche et qu'elles travaillent en coopération afin d'y parvenir. Par la suite, nous avons émis l'hypothèse que les interneurones inhibiteurs pourraient être à l'origine de la suppression liée à la récompense observée dans les entrées du NAc. En utilisant la photométrie à fibre à double site *in vivo* pour enregistrer simultanément l'activité des populations d'interneurones GABAergiques dans le mPFC et le vHip, nous trouvons de manière inattendue un motif d'intégration des conséquences antérieures d'une action, similaire à celui identifié dans les neurones glutamatergiques se projetant vers le NAc. En utilisant un modèle de Markov caché pour décomposer le comportement en périodes d'exploration et d'exploitation, nous montrons que la richesse de l'encodage est restreinte par l'état comportemental : un encodage plus granulaire des conséquences antérieures d'une action est observé lors de l'exploration. La corrélation de l'activité neuronale de ces populations et du comportement révèle que ces dernières modulent non seulement l'engagement à la tâche de la même façon que les entrées du NAc, mais aussi, de façon distribuée entre le mPFC et le vHip, le comportement de choix lorsque celui-ci dépend de l'état comportemental. Globalement, ces résultats identifient un nouveau mécanisme d'intégration des conséquences d'une action : la suppression de l'activité neuronale face à une récompense. Nous démontrons que l'encodage neuronal dépend de l'état comportemental et des demandes de la tâche tout en étant modulé par eux. Enfin, ce travail met en lumière les rôles coopératifs du mPFC et du vHip dans l'intégration des conséquences d'une action, de la récompense et de la

motivation, soulignant l’importance de prendre en compte plusieurs régions du cerveau et leurs interactions afin de mieux comprendre les bases neuronales du comportement.

Acknowledgments

When I landed in Montreal all those years ago on a hot August afternoon, I wasn't sure where I was going but I knew a PhD would be involved. I don't think I realized the scope of the endeavor at the time, but even looking back in hindsight, I would do it all again. These have been some of the most formative years of my life and as I step back to take it all in, I can see just how much I have grown, not only as a researcher but also as a person. For this, I would like to thank the many people that have helped me along the way. This thesis is a testament to their support.

I owe a great deal of my growth to my supervisor Dr. Rosemary Bagot. Thank you for always pushing me to dig deeper and understand more, and for always telling me the hard truth. I am grateful for the countless opportunities you provided me and all the time you dedicated to nurturing my development as a scientist. I am lucky to have your mentorship, it has made me the scientist I am today. Thank you for taking a chance on the American looking to step away from American hegemony (even if only for a bit).

Stewart Biology Building can be a chaotic place to do research and without the support of my fellow Bagots I'm not sure how I would have managed. Thank you to Dr. Vedrana Cvetkovska, Dr. Rand Eid, and Heike Schuler for all of the support and advice through the final stages of my PhD, to Serena Wu for inducting me into your cult of climbing and winter sports, and to Dr. Jessie Muir for the hijinks, reality checks, and "hot goss" (I guess a thank you to VWR is also in order for making that happen). Your friendships have made this PhD the experience of a lifetime, and one that I will cherish fondly. Thanks to you, all the burst pipes and asbestos were worth it (almost).

I want to express my sincere gratitude to Dr. Joëlle Lopez for making so much of this possible. Your help with everything from experiments to dealing with administrative headaches to abstract translations has been invaluable.

I would like to extend my deepest gratitude to Peter Vitaro, whose assistance with running the optogenetics experiments, as well as countless other behavioral experiments not included in this thesis, was indispensable.

To Dr. Becket Ebitz, thank you for always taking the time to provide insightful feedback whether that be about data, models, or manuscripts. This thesis and my computational skills are all the stronger for it.

I would like to thank my advisory committee for all their thoughtful feedback, advice, and guidance over the years. Thank you, Dr. Ross Otto, for first opening the door to the world of reinforcement learning for me, and thank you, Dr. Anna Weinberg, for always encouraging me to think translationally.

Thank you to all my friends who took the time to make the trip out to Montréal to see me: Léa, Kassandra, and Xing: I appreciate all the encouragement over the years, thank you for bringing a bit of California to me.

I'm incredibly grateful for the opportunity to have lived and made such amazing friends in Montreal. Thank you especially to Kat and Abram for always reminding me there is life outside of science. Your friendship has made Montreal feel like home. And to Mathieu, thank you for always pushing me to see the silver lining, this final mile would not have been possible without your support. I've loved running all over Quebec with you.

Finally, this degree would not have been possible without the unconditional love and support of my family. Thank you to my father Sriram Iyer for being there to listen whenever I need to complain and for reminding me that the future is in fact exciting. Thank you to my mother Varsha Naik-Iyer for your unconditional love and support and for starting me on this journey by driving me up and down all those hills in SF. And thank you to my sister, Era Iyer, for all the calls and visits, providing me with much needed distraction. You made home feel not so far away. You all have consistently encouraged my curiosity and love of science and always pushed me to strive for excellence. I would not be where I am today without you. I am lucky to have you all on my side.

Contribution to original knowledge

Chapter 2 presents an original manuscript accepted for publication at Nature Communications where we use *in vivo* fiber photometry to simultaneously record population-level activity in both medial prefrontal cortex (mPFC) and ventral hippocampus (vHip) inputs to the nucleus accumbens (NAc). We find a novel neural motif that encodes information about outcomes, with reward driving suppression of neural activity and unrewarded outcomes gradually restoring activity. We show how task demands can modulate outcome encoding in these regions, challenging the idea that reward-encoding is static and invariant. We also reveal that while mPFC projections to NAc invariantly encode reward, vHip projection to NAc are uniquely anchored to unrewarded outcomes. Using a novel dual-site optogenetic stimulation protocol, we show that both inputs co-operatively modulate task engagement through cumulative glutamatergic drive in NAc. Through this work I also implement a novel analysis approach for fiber photometry data, using a linear mixed modeling approach in conjunction with estimated marginal means to control for inter-session and inter-individual variance. While most analyses of fiber photometry data require the collapsing of trial-by trial data into averages, this new approach eliminates the need to average across trials and animals, conserving critical trial by trial information and increasing statistical power.

Chapter 3 presents an original manuscript submitted for publication wherein we use *in vivo* fiber photometry to simultaneously record GABAergic neurons in mPFC and vHip. We show that GABAergic neurons integrate information about outcomes with

relative suppression of neural activity mapping onto outcome history. Decomposing behavior into bouts of exploration and exploitation using a hidden Markov model, we show how behavioral state gates reward integration. We also demonstrate that while the primary role of these neural populations is to modulate engagement, they also exhibit control over choice behavior that is distributed between mPFC and vHip in a state-dependent manner. Together this work reveals a novel role for inhibitory interneurons in reward processing and in supporting reward-motivated behavior.

Contribution of authors

Chapter 1: Introduction

Eshaan Iyer wrote the introduction with input from Joëlle Lopez and Rosemary Bagot. Parts of the introduction were adapted from Iyer, Kairiss et al 2020.

Chapter 2: Reward integration in prefrontal-cortical and ventral-hippocampal nucleus accumbens inputs cooperatively modulates engagement

Eshaan Iyer and Rosemary Bagot conceived of experiments and methodology. Surgeries were performed by Eshaan Iyer, Jessie Muir, and Serena Wu. Operant training was performed by Eshaan Iyer and Peter Vitaro. Electrophysiology was performed by Yiu Chung Tse. Confocal imaging was performed by Vedrana Cvetkovska. Data analysis was performed by Eshaan Iyer. Eshaan Iyer and Rosemary Bagot wrote the manuscript.

Chapter 3: Behavioral state gates reward integration in GABAergic neurons in medial prefrontal cortex and ventral hippocampus in male mice

Eshaan Iyer and Rosemary Bagot conceived of experiments and methodology. Surgeries were performed by Eshaan Iyer. Operant training was performed by Eshaan Iyer. RNAscope was performed by Vedrana Cvetkovska. Confocal imaging was performed by Joëlle Lopez. Code and guidance for hidden-markov modeling was

provided by Becket Ebitz. Data analysis was performed by Eshaan Iyer. Eshaan Iyer and Rosemary Bagot wrote the manuscript with input from Becket Ebitz.

Chapter 4: Discussion

Eshaan Iyer wrote the discussion with limited input from Rosemary Bagot.

Chapter 5: Conclusion

Eshaan Iyer wrote the conclusion with limited input from Rosemary Bagot.

List of Figures and Tables

Chapter 2

Main Figures

Supplementary Figures

Chapter 3

Main Figures

Supplementary Figures

List of abbreviations

BLA	Basolateral Amygdala
HMM	Hidden Markov Model
CS	Conditioned Stimulus
D1	Dopamine Receptor 1
D2	Dopamine Receptor 2
PFC	Prefrontal Cortex
PVT	Paraventricular Thalamus
mPFC	Medial Prefrontal Cortex
MSN	Medium Spiny Neuron
NAc	Nucleus Accumbens
NAcC	Nucleus Accumbens Core
NAcS	Nucleus Accumbens Shell
RL	Reinforcement Learning
US	Unconditioned Stimulus
vHip	Ventral Hippocampus
VP	Ventral Pallidum
VTA	Ventral Tegmental Area
WSLS	Win Stay Lose Shift

Chapter 1

Introduction

Understanding Reward

The ability to learn from one's environment is crucial for survival. Learning within ever-changing environments is necessary to maximize reward while avoiding danger. Differences in how individuals use reward to navigate and learn from their environments may arise due to differences in the computations that underlie reward learning and integration in the brain. Deficits in how these latent processes affect behavior have been described in a range of psychiatric disorders including depression, schizophrenia, and anxiety disorders (Amir et al., 2012; Bishop & Gagne, 2018; Köther et al., 2021; Lawlor et al., 2020; Lloyd et al., 2024; Paulus & Yu, 2012).

Reward encompasses a wide range of stimuli from food to social interactions. Generally, we can understand reward to be an experience that induces a positive affective state. Behaviorally we can operationalize reward as a stimulus that drives approach behavior or increases the likelihood of emitting a behavior that produces this stimulus (Schultz et al., 1997; White, 1989). It is this behavioral consequence of reward that defines the concept of reinforcement. Reinforcement is the tendency of a stimulus to drive behavior and can be classified into either positive reinforcement or negative reinforcement. With positive reinforcement, behavior is driven by the presentation of an appetitive stimulus such as reward. In the case of negative reinforcement, behavior is driven by the removal of an aversive stimulus such as an unpleasant noise (Nevin &

Mandell, 2017). This is formalized in Thorndike's law of effect, where the consequences of a given action act to alter the probability of an individual performing that action again (Thorndike, 1927). It is through this relationship that we have come to talk about reward as reinforcing. For example, when encountering a block of cheese for the first time a mouse may be initially hesitant to consume. However, after a rewarding experience consuming this block of cheese, the mouse will now be more likely to approach and start reorienting its behavior towards seeking out blocks of cheese. How the brain utilizes rewarding information to drive this sort of adaptative reward-motivated behavior remains an outstanding question and is the focus of this thesis.

Reward and Behavior

To drive adaptive and reward-motivated behavior, reward can exert many different influences on a wide variety of behaviors. This can range from modulation of more complex behaviors such as foraging and exploration, to more simple behaviors such as the consumption of a food reward. Of these behaviors, the influence of reward on choice and decision-making represents one well studied area of research.

Frameworks for studying choice and decision-making

The influence of reward on decision-making can be understood intuitively as following the logic of Thorndike's law of effect wherein the results of an action act to alter the strength of the action itself (Thorndike, 1927). In rewarding environments, this means that the most recent outcome exerts the most influence on the current choice. Thus, choice behavior on a given trial, n, depends upon the choice and outcome on preceding trial, n − 1. This relationship has been formalized as a simple strategy known

as win stay/lose shift (WSLS) (Herrnstein, 1997). This framework captures whether the previous trial was rewarded or not and if the choice on the current trial repeats or switches from the previous choice (Dalton et al., 2014). For example, if the individual is rewarded on trial n − 1 and then chooses the same option on trial n, this is termed a 'win stay', whereas if a different choice is made on trial n following reward on n − 1, this is termed a 'win shift'. If the individual is not rewarded on trial n − 1 and makes the same choice on trial n, this is termed 'lose stay', and a 'lose shift' if a different option is chosen. The proportion of win stay and lose shift responses made is interpreted as a metric of sensitivity to positive and negative outcomes, respectively. That is, higher win stay probabilities are interpreted as increased sensitivity to positive feedback while higher lose shift scores are interpreted as increased sensitivity to negative feedback (Onge et al., 2011).

Reinforcement learning (RL) is another conceptual framework used to describe decision-making behavior in rewarding contexts. Here, learning accrues from the discrepancy between an expectation and an outcome, termed a prediction error (Rescorla & Wagner, 1972). This prediction error is used to incrementally update an individual's estimate of the value, or probability of reward, associated with a given choice. This estimate that a given choice is the most rewarding option drives choice. Recent years have seen a surge of interest within the field of behavioral neuroscience in applying RL models to probe fundamental questions about how individuals learn (Bathellier et al., 2013; Gustafson & Daw, 2011; Kuchibhotla et al., 2019; Langdon et al., 2019; Noworyta-Sokolowska et al., 2019; Stachenfeld et al., 2017). Within the RL framework, an 'action value' is learned through prediction errors that are modulated by a

learning rate, which weights the influence of prediction errors on yielding new action values. These action values are then transformed into actions using a choice rule such as the ‘softmax’ rule. At its core, RL is best suited to provide a model of how individuals interact with their environment, whether that be the rate at which they learn from feedback, internal representations of action values, or choice transition probabilities. (Langdon et al., 2019; Noworyta-Sokolowska et al., 2019; St-Amand et al., 2018; Verharen et al., 2019). This is accomplished by fitting different parameters to individual behavior. Then, to gauge the descriptive accuracy of various models at either the individual or group level, model fit can be used as a metric.

State-based models provide an alternative framework to understanding how reward interacts with choice behavior. Rather than constantly tracking relative value across choices, here, choice is conceptualized as a function of a belief about what action is the correct action to take. In rewarding contexts, this can be understood to be the action that will lead to reward. The conceptual framing behind these models relies on the assumption that individuals are performing a mental inference to arrive at a belief about which action is the most advantageous action to take. Commonly used approaches to decompose choice behavior into a state-based structure include Bayesian inference, and more specifically hidden Markov models (HMM) (C. S. Chen et al., 2021; Ebitz et al., 2018; Eckstein et al., 2022; Mishchanchuk et al., 2024). Bayesian inference utilizes Bayes’ theorem to infer the likelihood of being in a state (i.e. correct, incorrect) given a previous choice and its outcome. HMMs approach this by modeling sequences of choices as outputs generated by transitions through a series of latent states (i.e. state where left lever is rewarding). While simple state-based models assign

a singular state per possible action, not accounting for other types of choice behavior, it is possible to build on these models, incorporating states that can account for exploration wherein individuals choose to sample between options. This can be contrasted with states of exploitation, wherein individuals repeatedly sample a rewarding option foregoing opportunities to sample other options. There is increasing evidence that these models are more accurate descriptors of choice behavior in stochastically rewarded environments and provide greater flexibility by accounting for the influence of multiple states and contexts in driving behavior (Eckstein et al., 2022).

While this section covers several approaches that can be utilized to assess the influence of reward on choice and decision-making behavior, there is no clear consensus on which approach to use. This is due to the fact that reward does not have a static influence on behavior and so either of these approaches may be valid depending on the context and task examined. However, some general guidelines can be applied when considering which approach to implement. For high-level and blunt descriptions of reward's effect on choice behavior, WSLS is useful, with the expectation that under normal conditions, win-stay probabilities would be greater than chance. This approach, however, lacks the flexibility of describing the underlying processes that mediate reward's influence on behavior. In these circumstances RL and simple Bayesian inference models are appropriate. As these approaches are hypothesis-driven, in that researchers must define the model to which the data is then fit, care must be taken to fit the data to a wide range of models to improve the chance of selecting a model that closely aligns with the behavior and underlying processes. One major shortcoming of this approach is the lack of ability to account for the possibility that

individuals may be switching between a range of different latent processes to drive their behavior. It is in this case that hidden Markov models are useful. This approach allows that multiple latent processes can drive behavior, finding the probability that a behavior was performed given a latent state. Major caveats of this approach include the high computational cost, as well as the potential challenges in interpretability of results as the number of states increases. This approach also tends to be agnostic to the computational processes underlying the various latent states, making it harder to directly test hypotheses about the underlying mechanisms that drive reward's influence on decision and choice behavior.

Engagement

In addition to modulating choice and decision-making behavior, rewarding stimuli drive approach behavior and individuals will also work to obtain reward, suggesting a fundamental link between reward and engagement. However, the relationship between how reward dynamically modulates engagement has been less studied than choice behavior. Experimental and computational evidence points to a relationship between reward and engagement, with rats showing faster reaction times as the magnitude of expected rewards increases (Brown & Bowman, 1995; Niv et al., 2007). This effect has also been observed in humans: in a discrimination task as the average reward rate increases, the reaction times decreases, indicating increased engagement (Beierholm et al., 2013; Guitart-Masip et al., 2011). Behaviorally, engagement can be operationalized as the rate or latency of response, also known as vigor. Several groups have formalized the relationships between reward and vigor into computational models (Dezfouli et al., 2019; Niv et al., 2007) that capture phenomena observed in behavioral

experiments, with response latency scaling according to recently experienced reward (Cohen et al., 2015; Niv et al., 2007; Wang et al., 2013).

Given that reward modulates both choice behavior as well as engagement, models of choice (e.g. RL, HMM) can also be applied to conceptually integrate engagement. For example, reinforcement learning models show that estimated value of a given choice is inversely related to trial initiation time in rats (Wang et al., 2013). Likewise, state models of behavior show that exploitation states are associated with faster latencies to respond that exploratory states (C. S. Chen et al., 2021). Together this provides clear evidence of a relationship between reward and engagement and suggests the possibility that choice and engagement may also interact with each other, giving rise to complex behavioral control.

Studying Reward

The behavioral impacts of reward on choice and engagement reflect the integration of many component processes. To dissect these processes and their neural substrates, systems and behavioral neuroscientists utilize a variety of paradigms.

Pavlovian paradigms

The simplest of these paradigms is Pavlovian, or classical conditioning. These paradigms focus on stimulus-outcome learning and are often used in systems neuroscience to assess reward encoding across the brain. Here, a neutral stimulus, termed the *conditioned stimulus* (CS), is paired with a reward, i.e. the *unconditioned stimulus* (US), named for its ability to *unconditionally* produce innate responses (i.e. consumption, approach behavior). Through repeated presentations of a CS followed by

a US, animals learn to associate the CS and the US. As the CS develops predictive value for the outcome, the innate behavior associated with the US shifts to the CS. Once the CS-US association is learned, by probing responses to the CS, rather than the US, it is possible to dissociate the cognitive and affective state associated with reward from the basic sensory experience of the reward itself (e.g. the gustatory processing of chocolate milk). Pavlovian paradigms are powerful paradigms for probing the neural encoding of reward as well as how animals learn stimulus-outcome associations across relatively long timescales. However, these paradigms have limited ability to probe how reward dynamically shapes behavior in changing environments and do not consider how reward influences action-outcome learning or decision making.

Operant paradigms

To study action-outcome learning, operant, or instrumental conditioning paradigms are often used. Here, animals learn the association between emitting a certain action (i.e. lever press, nose poke) and obtaining a reward, (e.g. water or food). The simplest operant paradigms involve training animals to respond a fixed number of times to yield a reward. These are known as fixed ratio tasks and are used to study formation of action-outcome associations. Variable ratio tasks integrate probabilistic reward, for example in a variable ratio 5 schedule, 20% of responses will be rewarded such that on average only one in five responses will be rewarded. These tasks are useful for studying of how reward modulates engagement, and more specifically vigor. To study how reward modulates decision-making, behavioral neuroscientists use tasks such as n-armed bandit and reversal learning tasks which are well suited to exploring how animals utilize information about actions and their resulting outcomes to drive

future behavior (Lloyd et al., 2024). In these tasks, animals are confronted with a set of choices each with varying probabilities of reward that are unknown to the animal. Here animals must use and track feedback about reward to make an optimal decision given a set of choices. The need to constantly track information about reward and the omission of reward in order to inform decision-making behavior makes this task ideal for studying and modeling how reward influences choice.

Reward in the brain

To produce reward-motivated behavior, the brain must process reward and reward-associated information. This is controlled by several interconnected brain regions, including the ventral tegmental area (VTA), nucleus accumbens (NAc), prefrontal cortex (PFC), and hippocampus. The most thoroughly characterized of these circuits is the dopaminergic circuit from VTA to the NAc which plays an important role in signaling reward and orienting behavior towards reward and reward-predictive cues. VTA dopaminergic projections also innervate other reward regions, including the PFC, hippocampus, and amygdala. These regions are theorized to each play specialized roles in supporting various reward-motivated behaviors. For example, the PFC is thought to provide executive control over behavior while the hippocampus might support goal-directed behavior. The specific role of these regions in supporting reward processing can often be hard to parse due to complex interconnectivity of reward circuits. For example, the hippocampus sends projections to the PFC while the PFC sends projections to the amygdala and the thalamus (Collins et al., 2018; Liu & Carter, 2018; McGarry & Carter, 2017). Many of these rewards processing regions, including

mPFC, hippocampus, VTA, and amygdala, also send projections to the NAc, situating the NAc as a sort of information bottleneck with a critical role in integrating reward information from across the brain to support reward-motivated behavior.

The Nucleus Accumbens and Reward

The NAc is thought to bring together information about reward, salience, context, and emotional state by integrating glutamatergic and dopaminergic input from various regions throughout the brain to promote motivated and goal-directed behavior (Floresco, 2015). This section will examine the neuroanatomy and composition of the NAc and its role in reward processing and motivated behavior, drawing from the rodent literature.

Nucleus Accumbens Compartmentalization

Structurally, the NAc is divided into two main subregions: the core (NAcC), and the shell (NAcS). Experimental evidence suggests that these two subregions exert differential control over reward-motivated behavior. While the NAcC is implicated in promoting approach behavior, the NAcS is thought to play more of a role as a filter, suppressing behaviors and information that may interfere with goal-directed and reward-motivated behavior (Floresco, 2015; Floresco et al., 2018). Though both of these subregions are responsive to reward (G. Chen et al., 2023; Day et al., 2011; Loriaux et al., 2011; Roesch et al., 2009), the NAcC has been linked to encoding information about action selection, while NAcS activity has been more linked to value, likely tracking value fluctuations across time (Stopper & Floresco, 2011; West & Carelli, 2016). Further evidence of this value-tracking function of the NAcS is found in studies that observe

encoding of motivational value associated with relative value modulated either by food rewards or internal state (G. Chen et al., 2023; Loriaux et al., 2011; Sackett et al., 2017; West & Carelli, 2016). These studies point to complex but important roles for NAc subregions in supporting motivated behavior.

Nucleus Accumbens Medium Spiny Neurons

Within the NAc, medium spiny neurons (MSNs) represent the majority of cells, representing over 95% of NAc neurons (Gerfen et al., 1990). MSNs are GABAergic projection neurons, mainly targeting the ventral tegmental area (VTA), ventral pallidum (VP), and substantia nigra (Francis & Lobo, 2017), to suppress the activity of their target neurons. MSNs are characterized by their low excitability and low spontaneous activity, requiring excitatory drive from NAc inputs to elicit the firing of action potentials, consistent with the role of the NAc as an integrator of information from various upstream regions. Further, MSNs exist in “up” and “down” states, that refer to two membrane potentials, subthreshold to action potential generation, that they can oscillate between (Stern et al., 1997; C. J. Wilson & Groves, 1981; C. Wilson & Kawaguchi, 1996). The "down state” is characterized by more hyperpolarization, increasing the threshold to generate action potentials, while the “up state” is characterized by more depolarization, decreasing the threshold to generate action potentials (C. Wilson, 2008).

MSNs can also be characterized by the dopamine receptors they express, either dopamine D1 or D2 receptors, with some MSNs, notably in the NAc shell, expressing both (Kravitz & Kreitzer, 2012). These two receptors differ in their influence on MSN excitability, with D1 receptors increasing and D2 receptors decreasing excitability through Gi and Gq coupled receptors, respectively (Fisone et al., 2011; Kebabian &

Calne, 1979; Thibeault et al., 2019). D1 and D2 MSNs in the NAc also tend to differ in their projection targets. While D1 MSNs tend to target both the VTA and VP, D2 MSNs project only to the VP (Smith et al., 2013; Voorn et al., 2004). These functional and projection-specific differences in NAc MSNs provide a basis by which MSN subtypes can differentially modulate reward-motivated behavior.

Theories of NAc function ascribe opposing roles to D1 and D2 MSNs in supporting reward-motivated behavior, with D1 MSNs supporting reward-associated behavior and D2 MSNs supporting aversion (Calipari et al., 2016; Guillaumin et al., 2023; Kravitz et al., 2012; Lobo et al., 2010; Soares-Cunha, Coimbra, Sousa, et al., 2016; Thibeault et al., 2019). For example, in mice, lever-pressing paired with stimulation of D1 MSNs reinforces this lever-pressing behavior, whereas stimulation of D2 MSNs under these conditions biases animals towards a non-stimulation paired lever (Kravitz et al., 2012). Conversely, inhibiting D2 MSNs increases motivation, though with some deficits in goal-directed behavior (Baldo et al., 2002; Carvalho Poyraz et al., 2016; Gallo et al., 2018). Further supporting this, calcium imaging studies have found that interaction with hedonic reward increases D1 MSN activity but decreases D2 MSN activity (Guillaumin et al., 2023). Learning about reward also requires D1 and D2 MSNs, with D1 MSNs supporting integration of positive feedback and D2 MSNs supporting integration of negative feedback (Verharen et al., 2019). Similar relationships have been observed in studies of reward learning and drugs of abuse. D1 MSNs increase, while D2 MSNs decrease their activity to cocaine (Calipari et al., 2016). Stimulation of D1 MSNs also has been observed to increase drug-seeking behavior while stimulation of

D2 MSNs has been observed to decrease drug-seeking behavior (Koo et al., 2014; Lobo et al., 2010).

However, there is also evidence that within the NAc, D1 and D2 MSNs may cooperatively organize reward-motivated behavior, complicating the dichotomous role they are often thought to play in supporting behavior (Ikemoto et al., 1997; Soares-Cunha, Coimbra, Sousa, et al., 2016). For example, suppression of reward-oriented lever-pressing is observed following blockaded of both D1 and D2 receptors, implying cooperative control over reward-motivated behavior (Nowend et al., 2001). Brief stimulation of both D1 and D2 MSNs drives conditioned place preference while longer duration stimulation of both D1 and D2 MSNs drives aversion, demonstrating the importance of temporal factors in how MSNs contribute to behavior (Soares-Cunha et al., 2020). Recent studies have also shown that inhibition of both D1 and D2 MSNs decrease motivation to retrieve hedonic reward, while increases in motivation are associated with activation of D1 MSNs, and D2 MSNs under certain circumstances, suggesting a complicated and not fully bidirectional role of MSN subtypes in supporting reward-motivated behavior (Guillaumin et al., 2023; Soares-Cunha, Coimbra, David-Pereira, et al., 2016). Put together, these data point to an important, yet complex role for D1 and D2 MSNs in driving reward-motivated behavior, likely dependent on context, timing, and task.

Dopamine and reward

NAc MSNs neurons are characterized based on the dopamine receptors that they express, highlighting the critical role of this neuromodulator in NAc function. The VTA provides dopaminergic input to the NAc, and is critical in linking rewarding stimuli

with their associated cues and actions (Berridge, 2007; Jeong et al., 2022; Nestler & Carlezon, 2006). This has been most famously demonstrated by observations of the shifts in dopamine neuron activity across learning. While dopamine neuron firing is initially aligned to reward delivery, as a cue comes to predict reward, this activity shifts back in time to align with the reward predicting cue (Jeong et al., 2022; Schultz et al., 1997).

The importance of dopaminergic signaling has been demonstrated by ablation studies where global ablation of dopamine signaling leads to decreases in motivated behavior marked by extreme hypophagia and hypoactivity (Szczypka et al., 2001). More targeted modulation of dopamine signaling within the NAc specifically has been observed to decrease reward-motivated behavior (Aberman & Salamone, 1999; Salamone et al., 2018). Experimentally, dopamine release has been observed during a variety of both appetitive and aversive experiences (Bromberg-Martin et al., 2010; Hamid et al., 2016; Jeong et al., 2022; Mingote et al., 2019; Willmore et al., 2022). This dopamine release has been implicated in signaling salience of events and seems to encode various components of task structure (de Jong et al., 2019; Saddoris et al., 2015a). That dopamine release shifts to cue onset and correlates with the expected size of a given reward supports the possibility that dopamine signals a prediction, often formalized as reward prediction error (Keiflin & Janak, 2015; Sackett et al., 2017; Schultz, 2016; Schultz et al., 1997). Together these studies demonstrate the importance of dopamine transmission in bridging the gap between relevant cues and rewarding outcomes, suggesting that dopamine is likely playing a role encoding different dimensions of salient and reward-oriented features to drive reward-motivated behavior.

Nucleus accumbens glutamatergic input

While NAc dopamine modulates excitability of MSNs (Carter et al., 2007), ultimately, glutamatergic inputs drive MSN action potentials (Floresco, 2007; Goto & Grace, 2005). Within the NAc, dopamine axons are often found beneath glutamatergic synapses, providing a physical and localized basis for the interaction between glutamate and dopamine (Floresco, 2007). Together, this indicates an important role for glutamatergic input in supporting reward processing and reward-motivated behavior in conjunction with dopaminergic input. Glutamatergic input to the NAc is provided by the prefrontal cortex, hippocampus, amygdala, thalamus, and VTA. Despite theoretical perspectives that these inputs send qualitatively dissociable information to the NAc, the role of these afferents in supporting reward processing and motivated behavior remains relatively unexplored. This section will discuss current understandings of the role of these afferents in supporting these processes.

The Medial Prefrontal Cortex

The medial prefrontal cortex (mPFC) sends projections to both the NAcC and the NAcS . While the NAcC is innervated by prelimbic subregions of the mPFC, the NAcS preferentially is innervated by the infralimbic subregion of the mPFC (Piantadosi et al., 2020). Globally, the mPFC encodes information about previous actions and outcomes in reward-based tasks (Sul et al., 2010). In cue-based paradigms, mPFC projections to the NAc show increased activity to reward-paired cues when compared to neutral cues, and stimulation of this pathway can potentiate anticipatory licking behavior during task acquisition (Otis et al., 2017). mPFC projections to the NAc, specifically, have also been demonstrated to encode a variety of information about reward and choice,

bridging information about current actions and outcomes across multiple trials (Parker et al., 2022; Spellman et al., 2021). Different subregions within the PFC likely send different information to the NAc. For example, while activation of the medial orbital frontal cortex promotes appetitive behavior, activation of the infralimbic cortex suppresses appetitive behavior (Richard & Berridge, 2013). Animals will lever press for stimulation of the mPFC-NAc pathway and stimulation drives conditioned place preference (Britt et al., 2012; Lind et al., 2023). However, disruption in mPFC-NAc signaling through inhibition of pathway activity increases impulsive behavior, increases responding on nonreinforced cues, and decreases conditioned suppression, indicating an important role of this input in supporting and maintaining reward-motivated behavior (Hamel et al., 2022; Keistler et al., 2015; Wenzel et al., 2023). Though complicated and not always consistent, this body of literature suggests that the mPFC-NAc pathway is important for reward processing and provides evidence that disruptions in mPFC-NAc signaling also disrupt reward-motivated behavior. This suggests that the mPFC-NAc pathway may be integrating information about outcomes and task structure in order to regulate reward-seeking behavior.

The Ventral Hippocampus

Ventral hippocampal (vHip) inputs to the NAc primarily originate from the ventral subiculum, with some projections also originating from the CA1 region of the ventral hippocampus, primarily innervating the medial NAcS (Britt et al., 2012; Thierry et al., 2000). Behaviorally, the vHip input to the NAc is also implicated in reward motivated behavior, though the extant literature is considerably less elaborated than that of the mPFC-NAc pathway. Optogenetic stimulation and manipulations to facilitate the

strength of vHip inputs to the NAc reinforce instrumental behavior and conditioned place preference (Britt et al., 2012; LeGates et al., 2018). vHip-NAc activity also associates with reward-seeking behavior and is suppressed during consummatory behavior (Reed et al., 2018). This aligns with findings that suppression of vHip activity during reward-seeking is necessary to promote goal-directed behavior (Yoshida et al., 2019, 2021). Pharmacological and behavioral interventions that disrupt suppression of the vHip suppress or otherwise alter goal-directed behavior (Yoshida et al., 2019, 2021). Disruptions in vHip-NAc signaling also bias animals towards small and immediate rewards as opposed to larger delayed rewards (Abela et al., 2015). Together this suggests an important role for the vHip-NAc pathway in integrating information about outcomes and promoting efficient goal-directed behavior.

The Amygdala

Amygdalar inputs to the NAc primarily originate from the basolateral (BLA) region, with inputs terminating in both the NAcS and NAcC (Britt et al., 2012). There is also evidence of projection populations originating in the central amygdala which preferentially innervate the NAcC (Borrego et al., 2022). The BLA-NAc pathway is positively reinforcing, with stimulation of the projection promoting conditioned place preference as well as self-stimulation behavior (Britt et al., 2012; Dieterich et al., 2021; Stuber et al., 2011). Inhibition of this projection, however, results in reduced behavioral responding for a sucrose reward and overall decreased sucrose preference (Dieterich et al., 2021; Stuber et al., 2011). The BLA-NAc pathway has also been shown to preferentially encode information about rewarding outcomes and generally suppress activity during reward consumption (Ambroggi et al., 2008; Beyeler et al., 2016; Reed et

al., 2018). The BLA-NAc projection is also heavily implicated in supporting goal directed behavior, linking information about reward to the cues that predict it. NAc projecting cells in the BLA associate with behavioral response to both conditioned as well as unconditioned stimuli (He et al., 2023). Optogenetic manipulations of BLA-NAc activity also affect choice behavior with inhibition biasing animals away from preferred choices and towards riskier and larger rewards (Bercovici et al., 2018; van Holstein et al., 2020). This suggests the BLA-NAc pathway plays an important role in communicating information about reward to the NAc to help guide choice behavior and cue-outcome pairings.

The Thalamus

Within the thalamus, most nucleus accumbens projection cells arise from the paraventricular nucleus (PVT), with contributions also from the central medial nucleus (De Groote & de Kerchove d'Exaerde, 2021). Though these projections are primarily to the NAcS , the NAcC also receives some input from the PVT (Dong et al., 2017). This PVT input modulates MSN activity through the release of glutamate as well as through the stimulation of dopamine release in neighboring terminals (Parsons et al., 2007).

The PVT-NAc pathway has been implicated in mediating behavioral aversion with stimulation inducing real-time place aversion (Do-Monte et al., 2017; Engelke et al., 2021; Zhu et al., 2016). However, this simple interpretation is challenged by studies that find highly variable real-time place preference in response to PVT-NAc stimulation (Lafferty et al., 2020). In fact, animals will lever-press for PVT-NAc stimulation, suggesting that this pathway may be mediating both rewarding and aversive processing (Lafferty et al., 2020). This is supported by *in* vivo single-cell imaging studies which find

that projections from the thalamus to the NAc encode both rewarded as well as unrewarded outcomes in a reward-learning task (Otis et al., 2019; Parker et al., 2022). Interestingly, at a population level, reward suppresses activity in this projection (Otis et al., 2019). Inhibition of this pathway during omission of reward in a task where animals learn to lever press for a reward results in increased lever-pressing behavior, whereas chemogenetic excitation results in decreased behavioral responding (Do-Monte et al., 2017; Lafferty et al., 2020). This pattern of behavior is also seen in simpler approach tasks, where increased activity in the PVT-NAc pathway is associated with decreased food-seeking and decreased approach behavior (Engelke et al., 2021). This suggests that increased activity in this pathway may function as a sort of brake on behavior and decreasing activity through optogenetic inhibition or reward may increase behavioral engagement. Together, this points to a role of the PVT-NAc pathway in integrating information about outcome and task-relevant information to continuously modulate engagement.

The Ventral Tegmental Area

While the NAc input originating from the VTA is usually associated with dopaminergic modulation in the NAc, subpopulations of these cells release glutamate as well as co-release glutamate and dopamine (Tecuapetla et al., 2010; Warlow et al., 2024; Yamaguchi et al., 2011). This source of glutamate is thought to inhibit MSNs in the NAc by activating inhibitory parvalbumin-expressing GABAergic interneurons (Qi et al., 2016). The behavioral contribution of the glutamatergic VTA-NAc input is still not well understood, especially in the context of reward. Animals will nose-poke for stimulation of the glutamatergic VTA-NAc pathway, however, stimulation of the

glutamatergic VTA-NAc projection can also drive conditioned place aversion and active responding to terminate stimulation (Qi et al., 2016; Yoo et al., 2016). When mice are presented with two levers, one paired with the delivery of a food reward, and the other paired with both the delivery of a food reward as well as with stimulation of the glutamatergic VTA-NAc pathway, animals decrease their responding on the stimulation-paired lever (Qi et al., 2016). Stimulation of the glutamatergic VTA-NAc projection has also been shown to block instrumental and conditioned place preference reinstatement of drug-seeking behavior (Barbano et al., 2024). This suggests that increased activity of the glutamatergic VTA-NAc pathway may actually suppress behavior by activating parvalbumin-expressing GABAergic interneurons to inhibit MSNs.

A focus on medial prefrontal cortex and ventral hippocampus

Of the glutamatergic projections to the NAc, the projections from the mPFC and the vHip have a unique relationship and may play a joint role in reward processing and supporting motivated behavior. *In vivo* electrophysiological data demonstrates a complex competitive role between mPFC and vHip inputs to the NAc. vHip inputs to the NAc are able to switch cells in the NAc from a quiescent state to an ‘up’ state, allowing them to be depolarized by mPFC inputs (O’Donnell & Grace, 1995). Inactivation of the mPFC attenuates the ability of the vHip to depolarize cells in the NAc, however, induction of long-term potentiation in the vHip prior to mPFC inactivation prevents this attenuation (Belujon & Grace, 2008). Furthermore, individual NAc cells receive convergent inputs from both the vHip and the mPFC (French & Totterdell, 2002). Together this implies the need for balanced vHip input to enable mPFC control over NAc as well as the need for balanced mPFC input to enable vHip control over NAc. This

defines a relationship whereby NAc activity is dynamically influenced by co-operation and competition between mPFC and vHip inputs to NAc. The complex nature of this interaction highlights the necessity to study these inputs in conjunction, rather than in parallel and provides a physiological basis for joint control over reward processing and motivated behavior.

Rationale and Aims

The NAc integrates a variety of glutamatergic inputs with dopaminergic input from the ventral tegmental area to encode motivationally relevant stimuli to support learning, decision making and goal-directed behavior (Christoffel et al., 2021; French & Totterdell, 2002; Goto & Grace, 2005; Grace et al., 2007; O'Donnell & Grace, 1995; Yang et al., 2018). The importance of NAc dopaminergic signaling in these processes is widely appreciated, with a large body of research describing specific roles for dopamine in reward prediction error, saliency, and reward expectation (Howe et al., 2013; Kutlu et al., 2021; Mohebi et al., 2019; Saddoris et al., 2015b; Syed et al., 2016; Wassum et al., 2012). While NAc dopamine modulates excitability of medium spiny neurons (MSNs), ultimately, glutamatergic drive determines MSN action potential firing (Carter et al., 2007). Despite this fundamental physiological requirement for glutamatergic input, the role of these afferents in supporting reward processing remains relatively unexplored. Theoretical perspectives suggest that these inputs send qualitatively dissociable information to the NAc, implying input-specific roles in the computations that underlie reward processing and reward-motivated behavior. However, direct experimental evidence of this remains limited. The primary goal of this thesis is to address this gap, by determining 1) how two distinct glutamatergic inputs to the NAc process reward, 2) the redundancy and specificity of reward processing between these glutamatergic inputs, and 3) the behavioral contribution of activity in these inputs towards reward-motivated behavior.

Of the various glutamatergic inputs to NAc, the inputs from mPFC and vHip may play a unique and cooperative role in regulating the processing of reward-related

information in the NAc. Electrophysiological data demonstrates that activity in vHip can gate the mPFC input to NAc and vice versa, providing a circuit mechanism for joint control of reward-motivated behavior. (Belujon & Grace, 2008; O'Donnell & Grace, 1995). To examine this possibility, this thesis focuses on comparing these two regions which provide converging glutamatergic input to the NAc medial shell and also differ in projection strength, attributed behavioral and computational function. In the first set of studies, I examine how population-level activity in both mPFC-NAc and vHip-NAc encodes information about outcomes, with reward driving suppression of neural activity and unrewarded outcomes gradually restoring activity. I then expand this observation to describe how this motif can be used across many trials, tracking prior reward history as a graded function of unrewarded outcomes. Despite evidence of similar encoding motifs, I hypothesized that there should be differences in the information contained in mPFC-NAc and vHip-NAc, using task degradations and conditional entropy analyses to find that while mPFC-NAc consistently encodes outcome, vHip-NAc encoding is anchored to unrewarded outcomes. Examining the behavioral relevance of reward integration in mPFC-NAc and vHip-NAc revealed a co-operative and, to an extent, redundant role of these pathways in dynamically modulating engagement in rewarding environments.

The second study expands upon this work, based on the hypothesis that inhibitory populations in mPFC and vHip might contribute to the reward-mediated suppression identified in the first study. Surprisingly, I again find a similar pattern of reward integration as described in the first study. Inhibitory neurons in mPFC and vHip, suppress activity to reward, tracking prior reward history. However, in inhibitory neurons,

this seems to be gated by behavioral state. Examining the behavioral relevance of state-gated reward integration in mPFC and vHip again revealed a role in dynamically modulating task engagement in both regions, but a state-dependent role in modulating choice behavior in each region.

Together, this work identifies a novel motif of population-level neural activity that integrates outcome-information across trials in NAc projecting regions and highlights the relevance of task demands and behavioral state in efficient outcome encoding and reward processing.

Chapter 2

Reward integration in prefrontal-cortical and ventral-hippocampal nucleus accumbens inputs cooperatively modulates engagement

Eshaan S. Iyer[1], Peter Vitaro[2], Serena Wu[1], Jessie Muir[4,5], Yiu Chung Tse[2], Vedrana Cvetkovska[2], Rosemary C. Bagot[2,3]

1. Integrated Program in Neuroscience, McGill University, Montréal, Canada

2. Department of Psychology, McGill University, Montréal, QC, Canada

3. Ludmer Centre for Neuroinformatics and Mental Health, Montréal, QC, Canada

4. Center for Neuroscience, University of California, Davis; Davis, 95618, USA.

5. Department of Neurology, School of Medicine, University of California, Davis; Sacramento, 95817, USA.

http://doi.org/10.1101/2024.06.27.601063

Abstract

The NAc, a highly integrative brain region controlling motivated behavior, is thought to receive distinct information from various glutamatergic inputs yet strong evidence of functional specialization of inputs is lacking. While circuit neuroscience commonly seeks specific functions for specific circuits, redundancy can be highly adaptive and is a critical motif in circuit organization. Using dual-site fiber photometry in an operant reward task, we simultaneously recorded from two NAc glutamatergic afferents to assess circuit specialization. We identify a common neural motif that integrates reward history in medial prefrontal cortex (mPFC) and ventral hippocampus (vHip) inputs to NAc. Then, by systematically degrading task complexity, dissociating reward from choice and action, we identify key circuit-specificity in the behavioral conditions that recruit encoding. While mPFC-NAc invariantly encodes reward, vHip-NAc encoding is uniquely anchored to unrewarded outcomes. Ultimately, using optogenetic stimulation we demonstrate that both inputs co-operatively modulate task engagement. We illustrate how similar encoding, with differential gating by behavioral state, supports state-sensitive tuning of reward-motivated behavior.

Introduction

The NAc integrates glutamatergic inputs with dopaminergic input from the ventral tegmental area, with multiple glutamatergic inputs converging at the level of individual medium spiny neurons in the NAc medial shell (Britt et al., 2012; Carter et al., 2007; Christoffel et al., 2021; Floresco, 2015; French & Totterdell, 2002; Lind et al., 2023; Muir et al., 2024; O'Donnell & Grace, 1995). Prominent theoretical perspectives hold that these inputs send qualitatively distinct information which the NAc then integrates to orchestrate motivated behavior (Floresco, 2015; Grace et al., 2007; Lind et al., 2023; Parker et al., 2022). For example, the mPFC contributes information about rewarding events and executive control while the vHip contributes emotional context and behavioral inhibition (Bagot et al., 2015; Barker et al., 2019; Hamel et al., 2022; Lindenbach et al., 2022; Muir et al., 2020; Otis et al., 2017; Parker et al., 2022; Spellman et al., 2021; Wenzel et al., 2023; Yoshida et al., 2020). Despite predictions of distinct encoding and behavioral function for mPFC-NAc and vHip-NAc, strong evidence of functional specialization is lacking. To date, most studies have examined a single input and the few studies that examined one or more inputs in the same task compared across animals leaving open the possibility that inter-individual variation in behavior and other variables influence neural encoding (Britt et al., 2012; G. Chen et al., 2023; Reed et al., 2018).

To systematically interrogate functional redundancy versus specialization we simultaneously probed neural encoding using dual-site *in vivo* fiber photometry to record activity in two glutamatergic circuits during reward-guided choice in a two-armed bandit task. The mPFC-NAc is widely appreciated to mediate reward processing and, given

that vHip-NAc inputs converge with mPFC-NAc, we asked if the vHip-NAc might also contribute to this function (Otis et al., 2017; Parker et al., 2022; Spellman et al., 2021). Using trial-by-trial modeling of neural activity, we identify a novel mechanism for integrating outcome information across trials that is common to both circuits. Analyzing the redundancy across signals revealed an additional dimension of uniqueness to vHip-NAc encoding. By sequentially degrading task complexity we show that, despite sharing a common mechanism for outcome integration, each circuit is recruited in distinct behavioral states, with the vHip-NAc preferentially encoding reward after unrewarded outcomes. Optogenetically manipulating circuit-specific activity revealed that, once recruited, both inputs cumulatively mediate dynamic behavioral engagement. Our findings reveal co-operative circuit organization in NAc wherein redundant encoding in two inputs is gated by circuit-specific mechanisms for state-sensitive tuning of reward-motivated behavior.

Results

mPFC-NAc and vHip-NAc similarly encode outcomes in a probabilistically rewarded environment

To assess redundancy versus specificity in outcome encoding in two distinct circuits under matched conditions and trial histories, we injected retrograding AAV-GCaMP7f in NAc medial shell and implanted optic fibers in mPFC and vHip to record Ca^{2+}-associated fluorescence while mice engaged in reward-guided choice (Fig. 1E). We trained mice in a two-lever probabilistic reward learning task (i.e. a two-armed bandit task) in which lever pressing probabilistically earns a chocolate milk reward (Fig.

1A). Following each lever press, one of two different auditory cues signaled trial outcome (rewarded, unrewarded) and start of the inter-trial interval (ITI). To maintain a dynamic environment with robustly encountered rewarded and unrewarded outcomes, levers were probabilistically rewarded on 80% or 20% of presses with probabilities switched after five consecutive responses on the high probability lever. Female (n=10) and male (n=12) mice experienced similarly high numbers of unrewarded and rewarded trials and low numbers of omission trials (Fig. 1B-D). Examining behavior across sessions shows decreasing staying probability after unrewarded outcomes and increasing rewards earned, indicating animals use information about outcomes to guide behavior (Fig. S1).

Trial based tasks are ideal for probing neural encoding, generating large numbers of trials. However, standard analysis approaches either analyze individual trials, failing to account for the within animal nested data structure and inappropriately inflating effects, or average all trials within animals, thereby underestimating effects. Choice tasks are additionally challenging with the number of instances of each trial type varying across animals. To preserve the power of trial-by-trial data while accounting for the nested structure and unbalanced observations we used a linear mixed model approach (Yu et al., 2022). To examine how outcome is encoded in each projection, we modeled normalized Ca^{2+}-associated fluorescence change as a function of trial outcome, while controlling for inter-individual variability.

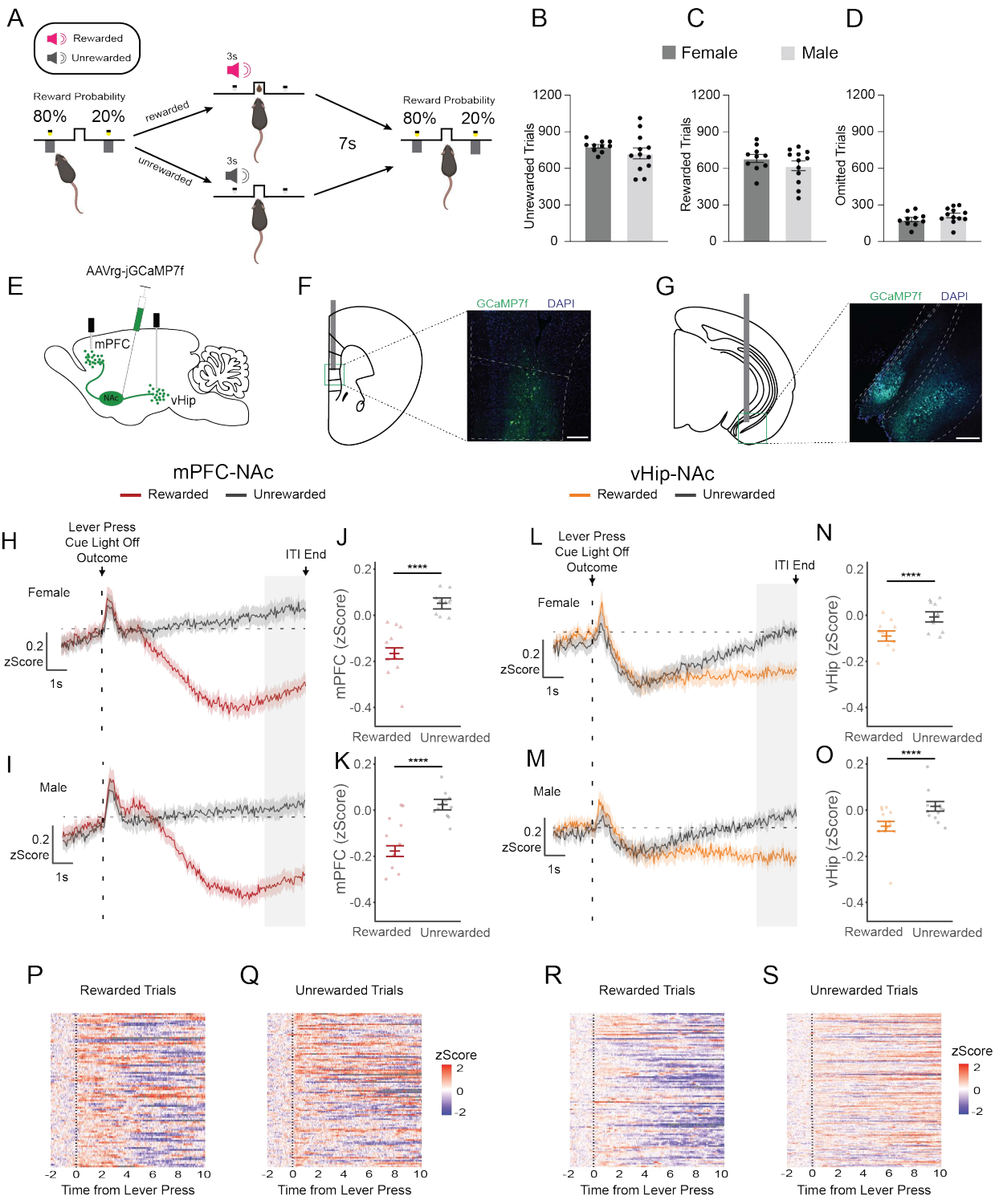


Figure 1. mPFC-NAc and vHip-NAc similarly encode reward in a probabilistically rewarded environment. *(A) Schematic of two-armed bandit task. Mice lever press in a two lever task in which one lever is rewarded with chocolate milk on 80% of trials, and the other on 20%. Following a lever press, levers retract, and auditory cues signal*

*outcome and start of a 10 sec inter-trial interval (ITI). Contingencies switch after five consecutive responses on the high probability lever. Female (n=10) and male (n=12) mice robustly engage with the task, experiencing similar numbers of (B) unrewarded (C) rewarded and (D) omission trials. (E) Retrograding jGCaMP7f is injected into the nucleus accumbens (NAc) medial shell and optic fibers implanted in medial prefrontal cortex (mPFC) and ventral hippocampus (vHip) to simultaneously probe neural activity indicated by Ca^{2+}-associated fluorescence changes in (F) mPFC neurons projecting to NAc (mPFC-NAc) and (G) vHip neurons projecting to NAc (vHip-NAc) as mice encounter reward and non-reward. Estimated mean mPFC-NAc activity across all rewarded and unrewarded trials in (H) female (n=10) and (I) male (n=12) mice. y=0 is indicated by a dashed horizontal line. Analysis focused on 8-10 sec after lever press (ITI end). At ITI end, mPFC-NAc activity is suppressed by rewarded outcomes in female (J; Z=21.348, p<0.0001) and male (K; Z=19.625, p<0.0001) mice. Estimated mean vHip-NAc activity across all rewarded and unrewarded trials in (L) female and (M) male mice. At ITI end, vHip-NAc activity is suppressed by rewarded outcomes in female (N; Z=8.161; p<0.0001) and male (O; Z=8.924; p<0.0001) mice. Heatmap of mPFC-NAc activity to (P) rewarded outcomes and (Q) unrewarded outcomes in a representative animal across one session. Heatmap of vHip-NAc activity to (R) rewarded outcomes and (S) unrewarded outcomes in a representative animal across one session. Error bars represent SEM around the estimated mean. ****p<0.0001*

Reward strongly suppressed mPFC-NAc and vHip-NAc activity in female and male mice. In mPFC-NAc, a peak following the lever press and outcome delivery is followed by gradually emerging reward-associated suppression across the ITI (Fig 1H, I, P,Q). In vHip-NAc, an initial peak is followed by suppression after the lever press and outcome delivery, with suppression sustained following reward or activity gradually increasing following unrewarded outcomes (Fig. 1 L, M, R,S). We focused analysis on the end of the ITI (8-10 sec after lever press) when trial outcome has been integrated prior to next trial start. By ITI end, reward robustly suppressed mPFC-NAc activity in female and male mice (Fig. 1J, K). Reward also robustly suppressed vHip-NAc activity in female and male mice (Fig. 1N, O). This indicates that outcome encoding emerges across the ITI with reward suppressing mPFC-NAc and vHip-NAc activity. To explore

modulation by other task factors, we examined neural encoding time-locked to licking and decision-relevant behaviors. We did not observe clear neural encoding of licking (Fig. S2A,D), the decision to stay or shift (Fig. S2B,E), or the identity of the chosen lever (Fig. S2C,F) suggesting that outcome is the primary source of modulation in mPFC-NAc and vHip-NAc in this task.

Having observed that mPFC-NAc and vHip-NAc are similarly modulated by reward, we then examined if one circuit leads the other. We found that the time-lag for the maximum cross-correlation between mPFC-NAc and vHip-NAc did not significantly differ from zero in rewarded or unrewarded trials in either sex. This shows that neither circuit drives outcome encoding in the other (Fig. S3A, B). Interestingly, we note that although suppression emerges earlier in vHip-NAc than mPFC-NAc (Fig. S3C), the utility of this suppression in distinguishing rewarded vs unrewarded outcome emerges earlier in mPFC-NAc than vHip-NAc (~ 3 seconds post lever press in mPFC-NAc vs. ~ 4 seconds post lever press in vHip-NAc; Fig. S3D,E). This suggests that while the overall informational encoding is comparable, the underlying dynamics likely vary considerably between pathways.

mPFC-NAc and vHip-NAc integrate reward history

We find that mPFC-NAc and vHip-NAc similarly encode outcomes. Visualizing this encoding across a longer timespan shows that reward-mediated suppression can last across 10s of seconds in mPFC-NAc and vHip-NAc (Figure S4). We thus speculated that this enduring modulation might integrate reward information across successive trials and that this integration might be more prominent in mPFC-NAc than vHip-NAc, given prior evidence of enduring representation in mPFC (Parker et al., 2022;

Spellman et al., 2021; Sul et al., 2010). To test this, we sorted trials by both prior and current outcome, identifying trial sequences that were rewarded then rewarded (R→R), rewarded then unrewarded (R→U), unrewarded then rewarded (U→R), and unrewarded then unrewarded (U→U). We then compared neural activity across the ITI on the most recent trial to determine how prior outcome modulates outcome encoding on the current trial. Analyzing males and females separately revealed similar modulation (Fig. S5) and we therefore report sex-combined analyses. Both previous and current outcome modulate mPFC-NAc activity (Fig. 2A). Following a given trial (t -1), reward suppresses mPFC-NAc activity (Fig. 2B) effectively resetting the baseline for the next trial. Reward on the subsequent trial (t0) similarly suppresses mPFC-NAc by ITI end, regardless of prior outcome. However, when mice are unrewarded on the subsequent trial (t0), suppression of mPFC-NAc by prior reward is maintained through ITI end (Fig. 2C). This suggests that a single reward maximally and enduringly suppresses mPFC-NAc activity and that, in the absence of subsequent reward, this suppression slowly dissipates.

We then examined if vHip-NAc similarly integrates outcomes (Fig. 2D). Following a given trial (t-1), reward suppresses vHip-NAc activity (Fig. 2E). As with mPFC-NAc, this resets the baseline for the next trial (t0), wherein reward suppresses vHip-NAc regardless of prior outcome. However, when the subsequent trial (t0) is unrewarded, suppression of vHip-NAc activity by prior reward is maintained through ITI end (Fig. 2F). Together, this shows that mPFC-NAc and vHip-NAc similarly integrate outcomes across trials. In both circuits, reward maximally suppresses neural activity and activity gradually increases following subsequent unrewarded outcomes, such that, by ITI end, the relative degree of suppression represents an integrated reward outcome history.

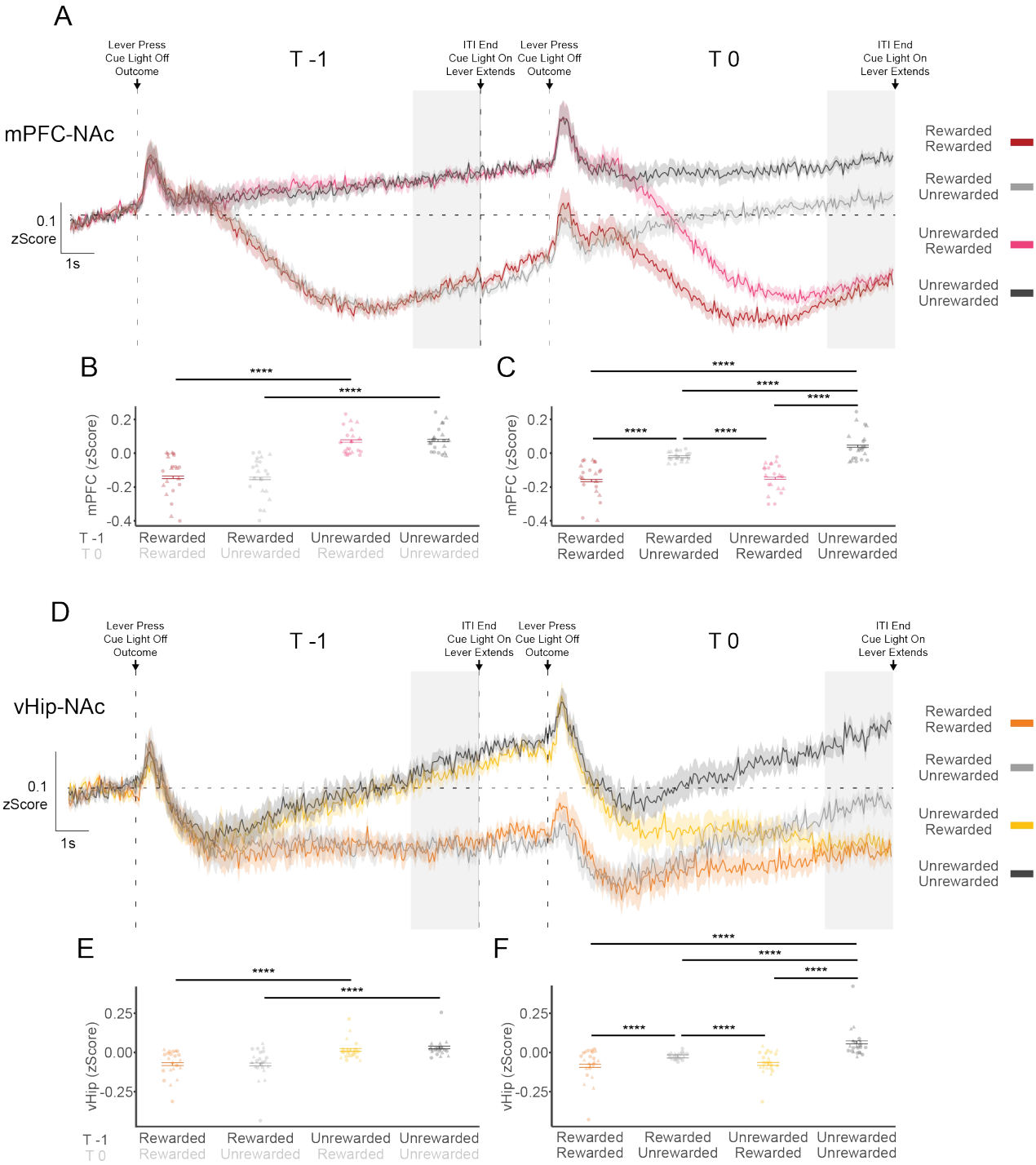


Figure 2. mPFC-NAc and vHip-NAc similarly integrate reward history. *(A) Estimated mean mPFC-NAc activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs in female (n=10) and male (n=12) mice. y=0 is indicated by a dashed horizontal line. Analysis focused on 8-10 sec after lever press (ITI end). (B) On trial t-1, mPFC-NAc activity is significantly suppressed by reward (U→U vs R→U: Z=28.99496, p<0.0001; U→R vs R→R: Z=25.6767,*

*p<0.0001). (C) On the subsequent trial, t0, mPFC-NAc activity is significantly suppressed by current reward (U→U vs U→R: Z=-28.5098, p<0.0001; R→U vs R→R: Z=-19.8981, p<0.0001, U→U vs R→R: Z=-29.0153, p<0.0001;). When trial t0 is unrewarded, mPFC-NAc activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 9.1965, p<0.0001; U→R vs R→R: Z= 1.9308, p=0.2811; R→U vs U→R: Z= -18.7786, p<0.0001). (D) Estimated mean vHip-NAc activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. y=0 is indicated by a dashed horizontal line. (E) On trial t-1, vHip-NAc activity is significantly suppressed by reward (U→U vs R→U: Z=14.9372, p<0.0001; U→R vs R→R: Z=11.6962, p<0.0001). (F) On the subsequent trial, t0, vHip-NAc activity is significantly suppressed by current reward (U→U vs U→R: Z=-17.4993, p<0.0001; R→U vs R→R: Z=-7.1005, p<0.0001; U→U vs R→R: Z=-18.0126, p<0.0001). When trial t0 is unrewarded, vHip-NAc activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 11.4112, p<0.0001; U→R vs R→R: Z= 1.4235, p=0.6349; R→U vs U→R: Z= 5.9394, p=0.6349). Individual-animal averages are indicated by circles for males and triangles for females. Error bars represent SEM around the estimated mean. ****p<0.0001*

mPFC-NAc and vHip-NAc are differentially sensitive to unrewarded outcomes

Analyzing neural encoding of reward and outcome integration revealed that mPFC-NAc and vHip-NAc similarly encode reward suggesting they may provide redundant information to the NAc. To test redundancy between mPFC-NAc and vHip-NAc we calculated the conditional entropy of mPFC-NAc given vHip-NAc (H(mPFC-NAc|vHip-NAc)) and vHip-NAc given mPFC-NAc (H(vHip-NAc|mPFC-NAc)). In this way, we assessed the information contributed by each circuit beyond that contributed by the other at ITI end, when outcome is fully integrated (Fig. 3A). We contrasted entropy between rewarded and unrewarded outcomes as a function of prior outcome. Relative to unrewarded outcomes, the entropy of mPFC-NAc given vHip-NAc was reduced by

rewarded outcomes, indicating that vHip-NAc and mPFC-NAc signals are more redundant after reward than non-reward (Fig. 3B). In contrast, following previous unreward, but not previous reward, current reward increased the entropy of vHip-NAc given mPFC-NAc (Fig. 3C), indicating that, under these conditions, mPFC-NAc explains less of the vHip-NAc signal. This shows that, after reward, vHip-NAc and mPFC-NAc encoding converges, becoming more redundant, but when reward is made more surprising by immediately following an unrewarded outcome, vHip-NAc carries additional information. That is, despite global redundancy in reward encoding motifs, we identify a dimension of circuit specificity and a potential unique role for vHip-NAc in encoding reward following unrewarded outcomes.

If this is true, across outcome histories, vHip-NAc encoding should be most apparent when reward follows an unrewarded outcome, whereas, following consecutive rewards, vHip-NAc should become insensitive to outcome as rewards become less surprising. In contrast, mPFC-NAc encoding is predicted to be relatively invariant across outcome histories. To test this, we examined current outcome encoding at ITI end while considering prior outcomes up to three trials back. Consistent with our prediction, mPFC-NAc encoded current outcome regardless of prior outcome history (Fig. 3D; Supplementary Table 1) while vHip-NAc failed to encode current outcome after two or more consecutive rewards (Fig. 3E; Supplementary Table 1). This effect seems to be mostly mediated by differences in how mPFC-NAc and vHip-NAc respond to unrewarded outcomes. While reward continues to suppress activity in both pathways regardless of reward history, when encountering an unrewarded outcome following

several rewarded outcomes, mPFC-NAc activity increases as expected but vHip-NAc activity fails to immediately increase.

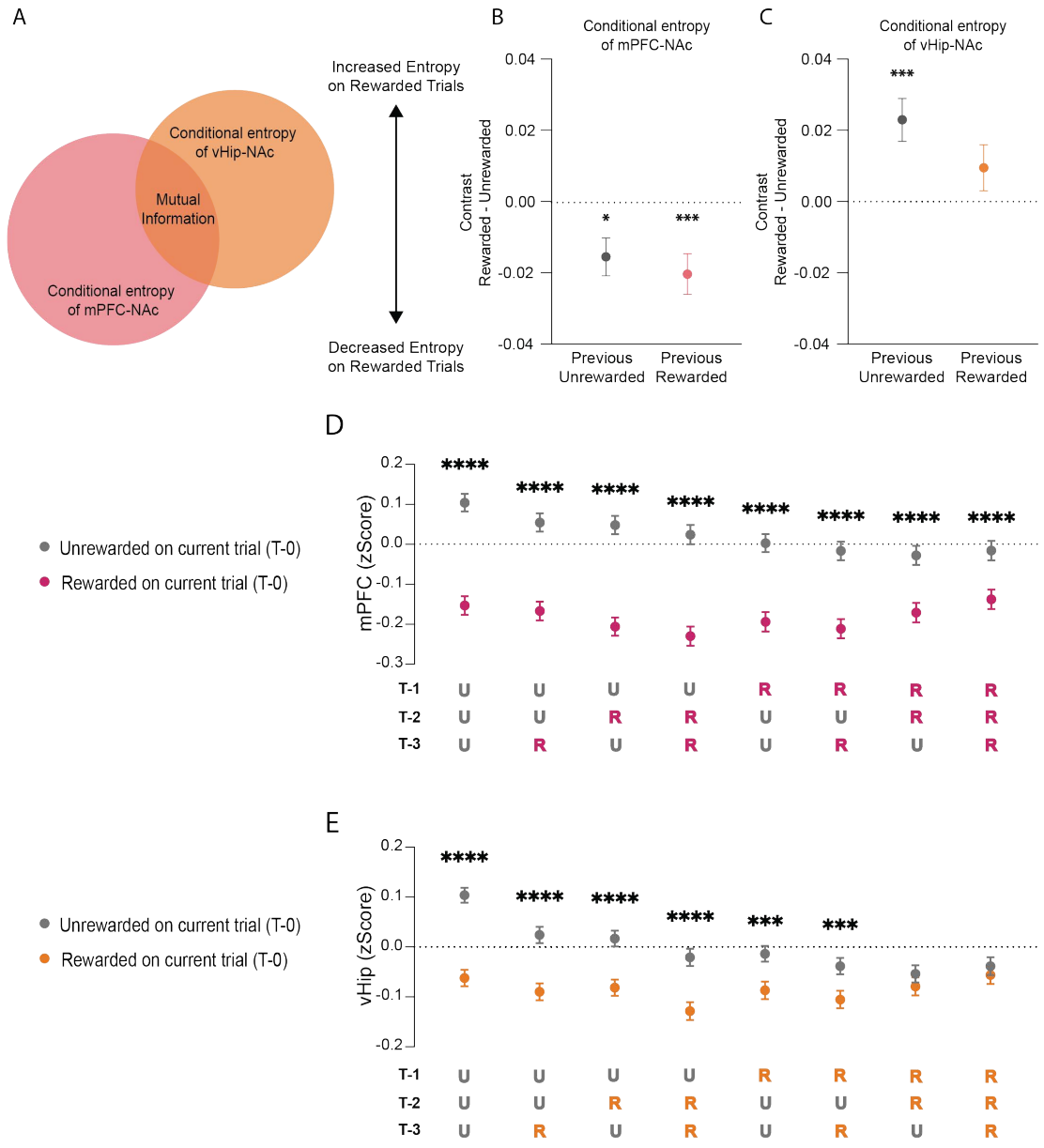


Figure 3. mPFC-NAc and vHip-NAc are differentially sensitive to unrewarded outcomes. *(A) Venn diagram representing the relationship between the mutual information and conditional entropy that exists between observed mPFC-NAc and vHip-NAc signals. Conditional entropy is a measure of the additional unique information contributed by a second signal given fully knowledge of a first signal. (B) Conditional entropy in mPFC-NAc is reduced on rewarded relative to unrewarded trials regardless*

*of previous outcome (U→U vs U→R: Z=2.8644, p<0.0001; R→U vs R→R: Z=3.5185, p<0.0001) indicating that less unique information is carried in mPFC-NAc after reward. (C) Conditional entropy in vHip-NAc is increased on rewarded relative to unrewarded trials only when the prior outcome was unrewarded (U→U vs U→R: Z=-3.7566, p=0.0003) indicating that more unique information is carried in vHip-NAc when reward follows nonreward. Comparison of activity at ITI end on currently rewarded or unrewarded trials considering prior outcome history up to three trials back shows that (D) mPFC-NAc activity is suppressed on every currently rewarded trial indicating that mPFC-NAc consistently encodes current outcome via relative suppression regardless of outcome history. In contrast, (E) vHip-NAc activity is suppressed on currently rewarded trials except when current reward is preceded by two (Z=1.2310, p=0.8606) or three (Z=0.8398, p=0.9834) prior consecutive rewards indicating that vHip-NAc ceases to encode current outcome via relative suppression after consistent reward. See supplementary table 1 for all comparisons. Error bars represent SEM around the estimated mean. *p<0.05, ***p<0.001, ****p<0.0001*

Degrading task requirements reveals circuit-specific roles in reward integration

Analyzing informational redundancy and encoding across varying outcome histories suggested that, while mPFC-NAc and vHip-NAc encode and integrate reward via a common mechanism, each may nevertheless serve distinct functions in reward processing. To isolate the specific conditions under which each circuit integrates outcomes we recorded neural activity while degrading task requirements to sequentially eliminate choice and action. We first eliminated choice, extending only a single lever while maintaining the requirement to press to elicit an outcome. To hold outcome experience constant, the specific sequence of reward and unreward was yoked to each animal's prior performance on the two-lever task (Fig. 4A). In the absence of choice, mPFC-NAc continued to encode previous and current outcome (Fig. 4B). On trial t0, by ITI end, current and prior outcomes were encoded, as in the two-lever task (Fig. 4C,

Fig. 2C). Examining vHip-NAc in the one-lever task also revealed largely similar outcome-mediated modulation (Fig. 4D, Fig. 2D). At ITI end, prior and current outcomes were integrated, similar to the two-lever task (Fig. 4E, Fig. 2F). Despite conserved information encoding in both circuits, the shape of the vHip-NAc signal was more visibly altered than the mPFC-NAc. In particular, the vHip-NAc signal in the one-lever task appeared noisier and blunted with the expected peak following lever press largely absent, potentially suggesting heightened sensitivity to task structure. Overall, we find that both mPFC-NAc and vHip-NAc maintain similar graded representations of reward history that are largely independent of choice requirements.

Removing lever choice minimally impacted reward integration. We then asked if neural integration of outcome history is entirely independent of response requirements by removing both levers in a choice-free response-free task. Trials continued to be signaled by cue-lights, but without lever extension and outcomes were passively delivered yoked to each animal's individual performance on the full two-lever task (Fig. 4F). Eliminating the response requirement markedly and distinctly altered reward integration in both circuits. In mPFC-NAc (Fig. 4G), encoding of prior outcome was erased and only the current outcome encoded (Fig. 4H). This differs from both the two-lever and one-lever tasks wherein mPFC-NAc encoded a graded representation of reward history and suggests that mPFC-NAc integrates reward history only in instrumental settings where a response elicits outcomes. However, even when rewards are passively encountered (i.e. when no lever press is required), mPFC-NAc continues to encode reward but with a shortened time constant, such that only the most recent outcome is retained.

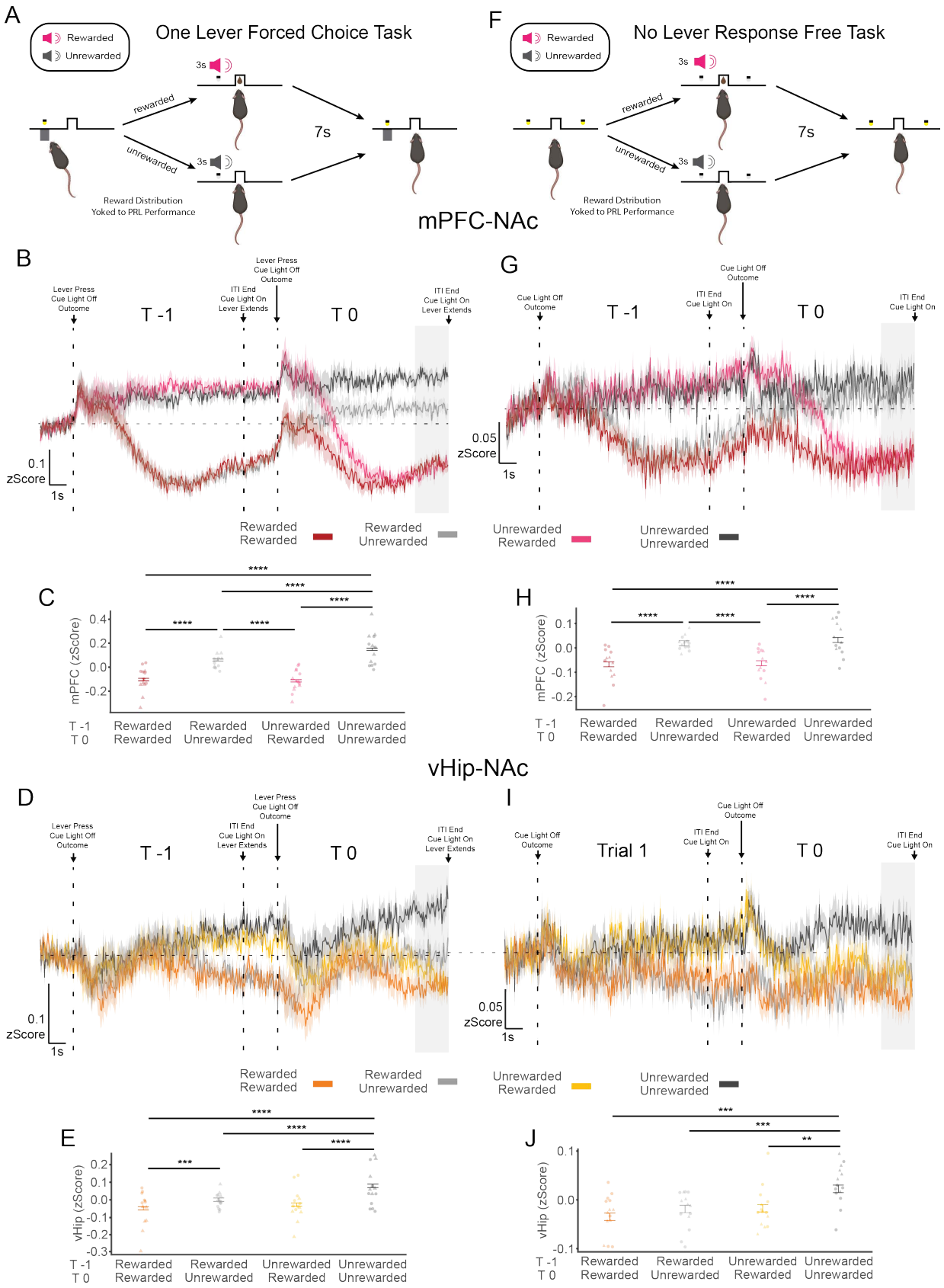

A
Rewarded
Unrewarded
One Lever Forced Choice Task
3s
rewarded
unrewarded
3s
7s
Reward Distribution
Yoked to PRL Performance
F
Rewarded
Unrewarded
No Lever Response Free Task
3s
rewarded
unrewarded
3s
7s
Reward Dstribution
Yoked to PRL Performance
mPFC-NAc
B
Lever Press
Cue Light Off
Outcome
ITI End
Cue Light On
Lever Extends
T -1
T 0
0.1
zScore
1s
Rewarded
Rewarded
Rewarded
Unrewarded
G
Cue Light Off
Outcome
ITI End
Cue Light On
T -1
T 0
0.05
zScore
1s
Unrewarded
Rewarded
Unrewarded
Unrewarded
C
mPFC (zSc0re)
T -1
T 0
Rewarded
Rewarded
Rewarded
Unrewarded
Unrewarded
Rewarded
Unrewarded
Unrewarded
H
mPFC (zScore)
T -1
T 0
Rewarded
Rewarded
Rewarded
Unrewarded
Unrewarded
Rewarded
Unrewarded
Unrewarded
vHip-NAc
D
Lever Press
Cue Light Off
Outcome
ITI End
Cue Light On
Lever Extends
T -1
T 0
0.1
zScore
1s
Rewarded
Rewarded
Rewarded
Unrewarded
I
Cue Light Off
Outcome
ITI End
Cue Light On
Trial 1
T 0
0.05
zScore
1s
Unrewarded
Rewarded
Unrewarded
Unrewarded
E
vHip (zScore)
T -1
T 0
Rewarded
Rewarded
Rewarded
Unrewarded
Unrewarded
Rewarded
Unrewarded
Unrewarded
J
vHip (zScore)
T -1
T 0
Rewarded
Rewarded
Rewarded
Unrewarded
Unrewarded
Rewarded
Unrewarded
Unrewarded

Figure 4. Degrading task requirements reveals circuit specialization in integrating reward history. *(A) Schematic of one-lever forced choice task in which lever presses are rewarded on a schedule yoked to each animal's individual performance in the final three days of the two-armed bandit task. Following a lever press, levers retract, and auditory cues signal outcome and start of a 10 sec inter-trial interval (ITI). (B) Estimated mean mPFC-NAc activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded unrewarded (U→U) trial pairs (male n= 8, female n=6). y=0 is indicated by a dashed horizontal line. Analysis focused on 8-10 sec after lever press (ITI end). (C) On trial t0, mPFC-NAc activity is suppressed by reward (U→U vs U→R: Z=18.8757, p<0.0001; R→U vs R→R: Z=12.0687, p<0.0001; U→U vs R→R: Z=18.2004, p<0.0001). When trial t0 is unrewarded, mPFC-NAc activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 6.3467, p<0.0001; U→R vs R→R: Z= -0.7826, p=0.9671; U→R vs R→U: Z= 12.7865, p<0.0001). (D) Estimated mean vHip-NAc activity across pairs of consecutive trials (t-1→t0) showing R→R, R→U, U→R, and U→U trial pairs (male n= 8, female n=6). y=0 is indicated by a dashed horizontal line. (E) On trial t0, vHip-NAc activity is suppressed by reward (U→U vs U→R: Z=8.5245, p<0.0001; R→U vs R→R: Z=4.0519, p=0.0003; U→U vs R→R: Z=10.2097, p<0.0001). When trial t0 is unrewarded, mPFC-NAc activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 6.2425, p<0.0001; U→R vs R→R: Z= 1.7019, p=0.4275; U→R vs R→U: Z= 2.3408, p=0.1101). (F) Schematic of no lever response free task. Mice are allowed to collect rewards delivered on a schedule yoked to each animal's individual trial statistics (latency and outcome) of the two-armed bandit task. Trial structure is signaled by cue-light illumination and after a predetermined delay auditory cues signal outcome and start of a 10 sec ITI. (G) Estimated mean mPFC-NAc activity across pairs of consecutive trials (t-1→t0) showing R→R, R→U, U→R, and U→U trial pairs (male n= 8, female n=6). y=0 is indicated by a dashed horizontal line. (H) On trial t0, mPFC-NAc activity is suppressed by reward (U→U vs U→R: Z=8.2136, p<0.0001; R→U vs R→R: Z=7.4647, p<0.0001; U→U vs R→R: Z=8.5242, p<0.0001; U→U vs R→U: Z= 1.1662, p=0.8126; U→R vs R→R: Z= 0.3493, p=0.9996; U→R vs R→U: Z= 7.1124, p<0.0001). (I) Estimated mean vHip-NAc activity across pairs of consecutive trials (t-1→t0) showing R→R, R→U, U→R, and U→U trial pairs (male n= 8, female n=6). y=0 is indicated by a dashed horizontal line. (J) On trial t0, vHip-NAc activity is suppressed by reward only if trial t-1 was unrewarded (U→U vs U→R: Z=3.7413, p=0.0011; R→U vs R→R: Z=1.5289, p=0.5551; U→U vs R→R: Z=5.3913, p<0.0001). When trial t0 is unrewarded, vHip-NAc activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 3.8661, p=0.0007; U→R vs R→R: Z= 1.6584, p=0.4587; U→R vs R→U: Z= -0.1282, p=0.9999). Individual-animal averages are indicated by circles for*

*males and triangles for females. Error bars represent SEM around the estimated mean. ***p<0.001, ****p<0.0001*

vHip-NAc representation of reward history was also degraded, yet in a distinct manner (Fig. 4I). Current outcomes were encoded only when the previous trial, t-1, was unrewarded (Fig. 4J). This shift in encoding translates into vHip-NAc effectively overlooking isolated instances of non-reward, likely reflecting an extended time constant. Critically, this cannot be explained by changes in task engagement given that mPFC-NAc continued to represent reward in these same animals (Fig. 4G,H) and licking bouts were similarly maintained across task variants (Fig. S6). Following the removal of response requirements, we returned animals to the two-lever task and again observed encoding of integrated reward history (Fig. S7) confirming that the modulation of encoding across task degradation is indeed attributable to altered task requirements and is not artifactual (e.g. potential signal degradation over time). This reveals that task demands differently shape neural encoding of reward in mPFC-NAc and vHip-NAc. When reward is passively encountered, independent of a required response, mPFC-NAc maintains a simplified reward representation across a shortened temporal window, limiting integration across trials. In contrast, vHip-NAc anchors encoding to unrewarded outcomes with an extended time constant, to preferentially represent surprising rewards. This suggests that while the fundamental function of mPFC-NAc in rewarding contexts is to encode outcomes, the fundamental function of vHip-NAc is to use information about unrewarded outcomes to tune outcome encoding.

mPFC-NAc and vHip-NAc modulate task engagement

Examining neural representation of outcomes identified both mechanistic redundancy and functional specificity in mPFC-NAc and vHip-NAc encoding. We then asked how this neural processing might integrate to modulate behavior. While in general encoding was similar in both circuits, reducing the requirement for engagement by making reward non-contingent revealed functional specialization. We hypothesized that outcome-associated neural activity in mPFC-NAc and vHip-NAc modulates task engagement. To test this, we examined if neural activity at ITI end predicted latency to lever press on the subsequent trial, a metric operationalizing engagement (Bari et al., 2019; Beierholm et al., 2013; Cox et al., 2023; Hamid et al., 2016; Niv et al., 2007). A linear mixed effects model revealed modest yet significant relationships between latency to lever press and mPFC-NAc, vHip-NAc, and the interaction of mPFC-NAc and vHip-NAc activity (Fig. 5A, Supplementary Table 2; Fig. S8). This suggests that increased activity during outcome integration in either circuit increases latency to lever press, indicating reduced behavioral engagement (Fig. S9).

From the association between neural activity and latency, we hypothesized that reward suppresses activity in mPFC-NAc and vHip-NAc to support behavioral engagement, defining a mechanism whereby recent reward history modulates engagement in reward-motivated behavior. We predicted that acutely increasing activity in either mPFC-NAc or vHip-NAc would suppress engagement. To test this, we injected retrograding AAV-ChR2 into NAc and implanted fibers above mPFC and vHip to deliver blue light stimulation during the ITI on a subset of trials in the two-armed bandit task (Fig. 5B-E; Fig. S10). To test if mPFC-NAc and vHip-NAc uniquely or redundantly

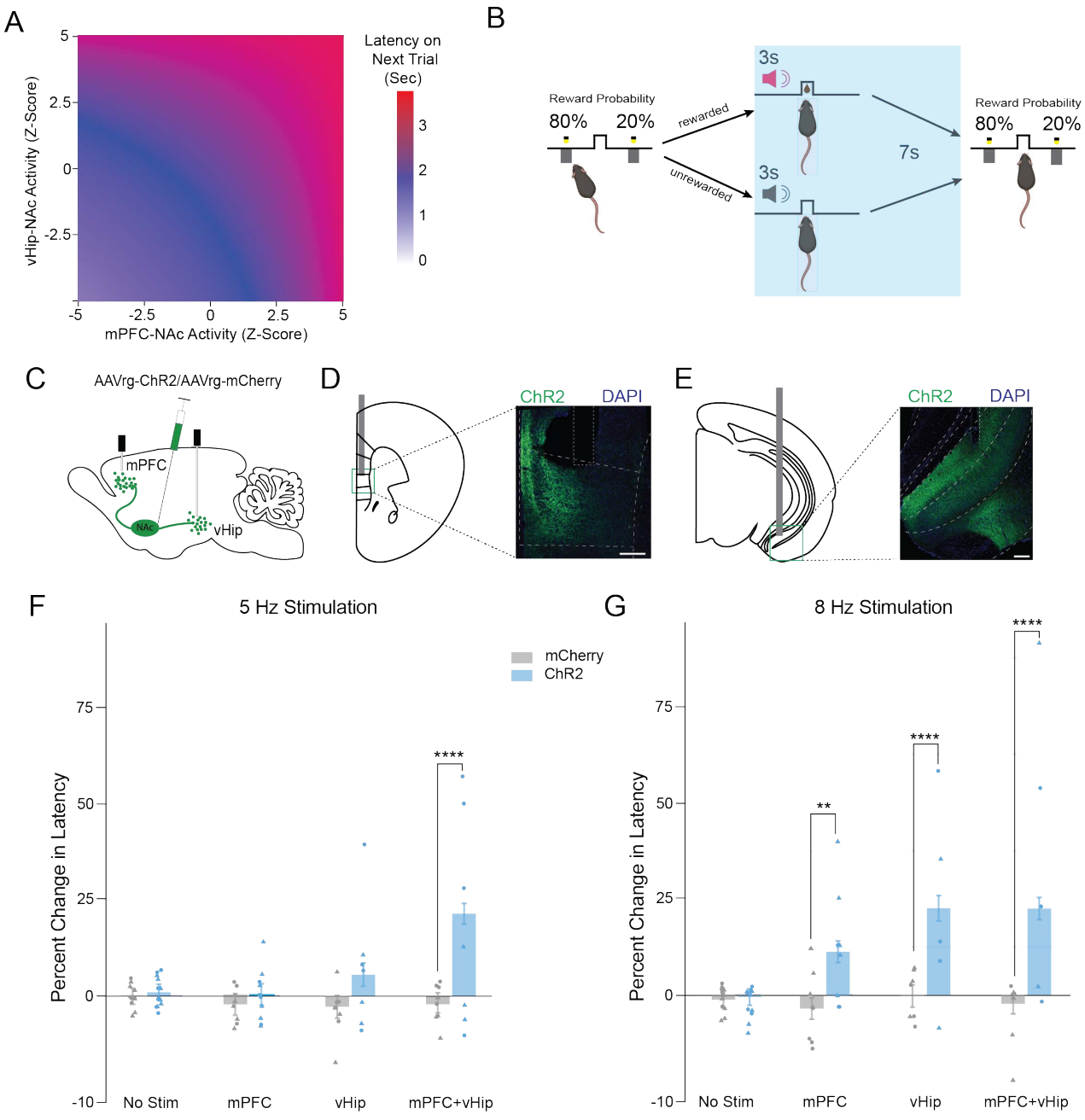


Figure 5. mPFC-NAc and vHip-NAc modulate task engagement. *(A) Heatmap of estimated latency to respond on the subsequent trial given mPFC-NAc and vHip-NAc activity at ITI end shows that increased activity associates with longer latency. (B) Optogenetic stimulation in the two-armed bandit task is delivered for the duration of the ITI to either mPFC-NAc, vHip-NAc, or simultaneously to both circuits. (C) AAVrg-ChR2-mCherry or AAVrg-mCherry is injected into the NAc and optic fibers implanted in mPFC and vHip to stimulate (D) mPFC-NAc neurons and (E) vHip-NAc neurons. (F) Simultaneous 5Hz stimulation of mPFC-NAc and vHip-NAc, but neither circuit individually, increased latency to respond in ChR2 animals (male n=6, female n=7) compared to mCherry controls (male n=6, female n=6; Z=-18.6984, p<0.0001). (G) 8 Hz stimulation of mPFC-NAc (Z=-12.6970, p=0.01354), vHip-NAc (Z=-23.8073, p<0.0001),*

*and simultaneous stimulation of both mPFC-NAc and vHip-NAc (Z=-24.1357, p<0.0001) all increased latency in ChR2 animals (male n= 5, female n=6) compared to mCherry controls (male n= 6, female n=6). Individual-animal averages are indicated by circles for males and triangles for females. Error bars represent SEM around the estimated mean. **p<0.01,****p<0.0001*

control behavior we stimulated each circuit alone or both simultaneously. Stimulating either circuit alone had no effect, whereas stimulating both simultaneously increased latency to lever press, but did not alter choice behavior (Fig. 5F; Fig. S11A). This could indicate either a threshold for sufficient cumulative glutamatergic drive or a requirement for synergistic interaction between inputs. To differentiate these possibilities, we repeated the experiment with stronger stimulation. Strong stimulation of either circuit alone increased latency to lever press, again with no effect on choice (Fig. 5G, Fig. S11B). This shows that total glutamatergic input modulates engagement, independent of input identity. mPFC-NAc stimulation yielded a slightly weaker effect than vHip-NAc, consistent with previous findings that mPFC projections to NAc medial shell are sparser than those from vHip (Britt et al., 2012). Stimulation during lever presentation did not yield any changes in latency or choice behavior, supporting the importance of neural integration of outcome during the ITI period, prior to action initiation (Fig. S12). Together our results demonstrate that mPFC-NAc and vHip-NAc dynamically track outcome information to modulate behavioral engagement according to recent history of reward. While each circuit is specialized to execute this function under distinct behavioral states, once engaged, they redundantly modulate behavior pointing to complementary roles in control of reward seeking.

Discussion

We examined redundancy and specificity in the function of two distinct glutamatergic inputs to the NAc. Using dual-site fiber photometry to probe trial-by-trial outcome encoding simultaneously in two circuits in the same animal during reward-guided choice, we find that mPFC-NAc and vHip-NAc similarly integrate reward via suppression of neural activity. By then systematically manipulating the conditions in which outcomes are encountered, we revealed that each circuit executes this common function under distinct behavioral states. While the mPFC-NAc invariantly encodes outcome, vHip-NAc uses information about unrewarded outcomes to tune outcome encoding, effectively amplifying surprising reward. By comparing independent or synchronous circuit-specific optogenetic stimulation we show that, once engaged, these circuits cooperatively execute a shared function, i.e. modulating task engagement. Taken together, we identify a redundant mechanism for outcome integration with circuit-specific gating. This supports convergence of multiple inputs in tuning behavioral engagement to recent history of reward.

Our finding that both mPFC-NAc and vHip-NAc integrate information about outcomes of reward-motivated actions is consistent with the well-established role of mPFC in reward processing. Critically, we demonstrate that this function is not specific or limited to the mPFC-NAc. Globally, the mPFC encodes information about previous actions and outcomes (Sul et al., 2010) and mPFC projections to the NAc bridge information about current actions and outcomes across trials (Parker et al., 2022; Spellman et al., 2021). Our findings suggest these functions are not unique to mPFC-NAc and are shared by vHip-NAc. However, we identify novel state dependent

specialization in how reward integration is engaged in each circuit. We show that the mPFC-NAc fundamentally functions as a reward ledger, with reward suppressing neural activity no matter the behavioral state. In contrast, we find that vHip-NAc is tuned to preferentially encode outcome information after unrewarded outcomes.

Differential encoding between mPFC-NAc and vHip-NAc emerged upon degrading task requirements, a manipulation that minimizes cognitive and behavioral demands effectively reducing the behavioral utility of representing integrated reward history. Under these circumstances, the base functionality of each circuit is revealed: mPFC-NAc encoding is anchored to reward whereas vHip-NAc is anchored to unrewarded outcomes. Layered on top of this base functionality, representation of reward history scales with task complexity in support of behavioral demands. When reward is passively encountered with limited utility for action-outcome associations, mPFC-NAc encoding is limited to the most recent outcome. In more complex environments wherein actions elicit reward and action-outcome associations have high utility, the mPFC-NAc encoding window extends to integrate reward history. In simpler task structures that no longer require active engagement with a lever to earn rewards, the time-constant of vHip-NAc encoding shifts such that activity no longer increases when a single unrewarded outcome follows reward. As a result, the vHip-NAc effectively comes to encode consecutive loss against other outcomes. Together, this suggests a role for vHip-NAc in providing information about the state of reward statistics in the environment, modulating behavior as a function of unrewarded outcomes, and revealing a novel role for this circuit as a parallel and distinct stream of outcome integration.

The NAc has long been implicated in reward processing yet the precise neural circuit mechanisms are still being resolved. In the NAc medial shell, reward predominantly suppresses neural activity (Chen et al., 2023). This suppression likely maintains reward seeking as stimulation of either D1 or D2 medium spiny neurons bidirectionally controls reward seeking behavior (Lafferty et al., 2020). Here we show that reward suppresses both mPFC-NAc and vHip-NAc, two major excitatory inputs to NAc medial shell. Reward-associated suppression of these inputs would lead to reduced NAc activity. As such, our findings are consistent with reports that optogenetic stimulation of diverse glutamatergic inputs inhibits motivated behavior and the idea that glutamatergic input to NAc medial shell functions as a brake on motivated behavior (Lafferty et al., 2020; Millan et al., 2017; Reed et al., 2018; Yoshida et al., 2019, 2021). We show that outcome integration in mPFC-NAc and vHip-NAc initiates parallel, temporally integrated, neural signaling that may engage this ‘brake’ to align ongoing behavior with recent reward history and so tune behavioral engagement to prevailing environmental conditions.

Employing a redundant mechanism in mPFC-NAc and vHip-NAc may serve several functions. A common mechanism makes for simple integration of multiple inputs and ensures the robustness of the fundamental function of reward-guided engagement against insults. Further, modulating redundant encoding with state-dependent circuit-specific sensitivity may increase the granularity and range of encoding to ultimately amplify the behavioral impact of surprising rewards. We demonstrate that high levels of reward suppress activity in both mPFC-NAc and vHip-NAc to favor continued engagement. In contrast, strong activation of either input suppresses engagement, but,

when weakly activated, synchronous recruitment of both circuits is required. Functionally, this may translate into a mechanism whereby moderate, balanced activity predominantly modulates task engagement while allowing for strong activation of either circuit to exert more direct behavioral control.

Preferential outcome encoding in vHip-NAc after unrewarded outcomes may serve to strengthen engagement in variably rewarding environments, driving increased engagement when reward is infrequently encountered. The sensitivity of vHip-NAc to continuous unrewarded outcomes, may also serve to gauge reward statistics of the environment, continually increasing with each consecutive unrewarded outcome to trigger task disengagement when activity reaches some threshold. Qualitatively, we see hints of this in the shape of the signal after experience with an unrewarded outcome: mPFC-NAc tends to plateau while vHip-NAc continues to increase. Ultimately, dysregulated outcome-encoding in either mPFC-NAc or vHip-NAc could alter behavioral sensitivity to reward. Relative to mPFC-NAc, The vHip-NAc is poised to exert an outsized effect on behavioral engagement both in the strength of its input to NAc medial shell (Britt et al., 2012) and in its role in signaling unrewarded outcomes. For example, hyperactivity of vHip-NAc may erroneously signal a large amount of consecutively unrewarded outcomes, causing premature disengagement. Given our finding that engagement is modulated by the cumulative glutamatergic input to NAc, a sufficiently strong vHip-NAc signal could effectively jam any reward signal from mPFC-NAc, compounding insensitivity to reward that manifests as anhedonia. Indeed, disruption of the balance between NAc inputs and increased vHip-NAc drive is observed following chronic stress (Bagot et al., 2015; Muir et al., 2020; Pignatelli et al., 2021; Williams et

al., 2020), as well as chronic alcohol (Griffin et al., 2023; Kircher et al., 2019) and cocaine intake (Barrientos et al., 2018; Cahill et al., 2016; Pascoli et al., 2014; Zinsmaier et al., 2022), manipulations associated with aberrant reward processing.

Here we examined the simultaneous encoding in two key neural circuits for motivated behavior. By considering outcome encoding within the context of recent outcome history and behavioral demands we identified a common neural mechanism of sustained temporal integration of reward outcomes and reveal how the external environment differentially shapes internal representations within two neural circuits. We also revealed critical circuit specificity: while mPFC-NAc consistently tracks outcomes, vHip-NAc preferentially encodes outcome information after unrewarded outcomes. By illustrating the interplay of redundancy and specificity in circuit control of motivated behavior we demonstrate the need to contextualize events within varied behavioral states to fully understand neural encoding. Overall, our findings point to the importance of balanced suppression of NAc glutamatergic inputs during outcome integration to maintain reward-modulated behavioral engagement.

Methods

Animals

Mice were maintained on a 12-h light-dark cycle (lights on at 7:00AM) at 22-25ºC, group-housed with 3-4 same-sex cage-mates with *ad libitum* access to food and water. All experimental manipulations occurred during the light cycle, in accordance with guidelines of McGill University's Comparative Medicine and Animal Resources Center and approved by the McGill Animal Care Committee. 7-week-old male and female

C57BL/6J mice were obtained from Jackson Laboratories and habituated to the colony room one week prior to start of manipulations. Mice were food restricted to 85% of their free-feeding body weight during experimentation.

Surgeries

Stereotaxic surgery was performed under ketamine (100 mg/kg)/xylazine (10 mg/kg) anesthesia. To achieve projection-specific GCaMP7f expression in glutamatergic NAc-projecting cells, 0.3µl pGP-AAVrg-syn-jGCaMP7f-WPRE virus (1.85× 10^{13}GC/ml; Addgene) was infused into the NAc (A/P: +1.3, M/L: +/-0.60, D/V: -4.9) at a rate of 0.1µl per min, before raising the needle to D/V: -4.7 and infusing a further 0.4µl virus, and allowed to diffuse for 10 min before withdrawing the needle. pGP-AAV-syn-jGCaMP7f-WPRE was a gift from Douglas Kim & GENIE Project (Addgene plasmid # 104488 ; http://n2t.net/addgene:104488 ; RRID:Addgene_104488) (Dana et al., 2019). Chronically implantable optic fibers (Neurophotometrics) with 200µm core and 0.37 NA threaded through ceramic ferrules were implanted above the ventral subiculum of the vHip (A/P: -3.40, M/L: +/-3.00, D/V: -4.75) and infralimbic mPFC (A/P: 1.90, M/L: +/-0.3, D/V: -2.80). Recordings began minimum 4 weeks after surgery to allow sufficient time for stable and robust retrograde virus expression. To achieve projection-specific ChR2 expression in glutamatergic NAc-projecting cells, 0.3µl pGP-AAVrg-hSyn-hChR2(H134R)-EYFP virus (7× 10^{12}GC/ml; Addgene) or a fluorophore only control, pGP-AAVrg-hSyn-mCherry (7× 10^{12}GC/ml; Addgene) was infused into the NAc (A/P: +1.3, M/L: +/-0.60, D/V: -4.9) at a rate of 0.1µl per min, before raising the needle to D/V: -4.7 and infusing a further 0.4µl virus, and allowed to diffuse for 10 min before withdrawing the needle. pAAV-hSyn-hChR2(H134R)-EYFP was a gift from Karl

Deisseroth (Addgene plasmid # 26973; http://n2t.net/addgene:26973 ; RRID:Addgene_26973). pAAV-hSyn-mCherry was a gift from Karl Deisseroth (Addgene plasmid # 114472; http://n2t.net/addgene:114472 ; RRID:Addgene_114472). Chronically implantable optic fibers (Neurophotometrics) with 200µm core and 0.22 NA threaded through ceramic ferrules were implanted above the ventral subiculum of the vHip (A/P: -3.40, M/L: +/-3.00, D/V: -4.75) and infralimbic mPFC (A/P: 1.90, M/L: +/-0.3, D/V: -2.80). Optogenetic manipulations began minimum 4 weeks after surgery to allow sufficient time for stable and robust retrograde virus expression.

Histology

After completion of all behavioral testing, mice were deeply anesthetized with ketamine/xylazine and transcardially perfused with phosphate buffered saline (PBS) and paraformaldehyde (4%). Brains were removed and post-fixed in paraformaldeyhde for 24h and stored in PBS until sectioning on a vibratome (50 µm). Sections were mounted with Vectashield with DAPI (Vector Laboratories) and examined under a fluorescent microscope (Leica DM6000 B) to confirm viral expression and fiber placement. A confocal microscope (Zeiss LSM800) was used to obtain fluorescent images. Images were acquired as tiles with a 20x air objective (NA 0.8) using Zeiss Zen Blue imaging software. Images were collected in the McGill University Advanced BioImaging Facility (ABIF), *RRID:SCR_017697*. Mistargeted animals were excluded from analysis.

Apparatus

Behavioral experiments were performed in standard Med Associates operant boxes (15.24 x 13.34 x 12.7 cm) enclosed in sound attenuating chambers outfitted with

a programmable audio generator, two retractable levers and cue lights either side of a food port for delivering a liquid chocolate milk reward (30µl, Nesquick) diluted with water in a 2:1 ratio. Boxes were controlled and data collected by a computer running MED-PC software (Med-Associates).

Lever Press Training

Training was completed in three stages, with all training sessions lasting 30 minutes. In the first stage, animals were presented with two levers, both of which delivered a chocolate milk reward with a 100% probability. To signal the start of the trial, both levers extended and the cue lights above the levers turned on, animals then had 60 seconds to make a response on either lever. A press on either lever resulted in lever retraction, immediate delivery of a 30 µL chocolate milk reward, and the start of a 3 second auditory cue (2kHz pure tone or white noise). Following either a lever press or 60 seconds with no press (i.e. an omission), a 10 second intertrial interval (ITI) was triggered. After one session with over 25 responses, animals progressed to the second stage. In this stage animals again were presented with two levers but reward was now delivered with a 50% probability on both levers. To signal the start of the trial, both levers extended and the cue lights above the levers turned on, animals then had 60 seconds to make a response. A lever press resulted in lever retraction and immediate delivery of the outcome, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise, counterbalanced across animals) or just a 3 second auditory cue (white noise or 2kHz pure tone). Following either a lever press or omission, a 10 second intertrial interval (ITI) was triggered. Following two consecutive sessions with over 40 responses, animals progressed to the third stage. This stage was the same

as stage two except that animals now had only 10 seconds to make a response before an omission was registered. Following two consecutive sessions with over 100 responses animals achieved criterion to progress to the two-armed bandit task.

Two-armed bandit Task

The two-armed bandit task was performed over of 6 days with each session lasting one hour. In this task, animals were presented with two levers with one lever rewarded on 80% of trials, and the other lever rewarded on 20% of trials. To signal the start of the trial, both levers extended and the cue lights above the levers turned on, animals then had 10 seconds to make a response on either lever or an omission was registered. A lever press resulted in lever retraction and immediate delivery of the outcome, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise, counterbalanced across animals) or simply a different 3 second auditory cue (white noise or 2kHz pure tone) signaling non-reward. Following either a lever press or an omission, a 10 second intertrial interval (ITI) was triggered. To maintain a dynamic learning environment and high rates of rewarded and unrewarded outcomes, probability of reward was switched between levers after five consecutive responses on the high probability lever. Four males and four females remained on the two-armed bandit task during the task degradation (data not shown).

One-Lever Forced Choice Task

The one lever forced choice task was performed over 3 days with each session lasting one hour. In this task, animals were presented with a single lever (counterbalanced across animals). Pressing this lever resulted in probabilistic reward on

a predetermined schedule. The outcome schedule was matched to each animal's individual performance in the final three days of the two-armed bandit task, such that the first session in the one-lever task was yoked to the reward schedule experienced by the animal on day four in the two-armed bandit task, the second to day five, and the third to day six. To signal the start of the trial, the lever extended and the cue light above the lever turned on. Animals then had 10 seconds to make a response. A lever press resulted in lever retraction and immediate delivery of the outcome, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise) or simply a different 3 second auditory cue (white noise or 2kHz pure tone). Following either a lever press or an omission, a 10 second intertrial interval (ITI) was triggered.

No Lever Response Free Task

The no lever response free task was performed over the course of 3 days with each session lasting one hour. In this task, animals were able to retrieve non-contingently delivered rewards under a similar trial structure to both the two-armed bandit task and the one-lever forced choice task but with no levers available. To signal the start of the trial, cue lights above both levers turned on and remained illuminated for a period of time matched to each animal's response time in the last three days of the two-armed bandit task. After cue lights turned off, outcomes were delivered, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise) or simply a different 3 second auditory cue (white noise or 2kHz pure tone). As in the one-lever task, the outcome schedule was matched to each animal's performance in the final three days of the two-armed bandit task now also matching the latency to receive

the outcome to the trial-by-trial latency to lever press on the two-armed bandit task with a 10 second intertrial interval (ITI).

Frame Independent Projected Fiber Photometry

To measure calcium-associated changes in fluorescence in real time, recordings were made from vHip-NAc and mPFC-NAc-projecting cells during the two-armed bandit task, the one-lever forced choice task, and the no lever response free task. Samples were collected at a frequency of 20 Hz using Neurophotometrics hardware through Bonsai and FlyCap software. Recordings were coupled to the start of behavioral analysis by interfacing Bonsai with MED-PC using a custom DAQ box (Neurophotometrics).

Photometry data extraction and normalization

Photometry data were extracted and analyzed using custom-written scripts in Python. To normalize the data, the control channel (415nm) was fitted to the raw (470nm). The fitted control was then subtracted from the raw trace. The resultant trace was divided by the fitted control giving the ΔF/F and converted to a Z-score. This calculation was performed over the entirety of the session to preserve dynamic fluctuation in population activity that persists beyond individual trials to allow comparison across trials. For heatmaps Z-scores were baseline subtracted from average activity in the two seconds prior to lever press to accommodate moving baselines. For analyses of reward history, Z-scores were baseline subtracted from average activity in the two seconds prior to lever press on trial t-1 to account for shifted baselines in trial t0.

Optogenetics in Two-armed Bandit Task

Following lever press training, animals started the two-armed bandit task with optogenetic manipulations of mPFC-NAc and vHip-NAc activity for the duration of the ITI. Each day animals received either mPFC-NAc, vHip-NAc, or simultaneous mPFC-NAc and vHip-NAc stimulation on a subset of trials over the course of 9 days such that they received a total of 3 days of stimulation per condition for each stimulation protocol tested (5 Hz, 10 ms, 1-2 mW; 8 Hz, 10 ms, 2-3 mW). Order of stimulation days was fully counterbalanced within and between mice to avoid any order effects. Stimulation was delivered by 450 nm lasers controlled by a laser driver (Doric) running Doric studios software and triggered via a TTL (Med-Associates) at ITI start on a random subset of trials (30%) and terminated immediately prior to lever extension.

Ex vivo current-clamp electrophysiology

Brain slice preparation

Mice were deeply anesthetized with isofluorane. Transcardial perfusion was performed with 25-30 ml of ice-chilled carbogenated NNMDG artificial cerebrospinal fluid (aCSF: containing in mM: 92 NMDG, 2.5 KCl, 1.25 NaH_2PO_4, 30 $NaHCO_3$, 20 HEPES, 25 glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate, 0.5 $CaCl_2 \cdot 4H_2O$ and 10 $MgSO_4 \cdot 7H_2O$; titrated to pH 7.3–7.4 with concentrated hydrochloric acid). Brain slices (200 μm) were prepared in ice-chilled carbogenated NMDG aCSF by a vibratome (Lecia VT 1200S). All brain slices were recovery in 32–34 °C carbogenated NMDG aCSF for 10 min and then were transferred into room-temperature carbogenated HEPES holding aCSF (containing in mM: 92 NaCl, 2.5 KCl, 1.25 NaH_2PO_4, 30 $NaHCO_3$, 20 HEPES, 25

glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate, 2 $CaCl_2 \cdot 4H_2O$ and 2 $MgSO_4 \cdot 7H_2O$; titrated to pH to 7.3–7.4 with NaOH) for at least 1 hour before current-clamp recording.

Electrophysiology recordings

Current-clamp recordings were performed in room-temperature carbogenated aCSF (containing in mM: mM: 128 NaCl, 3 KCl, 1.25 NaH_2PO_4, 2 $MgCl_2$, 2 $CaCl_2$, 24 $NaHCO_3$ and 10 glucose; pH 7.2). The patch pipette solution was composed of (in mM) 115 K-gluconate, 20 KCl, 1.5 $MgCl_2$, 10 Phosphocreatine-Tris, 2 Mg-ATP, 0.54 Na-GTP and 10 HEPES. Blue light (wavelength: 470 nm) from a LED system (DC4100, Thorlabs) was used for optogenetic stimulation to evoke action potentials. The optogenetic stimulation protocol consisted of trains of 5 Hz (1-2 mW) or 8 Hz (2-3 mW) 10ms light pulses for 5 s. All signals were amplified and digitized by Multiclamp 700B (Molecular Device) and Digidata 1550B (Molecular Device) respectively. Series and access resistance were monitored during the experiments and signals were bessel filtered at 2 kHz.

Data Analysis & Statistics

Linear Mixed Effects Regression

Linear Mixed Effects Regression Models are a powerful approach to probe variance attributable to variables of interest (e.g. trial outcome) while simultaneously controlling for random effects (e.g. session ID) (Fetcho et al., 2023; Kato et al., 2022; Yu et al., 2022). This is useful for modeling instances where there is nonindependence in the structure of data e.g. multiple trials recorded within multiple animals. Models were fit using the full interaction of the factors of interest (trial outcome, previous trial outcome,

sex) and using animal ID and session ID as random effects using the *lme4* package in R (Bates et al., 2014). Where the dependent variable was latency, a Gamma link function was used to approximate the non-gaussian distribution. The fitted models were used to calculate estimated marginal means using the *emmeans* package in R (Lenth et al., 2021). The effect of variables of interest were then examined by comparing estimated marginal means. Given the large number of samples generated using this approach (all trials x all animals), comparisons of estimated marginal means were conducted using a *Z*-test and Sidak's method to adjust for multiple comparisons.

Cross-Correlation Time Delay Analysis

Time delay analysis was performed by first calculating the cross-correlation between mPFC-NAc and vHip-NAc during the ITI across a maximum lag of ± 5 seconds using the CCF function in R. The argument of the maximum (i.e. the time offset of peak correlation) of the resulting cross-correlation function was used to estimate the delay between mPFC-NAc and vHip-NAc on a trial-by-trial basis (Abboud & Sadeh, 1984). Linear mixed effects models were then fit to assess if the delay was non-zero (i.e. non-synchronous) using the following models to test for effects of sex [Time Delay~Sex-1+(1|ID)+(1|Day)] and for the interaction between sex and reward [Time Delay~Rewards:Sex-1+(1|ID)+(1|Day)]. The resulting regression coefficients from each model were examined to determine if the time delay was non-zero in any group (i.e. regression coefficient significantly different from zero).

Conditional Entropy Analysis

Conditional entropy is an information measure used to estimate the amount of additional information needed to explain one signal given full knowledge of a second signal. This

can be interpreted as the unique information contributed by a second signal beyond that contributed by a first with smaller conditional entropy suggesting less unique information carried by the second signal. Conditional entropy was calculated on the first two seconds and the last two seconds of the ITI using the PyInform package in Python to calculate the entropy (H) of the mPFC circuit given the vHip-NAc circuit, H(mPFC-NAc|vHip-NAc), and the entropy of the vHip-NAc circuit given the mPFC-NAc circuit, H(vHip-NAc|mPFC-NAc) (Cover & Thomas, 1991; Moore et al., 2018).

Code Availability

Code used to perform analyses for all figures available at https://github.com/eshaaniyer/mPFCvHip-NAc_RewardIntegration

Acknowledgments

We would like to thank Dr. Becket Ebitz, Dr. Mihaela Iordanova and Heike Schuler for their helpful comments and feedback throughout this project. Renderings of mice were created with BioRender.com. Research was supported by grants to RCB from NSERC and the Ludmer Centre for Neuroinformatics and Mental Health.

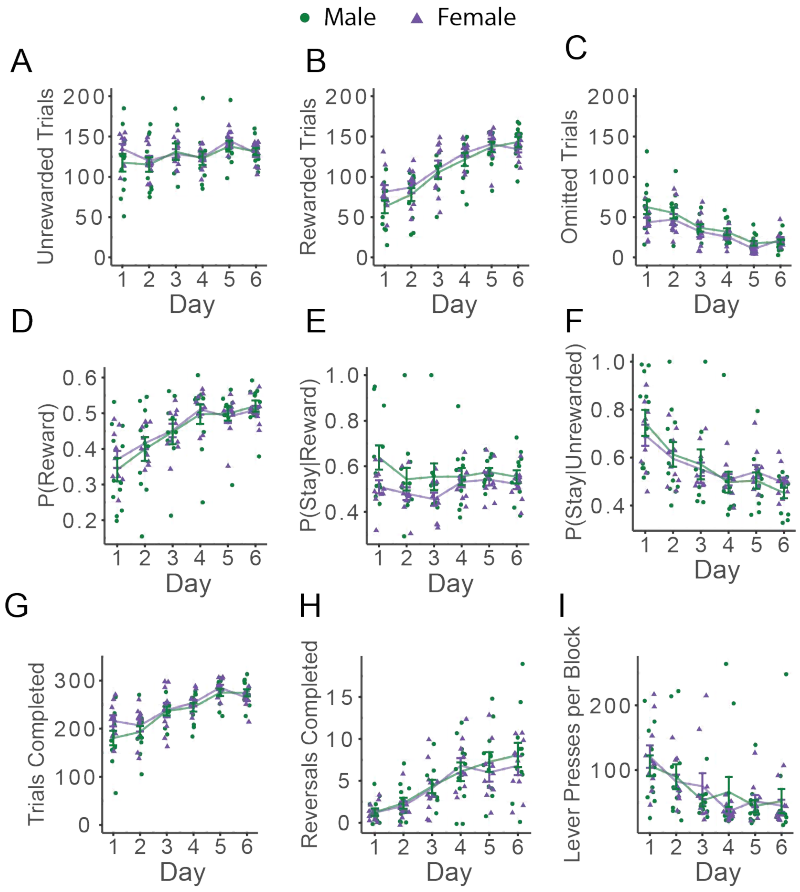
Author contributions

Conceptualization, E.S.I and R.C.B.; Methodology, E.S.I and R.C.B.; Investigation, E.S.I, P.V., S.W., J.M., Y.C.T., V.C.; Writing – Original Draft, E.S.I and R.C.B.; Writing – Review & Editing, E.S.I. and R.C.B.; Funding Acquisition, R.C.B.; Resources, R.C.B.; Supervision, R.C.B.

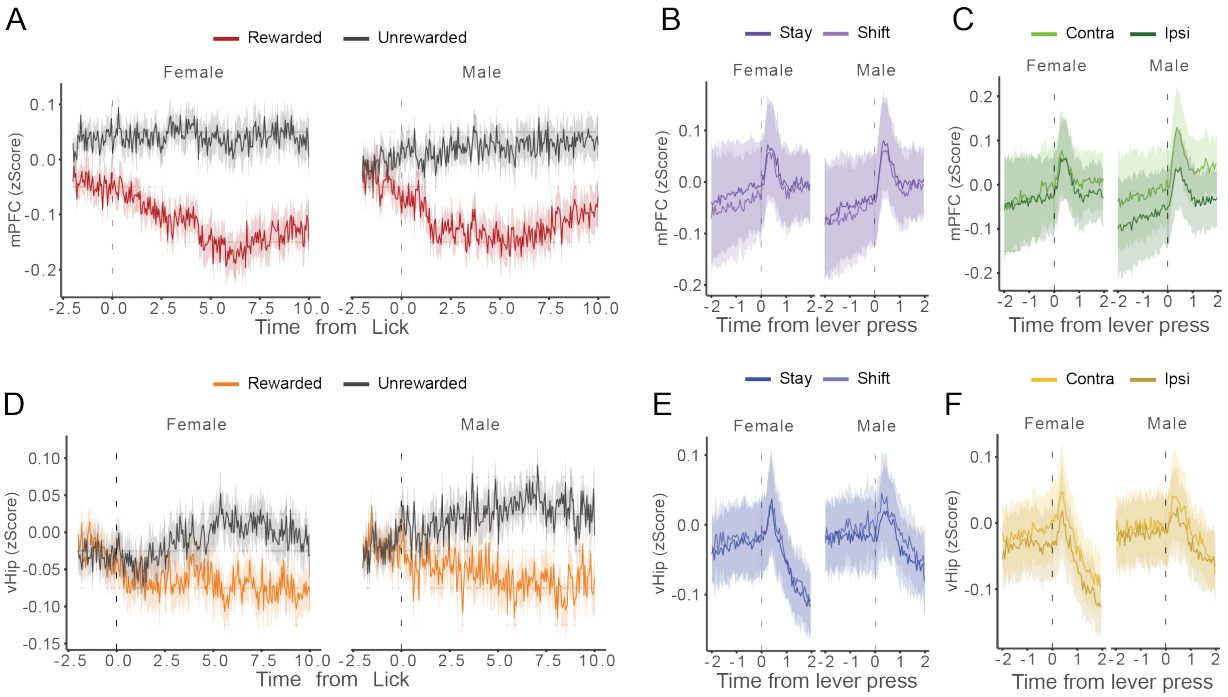
Declaration of interests

The authors declare no competing interests.

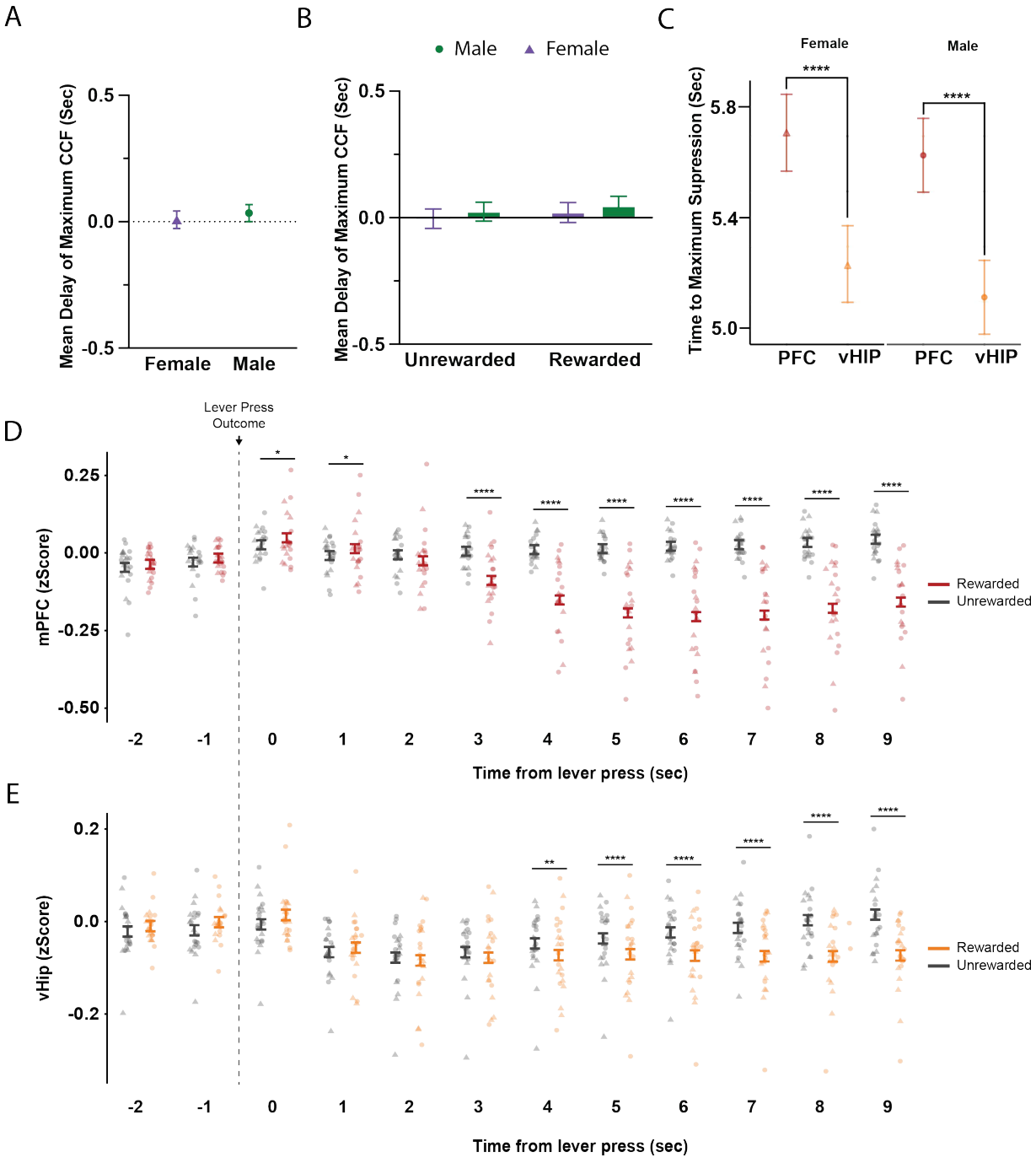
Supplementary Materials



Supplementary Figure 1. Behavior in the two-arm bandit task across days. *Male (n=12) and female (n=10) mice trend towards experiencing more (A) unrewarded trials across days (F=3.9919, p=0.0595, and experience more (B) rewarded trials across days (F=138.7239, p<0.0001), and fewer (C) omitted trials across days (F=51.5277, p<0.0001). (D) Across days, animals are more likely to earn rewards. Across days, animals do not change their staying probability following a (E) rewarded outcome but decrease their staying probability following (F) unrewarded outcomes (F=50.0712, p<0.0001). (G) Animals complete more trials across days (F=150.3925, p<0.0001), and increase the number of (H) reversals completed (F=40.3985, p<0.0001), and (I) decrease the number of lever presses performed to trigger a reversal (F=16.5760, p=0.0006). Individual-animals averages are indicated by circles for males and triangles for females. Error bars represent SEM around the estimated mean*

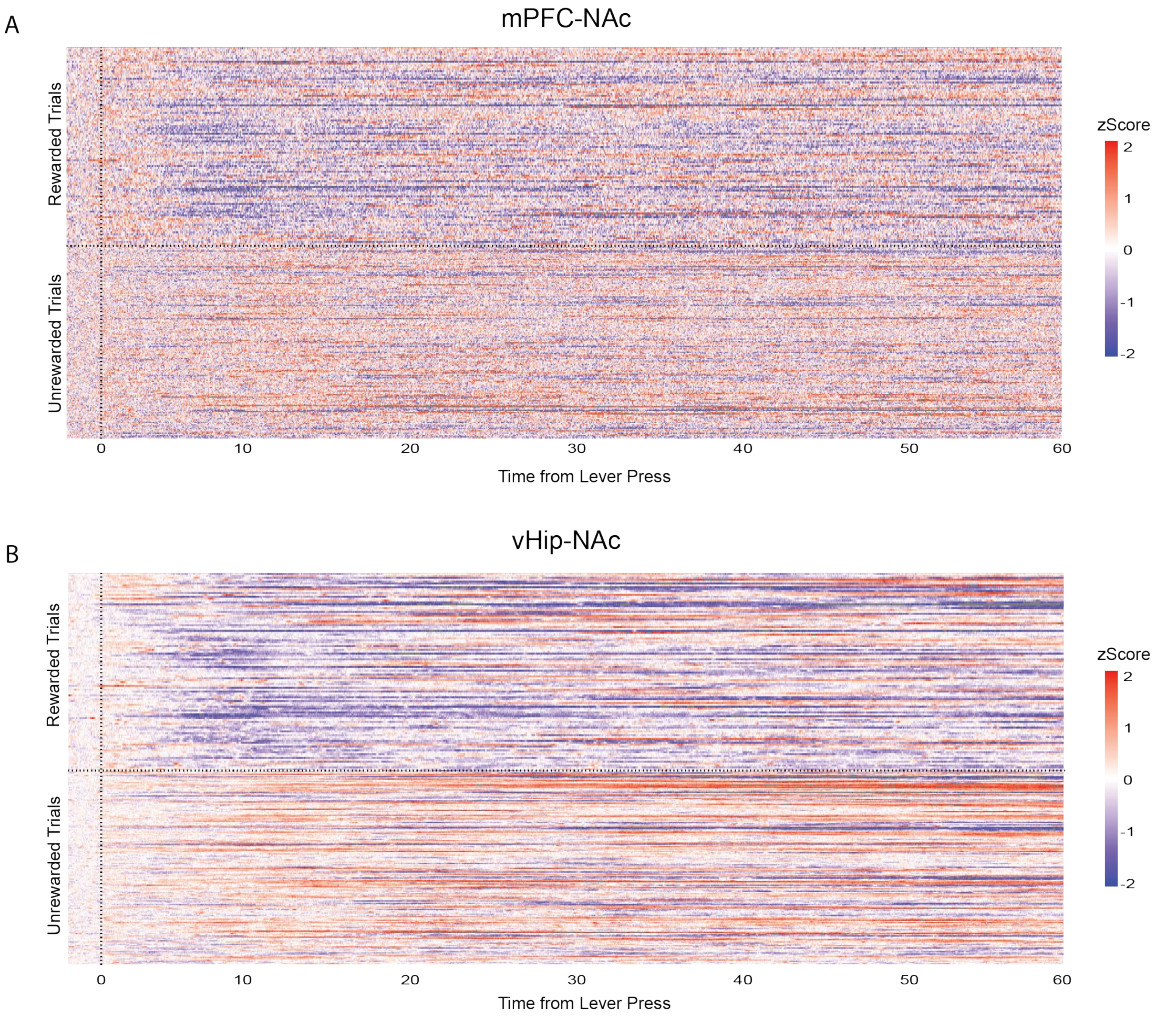


Supplemental Figure 2. mPFC-NAc and vHip-NAc activity time-locked to task-relevant features. *(A) Estimated mean mPFC-NAc activity time-locked to the first rewarded and unrewarded lick during the ITI in female (n=10) and male (n=12) mice. (B) Estimated mean mPFC-NAc activity time-locked to lever press on trials in which the previous choice is repeated (stay) and when a different choice is made (shift) (C) Estimated mean mPFC-NAc activity time-locked to lever press on trails in which animals choose the lever contralateral or ipsilateral to their implant. (D) Estimated mean vHip-NAc activity time-locked to the first rewarded and unrewarded lick during the ITI in female (n=10) and male (n=12) mice. (E) Estimated mean vHip-NAc activity time-locked to lever press on trials in which the previous choice is repeated (stay) and when a different choice is made (shift) (F) Estimated mean mPFC-NAc activity time-locked to lever press on trails in which animals choose the lever contralateral or ipsilateral to their implant.*

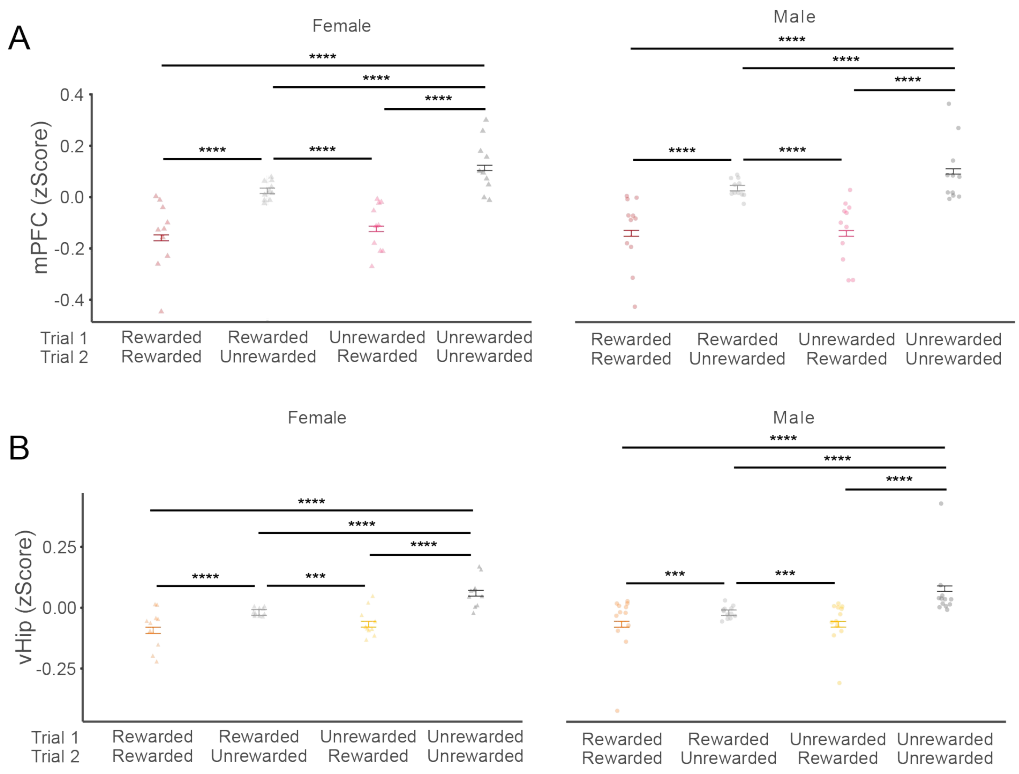


Supplementary Figure 3. Reward-associated encoding in mPFC-NAc and vHip-NAc is not led by either mPFC-NAc or vHip-NAc. *(A) Linear mixed effects models of the cross-correlation between mPFC-NAc and vHip-NAc shows that the time delay of the maximum cross-correlation does not significantly differ from zero in either females (n=10; Delay: 0.0074 ± 0.0349; Correlation: 0.2326 ± 0.0205) or males (n=12; Delay=0.0335 ± 0.0338; Correlation: 0.2255 ± 0.0196) (B) and is not modulated by trial outcome (Female Unrewarded: Delay = -0.0035 ± 0.0382, Correlation = 0.2340 ±*

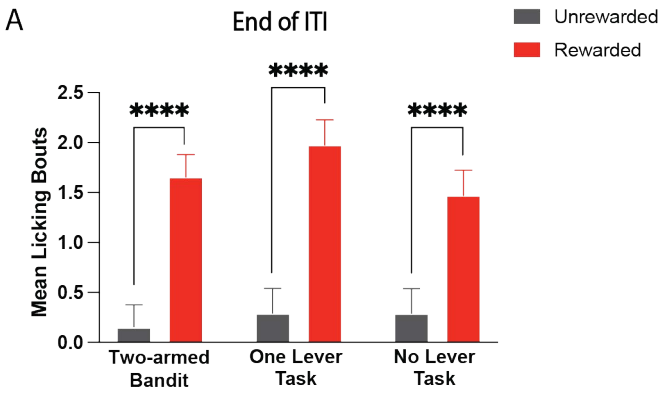
*0.0205; Female Rewarded: Delay = 0.0202 ± 0.0393, Correlation = 0.2310 ± 0.0205; Male Unrewarded: Delay =0.0237 ± 0.0373, Correlation = 0.2228 ± 0.0196; Male Rewarded: Delay = 0.0463 ± 0.0385, , Correlation = 0.2288 ± 0.0196). (C) Maximum suppression emerges at shorter latencies in vHip-NAc than mPFC-NAc in both females (Z=9.8839, p<0.0001) and males (Z=10.6193, p<0.0001). (D) Probing mPFC-NAc neural activity in 1sec bins reveals that, following a reward, activity initially increases, compared to unrewarded outcomes, before then suppressing activity to reward starting approximately three seconds following outcome delivery (-2s: Z=-1.3533, p=0.9020; -1s: Z=-1.7419, p=0.6396; 0s: Z=-3.0654, p=0.0258; 1s: Z=-3.0142, p=0.0305; 2s: Z=26344, p=0.0966; 3s: Z=12.6813, p<0.0001; 4s: Z=22.0057, p<0.0001; 5s: Z=28.0269, p<0.0001; 6s: Z=30.6498, p<0.0001; 7s: Z=30.6759, p<0.0001; 8s: Z=28.6823, p<0.0001; 9s: Z=27.3802, p<0.0001). (E) Probing vHip-NAc neural activity in 1sec bins reveals that, following a reward, activity initially decreases to both outcomes, with activity gradually increasing following unrewarded outcomes starting approximately four seconds following outcome delivery (-2s: Z=-1.6402, p=0.7212; -1s: Z=-2.4254, p=0.1688; 0s: Z=-2.8276, p=0.0548; 1s: Z=-1.3310, p=0.9118; 2s: Z=0.8472, p=0.9968; 3s: Z=1.6955, p=0.6774; 4s: Z=3.5239, p=0.0051; 5s: Z=4.7739, p<0.0001; 6s: Z=6.9466, p<0.0001; 7s: Z=8.5257, p<0.0001; 8s: Z=10.8820, p<0.0001; 9s: Z=12.2713, p<0.0001) Error bars represent SEM around the estimated mean. * p<0.05, **p<0.01, ****p<0.0001*



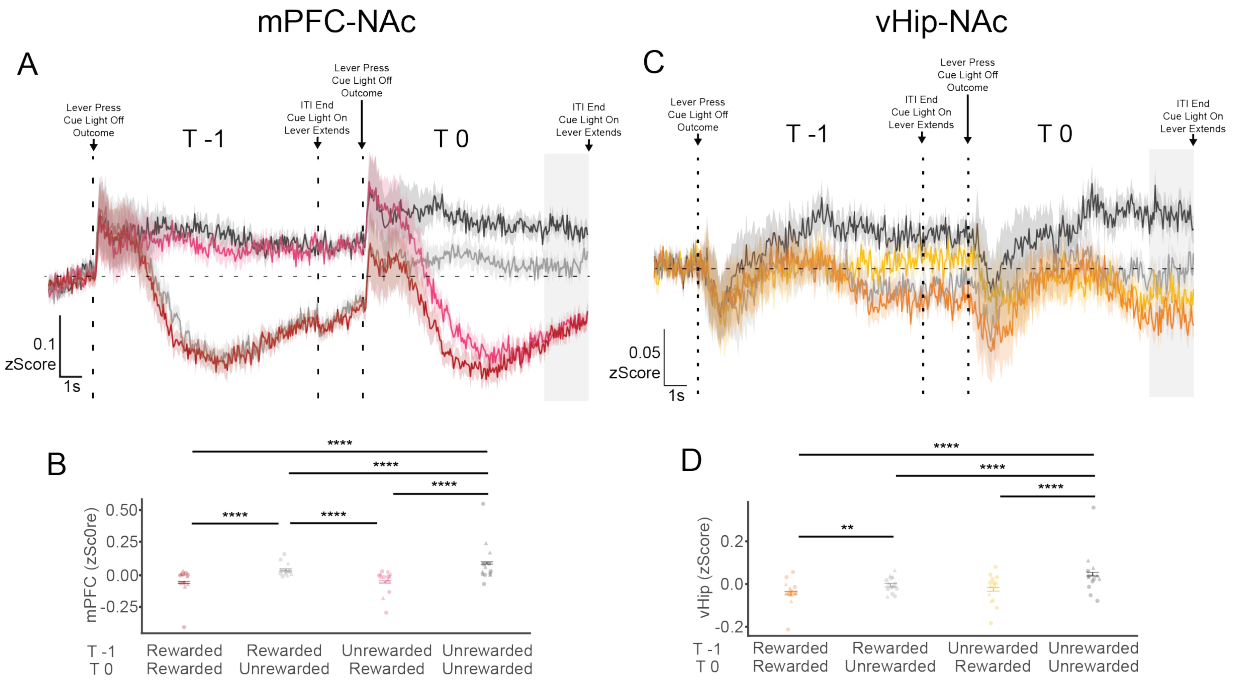
Supplementary Figure 4. Example traces of mPFC-NAc and vHip-NAc activity illustrating baseline fluctuations across an extended timescale. *(A) Heatmap of mPFC-NAc activity in the 60 seconds following lever press, sorted by rewarded and unrewarded outcomes in a representative animal across one session. (B) Heatmap of vHip-NAc activity in the 60 seconds following lever press, sorted by rewarded and unrewarded outcomes in a representative animal across one session.*



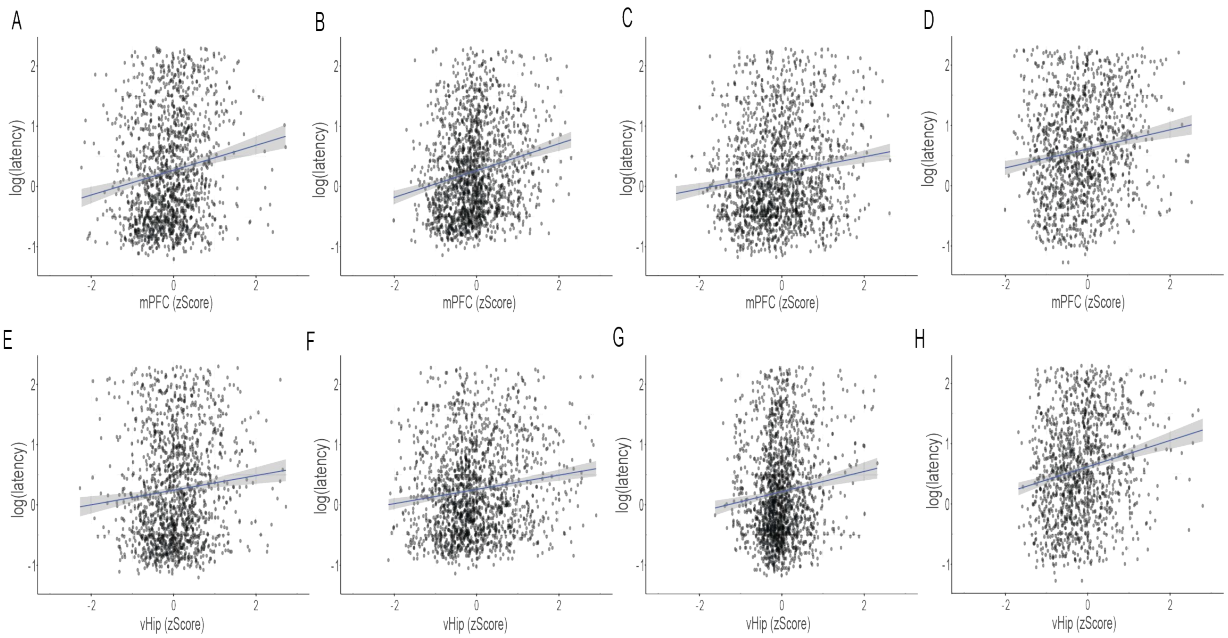
Supplementary Figure 5. mPFC-NAc and vHip-NAc similarly integrate reward history in females and males. *Estimated mean neural activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs in female (n=10) and male (n=12) mice. Analysis focused on 8-10 sec after lever press (ITI end) on trial t0. (A) On trial t0, mPFC-NAc activity is significantly suppressed by current reward in females (U→U vs U→R: Z=-20.0296, p<0.0001; R→U vs R→R: Z=-14.4311, p<0.0001; U→U vs R→R: Z=-21.8708, p<0.0001) and males (U→U vs U→R: Z=-20.3516, p<0.0001; R→U vs R→R: Z=-13.7364, p<0.0001; U→U vs R→R: Z=-19.3064, p<0.0001). When trial t0 is unrewarded, mPFC-NAc activity remains significantly suppressed by reward experienced on the previous trial, t-1, in females (U→U vs R→U: Z= 7.5245, p<0.0001; U→R vs R→R: Z= 2.7695, p=0.0653; U→R vs R→U: Z= -12.2087, p<0.0001) and males (U→U vs R→U: Z= 5.4925, p<0.0001; U→R vs R→R: Z= -0.0498, p=1.0000; U→R vs R→U: Z= -14.3616, p<0.0001). (B) On trial t0, vHip-NAc activity is significantly suppressed by current reward in females (U→U vs U→R: Z=-11.2415, p<0.0001; R→U vs R→R: Z=-6.0202, p<0.0001; U→U vs R→R: Z=-12.7816, p<0.0001) and males (U→U vs U→R: Z=-13.4823, p<0.0001; R→U vs R→R: Z=-4.0837, p=0.0005; U→U vs R→R: Z=-12.8133, p<0.0001). When trial t0 is unrewarded, vHip-NAc activity remains significantly suppressed by reward experienced on the previous trial, t-1, in females (U→U vs R→U: Z= 6.9947, p<0.0001; U→R vs R→R: Z= 2.0602, p=0.3825; U→R vs R→U: Z= -4.1470, p=0.0004) and males (U→U vs R→U: Z= 9.1069, p<0.0001; U→R vs R→R: Z= 0.0066, p=1; U→R vs R→U: Z= -4.2526, p=0.0002). Individual animal averages are indicated by circles for males and triangles for females. Error bars represent SEM around the estimated mean. ***p<0.001, ****p<0.0001*



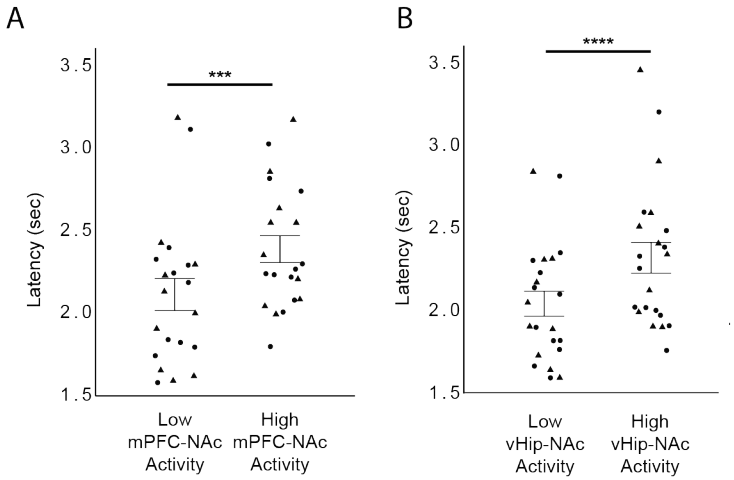
Supplementary Figure 6. Licking behavior is similar across task variants. *(A) Licking bouts are similarly increased at ITI end on rewarded compared to unrewarded trials across task variants with no significant differences in licking behavior across task variants (male n= 8, female n=6; Two-lever Task: Z=-22.415, p<0.0001; One-lever Task: Z=-18.599, p<0.0001, No Lever Task: Z=-22.415, p=0.0001). ****p<0.0001*



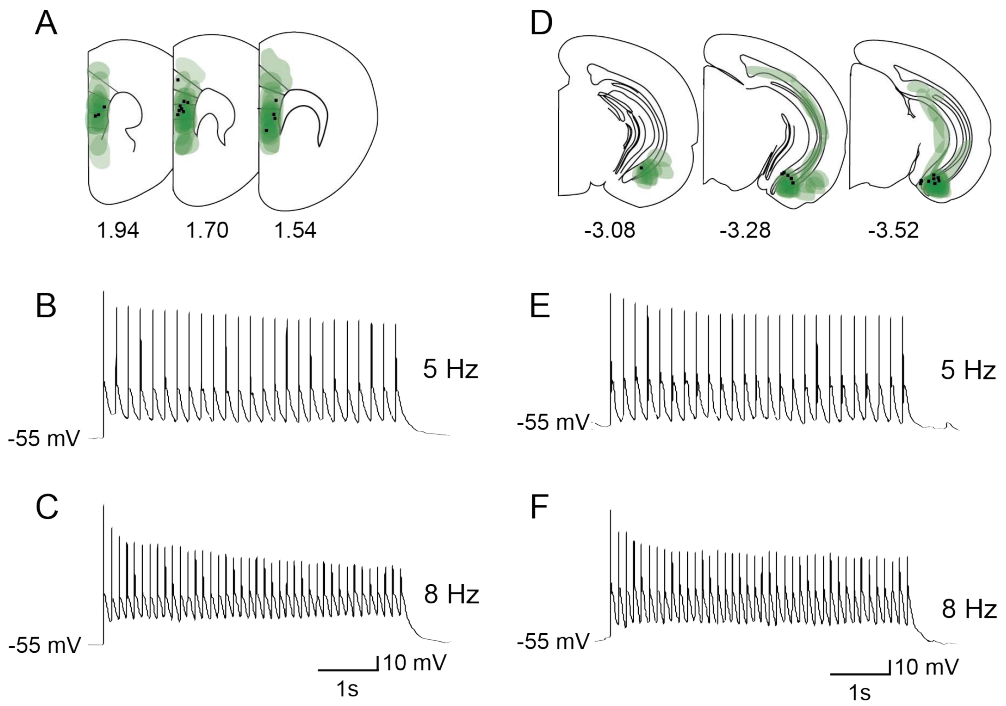
Supplementary Figure 7. mPFC-NAc and vHip-NAc continue to integrate reward history in the two-lever task following task degradation. *(A) Estimated mean mPFC-NAc activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs (male n= 8, female n=6). y=0 is indicated by a dashed horizontal line. Analysis focused on 8-10 sec after lever press (ITI end). (B) On trial t0, mPFC-NAc activity is suppressed by reward (U→U vs U→R: Z=12.6023, p<0.0001; R→U vs R→R: Z=8.3462, p<0.0001; U→U vs R→R: Z=12.8094, p<0.0001). When trial t0 is unrewarded, mPFC-NAc activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 4.5712, p<0.0001; U→R vs R→R: Z= 0.2476, p=0.9999; U→R vs R→U: Z= 8.1228, p<0.0001). (C) Estimated mean vHip-NAc activity across pairs of consecutive trials (t-1→t0) showing R→R, R→U, U→R, and U→U trial pairs (male n= 8, female n=6). y=0 is indicated by a dashed horizontal line. (D) On trial t0, vHip-NAc activity is suppressed by reward (U→U vs U→R: Z=6.0444, p<0.0001; R→U vs R→R: Z=3.2567, p=0.0067; U→U vs R→R: Z=7.5496, p<0.0001). When trial t0 is unrewarded, vHip-NAc activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 4.3316, p<0.0001; U→R vs R→R: Z= 1.5252, p=0.5580; U→R vs R→U: Z= 1.7330, p=0.4058). Error bars represent SEM around the estimated mean. **p<0.01, ***p<0.001, ****p<0.0001*



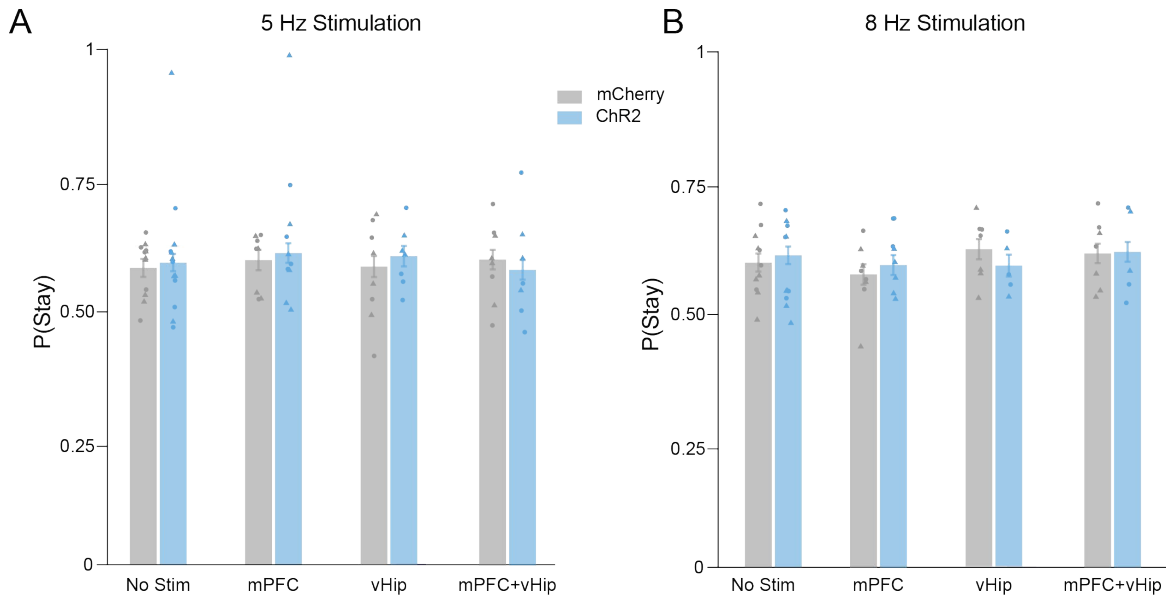
Supplemental Figure 8. Correlation between latency and mPFC-NAc or vHip-NAc activity at ITI end in representative animals. *(A) Correlation between mPFC-NAc and latency at ITI end in animal #207 (male, r=0.1542, p<0.0001). (B) Correlation between mPFC-NAc and latency at ITI end in animal #208 (male, r=0.1868, p<0.0001). (C) Correlation between mPFC-NAc and latency at ITI end in animal #215 (female, r=0.1368, p<0.0001). (D) Correlation between mPFC-NAc and latency at ITI end in animal #215 (female, r=0.1408, p<0.0001). (E) Correlation between vHip-NAc and latency at ITI end in animal #207 (male, r=0.1029, p<0.0001). (F) Correlation between vHip -NAc and latency at ITI end in animal #208 (male, r=0.1373, p<0.0001). (G) Correlation between vHip-NAc and latency at ITI end in animal #215 (female, r=0.1082, p<0.0001). (H) Correlation between vHip -NAc and latency at ITI end in animal #215 (female, r=0.1838, p<0.0001).*



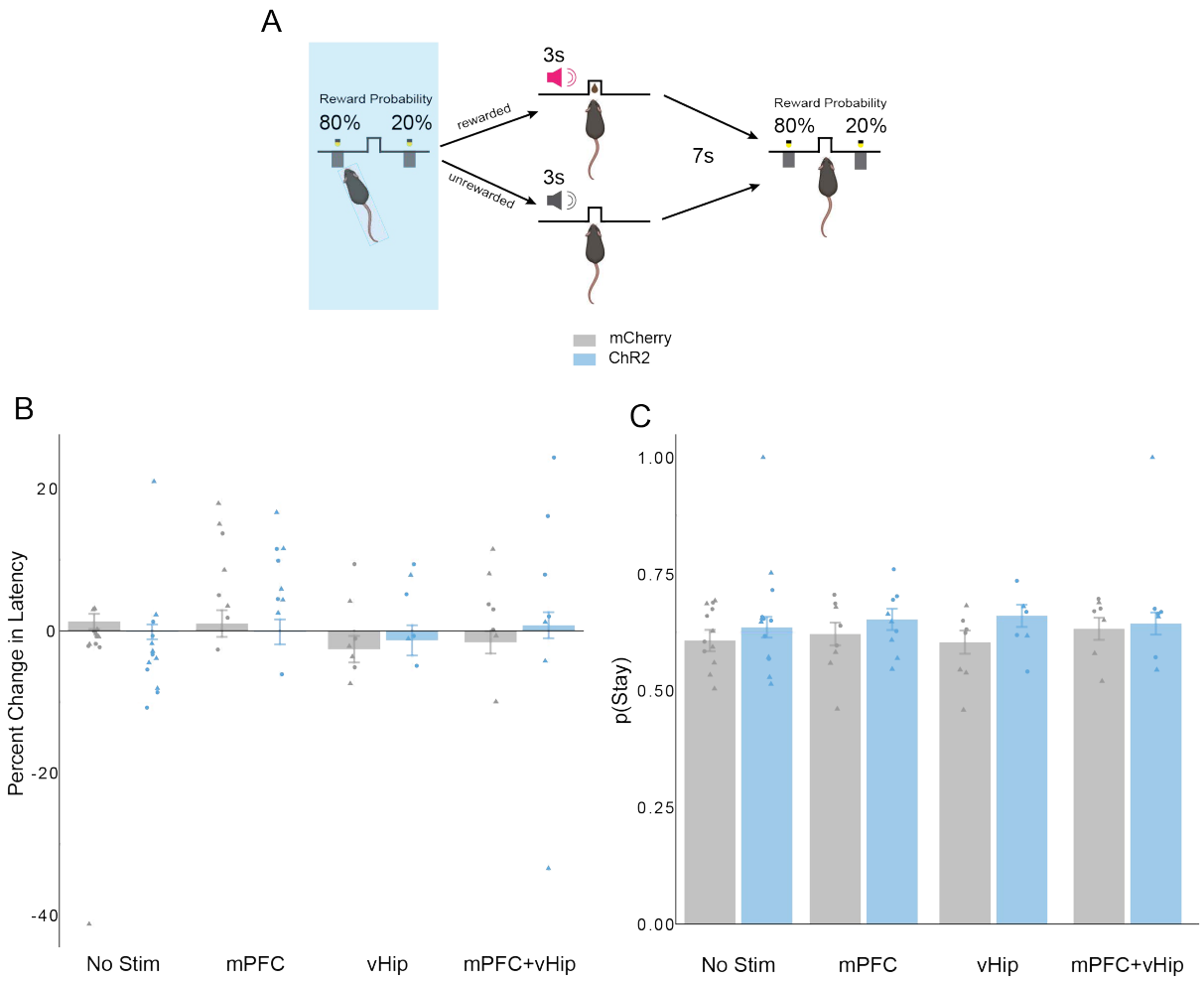
Supplemental Figure 9. Median-split mPFC-NAc and vHip-NAc activity associates with distinct response latencies. *A median split of average neural activity reveals that (A) high mPFC-NAc activity (t(20)=4.555, p=0.0002) and (B) high vHip-NAc activity (t(21)=5.48, p<0.0001) at ITI end increases latency to respond. Error bars represent SEM. ***p<0.001, ****p<0.0001*



Supplementary Figure 10. Histology and optogenetic validation. *(A) Plates indicate viral spread and location of fiber tips in the mPFC. In vitro optical stimulation of mPFC-NAc cell bodies induces spiking that reliably tracks stimulation at (B) 5Hz and (C) 8Hz. (D) Plates indicate viral spread and location of fiber tips in the vHip. In vitro optical stimulation of vHip-NAc cell bodies induces spiking that reliably tracks stimulation at (E) 5Hz and (F) 8Hz.*



Supplementary Figure 11. Optogenetic stimulation does not impact choice behavior. *(A) Neither simultaneous nor individual 5Hz stimulation of mPFC-NAc and vHip-NAc changed staying probabilities in ChR2 animals (male n=6, female n=7) compared to mCherry controls (male n=6, n=6 females). (B) Neither simultaneous nor individual 8Hz stimulation of mPFC-NAc and vHip-NAc changed staying probabilities in ChR2 animals (male n=6, female n=7) compared to mCherry controls (male n=6, female n=6). Individual-animal averages are indicated by circles for males and triangles for females. Error bars represent SEM around the estimated mean.*



Supplementary Figure 12. Optogenetic stimulation during lever-pressing does not impact behavior. *(A) Optogenetic stimulation protocol in the two-armed bandit task is delivered for the duration of the lever extension to either mPFC-NAc, vHip-NAc, or simultaneously to both circuits. Neither simultaneous nor individual 5Hz stimulation of mPFC-NAc and vHip-NAc affected (B) latency to press or (C) staying probabilities in ChR2 animals (male n=6, female n=7) compared to mCherry controls (male n=6, female n=6). Individual-animal averages are indicated by circles for males and triangles for females. Error bars represent SEM around the estimated mean.*

Results of pairwise comparisons assessing outcome encoding on current trial (T0) given specified histories of reward over previous three trials (T -1, T-2, T -3).

	T0 Pairwise Comparison	T -1 Outcome	T -2 Outcome	T -3 Outcome	Z Ratio	p Value
mPFC-NAc	Rewarded vs Unrewarded	Unrewarded	Unrewarded	Unrewarded	14.709144	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Unrewarded	Unrewarded	Rewarded	11.342379	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Unrewarded	Rewarded	Unrewarded	13.708781	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Unrewarded	Rewarded	Rewarded	12.042946	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Rewarded	Unrewarded	Unrewarded	9.963937	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Rewarded	Unrewarded	Rewarded	9.755649	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Rewarded	Rewarded	Unrewarded	6.794145	p<0.0001
mPFC-NAc	Rewarded vs Unrewarded	Rewarded	Rewarded	Rewarded	5.635762	p<0.0001
vHip-NAc	Rewarded vs Unrewarded	Unrewarded	Unrewarded	Unrewarded	9.6733623	p<0.0001
vHip-NAc	Rewarded vs Unrewarded	Unrewarded	Unrewarded	Rewarded	5.9858316	p<0.0001
vHip-NAc	Rewarded vs Unrewarded	Unrewarded	Rewarded	Unrewarded	5.4500847	p<0.0001
vHip-NAc	Rewarded vs Unrewarded	Unrewarded	Rewarded	Rewarded	5.2443238	p<0.0001
vHip-NAc	Rewarded vs Unrewarded	Rewarded	Unrewarded	Unrewarded	3.801793	p=0.0011
vHip-NAc	Rewarded vs Unrewarded	Rewarded	Unrewarded	Rewarded	3.4495963	p=0.0045
vHip-NAc	Rewarded vs Unrewarded	Rewarded	Rewarded	Unrewarded	1.2310677	p=0.8605
vHip-NAc	Rewarded vs Unrewarded	Rewarded	Rewarded	Rewarded	0.8397556	p=0.9834

Supplementary Table 1. *Results of pairwise comparisons assessing outcome encoding on current trial (T0) given specified histories of reward over previous three trials (T -1, T-2, T -3).*

Linear mixed model of latency on subsequent trial based on mPFC-NAc and vHip-NAc activity at ITI end

Model: Latency ~ mPFC * vHip + (1 | Animal ID) + (1 | Day), family = Gamma

	Estimate	Standard Err(	t value	p value
Intercept	0.459593	0.036952	12.438	2E-16
PFC	-0.033814	0.004059	-8.331	2E-16
vHip	-0.030279	0.003969	-7.629	2.36E-14
PFC:vHip	0.006913	0.00244	2.834	0.0046

Random Effects

	Variance	Standard Deviation
Animal ID	0.004705	0.06859
Day	0.006334	0.07959

Supplementary Table 2. *Linear mixed model of latency on subsequent trial based on mPFC-NAc and vHip-NAc activity at ITI end. Latency ~ mPFC * vHip + (1\|Animal ID) + (1\|Day), family = Gamma.*

References

Abboud, S., & Sadeh, D. (1984). The use of cross-correlation function for the alignment of ECG waveforms and rejection of extrasystoles. *Computers and Biomedical Research*, *17*(3), 258–266. https://doi.org/10.1016/S0010-4809(84)80017-8

Abela, A. R., Duan, Y., & Chudasama, Y. (2015). Hippocampal interplay with the nucleus accumbens is critical for decisions about time. *The European Journal of Neuroscience*, *42*(5), 2224–2233. https://doi.org/10.1111/ejn.13009

Aberman, J. E., & Salamone, J. D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience*, *92*(2), 545–552. https://doi.org/10.1016/S0306-4522(99)00004-4

Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron*, *59*(4), 648–661. https://doi.org/10.1016/j.neuron.2008.07.004

Amir, N., Prouvost, C., & Kuckertz, J. M. (2012). Lack of a Benign Interpretation Bias in Social Anxiety Disorder. *Cognitive Behaviour Therapy*, *41*(2), 119–129. https://doi.org/10.1080/16506073.2012.662655

Bagot, R. C., Parise, E. M., Peña, C. J., Zhang, H.-X., Maze, I., Chaudhury, D., Persaud, B., Cachope, R., Bolaños-Guzmán, C. A., Cheer, J. F., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2015). Ventral hippocampal afferents to the nucleus accumbens

regulate susceptibility to depression. *Nature Communications*, *6*, 7062. https://doi.org/10.1038/ncomms8062

Baldo, B. A., Sadeghian, K., Basso, A. M., & Kelley, A. E. (2002). Effects of selective dopamine D1 or D2 receptor blockade within nucleus accumbens subregions on ingestive behavior and associated motor activity. *Behavioural Brain Research*, *137*(1), 165–177. https://doi.org/10.1016/S0166-4328(02)00293-0

Barbano, M. F., Qi, J., Chen, E., Mohammad, U., Espinoza, O., Candido, M., Wang, H., Liu, B., Hahn, S., Vautier, F., & Morales, M. (2024). VTA glutamatergic projections to the nucleus accumbens suppress psychostimulant-seeking behavior. *Neuropsychopharmacology*, 1–11. https://doi.org/10.1038/s41386-024-01905-3

Bari, B. A., Grossman, C. D., Lubin, E. E., Rajagopalan, A. E., Cressy, J. I., & Cohen, J. Y. (2019). Stable representations of decision variables for flexible behavior. *Neuron*, *103*(5), 922-933.e7. https://doi.org/10.1016/j.neuron.2019.06.001

Barker, J. M., Bryant, K. G., & Chandler, L. J. (2019). Inactivation of ventral hippocampus projections promotes sensitivity to changes in contingency. *Learning & Memory*, *26*(1), 1–8.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv Preprint arXiv:1406.5823*.

Bathellier, B., Tee, S. P., Hrovat, C., & Rumpel, S. (2013). A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. *Proceedings of the National Academy of Sciences*, *110*(49), 19950–19955. https://doi.org/10.1073/pnas.1312125110

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., & Dayan, P. (2013). Dopamine Modulates Reward-Related Vigor. *Neuropsychopharmacology*, *38*(8), 1495–1503. https://doi.org/10.1038/npp.2013.48

Belujon, P., & Grace, A. A. (2008). Critical role of the prefrontal cortex in the regulation of hippocampus-accumbens information flow. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(39), 9797–9805. https://doi.org/10.1523/JNEUROSCI.2200-08.2008

Bercovici, D. A., Princz-Lebel, O., Tse, M. T., Moorman, D. E., & Floresco, S. B. (2018). Optogenetic Dissection of Temporal Dynamics of Amygdala-Striatal Interplay during Risk/Reward Decision Making. *eNeuro*, *5*(6), ENEURO.0422-18.2018. https://doi.org/10.1523/ENEURO.0422-18.2018

Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology*, *191*(3), 391–431. https://doi.org/10.1007/s00213-006-0578-x

Beyeler, A., Namburi, P., Glober, G. F., Simonnet, C., Calhoon, G. G., Conyers, G. F., Luck, R., Wildes, C. P., & Tye, K. M. (2016). Divergent routing of positive and negative information from the amygdala during memory retrieval. *Neuron*, *90*(2), 348–361. https://doi.org/10.1016/j.neuron.2016.03.004

Bishop, S. J., & Gagne, C. (2018). Anxiety, Depression, and Decision Making: A Computational Perspective. *Annual Review of Neuroscience*, *41*, 371–388. https://doi.org/10.1146/annurev-neuro-080317-062007

Bloem, B., Poorthuis, R. B., & Mansvelder, H. D. (2014). Cholinergic modulation of the medial prefrontal cortex: The role of nicotinic receptors in attention and regulation of neuronal activity. *Frontiers in Neural Circuits*, *8*, 17. https://doi.org/10.3389/fncir.2014.00017

Borrego, M. B., Grigsby, K. B., Townsley, K. G., Chan, A., Firsick, E. F., Tran, A., & Ozburn, A. R. (2022). Central nucleus of the amygdala projections onto the nucleus accumbens core regulate binge-like alcohol drinking in a CRF-dependent manner. *Neuropharmacology*, *203*, 108874. https://doi.org/10.1016/j.neuropharm.2021.108874

Bouarab, C., Thompson, B., & Polter, A. M. (2019). VTA GABA Neurons at the Interface of Stress and Reward. *Frontiers in Neural Circuits*, *13*. https://doi.org/10.3389/fncir.2019.00078

Bouras, N. N., Mack, N. R., & Gao, W.-J. (2023). Prefrontal modulation of anxiety through a lens of noradrenergic signaling. *Frontiers in Systems Neuroscience*, *17*, 1173326. https://doi.org/10.3389/fnsys.2023.1173326

Bouret, S., & Richmond, B. J. (2015). Sensitivity of Locus Ceruleus Neurons to Reward Value for Goal-Directed Actions. *The Journal of Neuroscience*, *35*(9), 4005–4014. https://doi.org/10.1523/JNEUROSCI.4553-14.2015

Britt, J. P., Benaliouad, F., McDevitt, R. A., Stuber, G. D., Wise, R. A., & Bonci, A. (2012). Synaptic and Behavioral Profile of Multiple Glutamatergic Inputs to the Nucleus Accumbens. *Neuron*, *76*(4), 790–803. https://doi.org/10.1016/j.neuron.2012.09.040

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Brown, V. J., & Bowman, E. M. (1995). Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. *The European Journal of Neuroscience*, *7*(12), 2479–2485. https://doi.org/10.1111/j.1460-9568.1995.tb01046.x

Calipari, E. S., Bagot, R. C., Purushothaman, I., Davidson, T. J., Yorgason, J. T., Peña, C. J., Walker, D. M., Pirpinias, S. T., Guise, K. G., Ramakrishnan, C., Deisseroth, K., & Nestler, E. J. (2016). In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proceedings of the National Academy of Sciences*, *113*(10), 2726–2731. https://doi.org/10.1073/pnas.1521238113

Carter, A. G., Soler-Llavina, G. J., & Sabatini, B. L. (2007). Timing and Location of Synaptic Inputs Determine Modes of Subthreshold Integration in Striatal Medium Spiny Neurons. *The Journal of Neuroscience*, *27*(33), 8967–8977. https://doi.org/10.1523/JNEUROSCI.2798-07.2007

Carvalho Poyraz, F., Holzner, E., Bailey, M. R., Meszaros, J., Kenney, L., Kheirbek, M. A., Balsam, P. D., & Kellendonk, C. (2016). Decreasing Striatopallidal Pathway Function Enhances Motivation by Energizing the Initiation of Goal-Directed Action. *The Journal of Neuroscience*, *36*(22), 5988–6001. https://doi.org/10.1523/JNEUROSCI.0444-16.2016

Chen, C. S., Knep, E., Han, A., Ebitz, R. B., & Grissom, N. M. (2021). Sex differences in learning from exploration. *eLife*, *10*, e69748. https://doi.org/10.7554/eLife.69748

Chen, C. S., Mueller, D., Knep, E., Ebitz, R. B., & Grissom, N. M. (2024). Dopamine and Norepinephrine Differentially Mediate the Exploration-Exploitation Tradeoff. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *44*(44), e1194232024. https://doi.org/10.1523/JNEUROSCI.1194-23.2024

Chen, G., Lai, S., Bao, G., Ke, J., Meng, X., Lu, S., Wu, X., Xu, H., Wu, F., Xu, Y., Xu, F., Bi, G.-Q., Peng, G., Zhou, K., & Zhu, Y. (2023). Distinct reward processing by subregions of the nucleus accumbens. *Cell Reports*, *42*(2). https://doi.org/10.1016/j.celrep.2023.112069

Christoffel, D. J., Walsh, J. J., Hoerbelt, P., Heifets, B. D., Llorach, P., Lopez, R. C., Ramakrishnan, C., Deisseroth, K., & Malenka, R. C. (2021). Selective filtering of excitatory inputs to nucleus accumbens by dopamine and serotonin. *Proceedings of the National Academy of Sciences*, *118*(24), e2106648118. https://doi.org/10.1073/pnas.2106648118

Cid-Pellitero, E. del, & Garzón, M. (2011). Medial prefrontal cortex receives input from dorsal raphe nucleus neurons targeted by Hypocretin1/OrexinA-containing axons. *Neuroscience*, *172*, 30–43. https://doi.org/10.1016/j.neuroscience.2010.10.058

Cohen, J. Y., Amoroso, M. W., & Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. *eLife*, *4*, e06346. https://doi.org/10.7554/eLife.06346

Collins, D. P., Anastasiades, P. G., Marlin, J. J., & Carter, A. G. (2018). Reciprocal circuits linking the prefrontal cortex with dorsal and ventral thalamic nuclei. *Neuron*, *98*(2), 366-379.e4. https://doi.org/10.1016/j.neuron.2018.03.024

Cover, T. M., & Thomas, J. A. (1991). Entropy, relative entropy and mutual information. *Elements of Information Theory*, *2*(1), 12–13.

Cox, J., Minerva, A. R., Fleming, W. T., Zimmerman, C. A., Hayes, C., Zorowitz, S., Bandi, A., Ornelas, S., McMannon, B., Parker, N. F., & Witten, I. B. (2023). A neural substrate of sex-dependent modulation of motivation. *Nature Neuroscience*, *26*(2), Article 2. https://doi.org/10.1038/s41593-022-01229-9

Dalton, G. L., Phillips, A. G., & Floresco, S. B. (2014). Preferential Involvement by Nucleus Accumbens Shell in Mediating Probabilistic Learning and Reversal Shifts. *Journal of Neuroscience*, *34*(13), 4618–4626. https://doi.org/10.1523/JNEUROSCI.5058-13.2014

Dana, H., Sun, Y., Mohar, B., Hulse, B. K., Kerlin, A. M., Hasseman, J. P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., Macklin, J. J., Chen, Y., Konnerth, A., Jayaraman, V., Looger, L. L., Schreiter, E. R., Svoboda, K., & Kim, D. S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods*, *16*(7), 649–657. https://doi.org/10.1038/s41592-019-0435-6

Day, J. J., Jones, J. L., & Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs. *The European Journal of Neuroscience*, *33*(2), 308–321. https://doi.org/10.1111/j.1460-9568.2010.07531.x

De Groote, A., & de Kerchove d'Exaerde, A. (2021). Thalamo-Nucleus Accumbens Projections in Motivated Behaviors and Addiction. *Frontiers in Systems Neuroscience*, *15*, 711350. https://doi.org/10.3389/fnsys.2021.711350

de Jong, J. W., Afjei, S. A., Dorocic, I. P., Peck, J. R., Liu, C., Kim, C. K., Tian, L., Deisseroth, K., & Lammel, S. (2019). A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron*, *101*(1), 133-151.e7. https://doi.org/10.1016/j.neuron.2018.11.005

Dezfouli, A., Balleine, B. W., & Nock, R. (2019). Optimal response vigor and choice under non-stationary outcome values. *Psychonomic Bulletin & Review*, *26*(1), 182–204. https://doi.org/10.3758/s13423-018-1500-3

Dieterich, A., Floeder, J., Stech, K., Lee, J., Srivastava, P., Barker, D. J., & Samuels, B. A. (2021). Activation of Basolateral Amygdala to Nucleus Accumbens Projection Neurons Attenuates Chronic Corticosterone-Induced Behavioral Deficits in Male Mice. *Frontiers in Behavioral Neuroscience*, *15*. https://doi.org/10.3389/fnbeh.2021.643272

Do-Monte, F. H., Minier-Toribio, A., Quiñones-Laracuente, K., Medina-Colón, E. M., & Quirk, G. J. (2017). Thalamic regulation of sucrose-seeking during unexpected reward omission. *Neuron*, *94*(2), 388-400.e4. https://doi.org/10.1016/j.neuron.2017.03.036

Dong, X., Li, S., & Kirouac, G. J. (2017). Collateralization of projections from the paraventricular nucleus of the thalamus to the nucleus accumbens, bed nucleus of the stria terminalis, and central nucleus of the amygdala. *Brain Structure & Function*, *222*(9), 3927–3943. https://doi.org/10.1007/s00429-017-1445-8

Ebitz, R. B., Albarran, E., & Moore, T. (2018). Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex. *Neuron*, *97*(2), 450-461.e9. https://doi.org/10.1016/j.neuron.2017.12.007

Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). Reinforcement learning and Bayesian inference provide complementary models for the unique advantage of adolescents in stochastic reversal. *Developmental Cognitive Neuroscience*, *55*, 101106. https://doi.org/10.1016/j.dcn.2022.101106

Engelke, D. S., Zhang, X. O., O'Malley, J. J., Fernandez-Leon, J. A., Li, S., Kirouac, G. J., Beierlein, M., & Do-Monte, F. H. (2021). A hypothalamic-thalamostriatal circuit that controls approach-avoidance conflict in rats. *Nature Communications*, *12*, 2517. https://doi.org/10.1038/s41467-021-22730-y

Fetcho, R. N., Hall, B. S., Estrin, D. J., Walsh, A. P., Schuette, P. J., Kaminsky, J., Singh, A., Roshgodal, J., Bavley, C. C., Nadkarni, V., Antigua, S., Huynh, T. N., Grosenick, L., Carthy, C., Komer, L., Adhikari, A., Lee, F. S., Rajadhyaksha, A. M., & Liston, C. (2023). Regulation of social interaction in mice by a frontostriatal circuit modulated by established hierarchical relationships. *Nature Communications*, *14*, 2487. https://doi.org/10.1038/s41467-023-37460-6

Fisone, G., Bonito-Oliva, A., & Feyder, M. (2011). Deciphering the Actions of Antiparkinsonian and Antipsychotic Drugs on cAMP/DARPP-32 Signaling. *Frontiers in Neuroanatomy*, *5*. https://doi.org/10.3389/fnana.2011.00038

Floresco, S. B. (2007). Dopaminergic regulation of limbic-striatal interplay. *Journal of Psychiatry & Neuroscience : JPN*, *32*(6), 400–411.

Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of Psychology*, *66*(1), 25–52. https://doi.org/10.1146/annurev-psych-010213-115159

Floresco, S. B., Montes, D. R., Tse, M. M. T., & van Holstein, M. (2018). Differential Contributions of Nucleus Accumbens Subregions to Cue-Guided Risk/Reward Decision Making and Implementation of Conditional Rules. *The Journal of Neuroscience*, *38*(8), 1901–1914. https://doi.org/10.1523/JNEUROSCI.3191-17.2018

Francis, T. C., & Lobo, M. K. (2017). Emerging role for nucleus accumbens medium spiny neuron subtypes in depression. *Biological Psychiatry*, *81*(8), 645–653. https://doi.org/10.1016/j.biopsych.2016.09.007

French, S. J., & Totterdell, S. (2002). Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *Journal of Comparative Neurology*, *446*(2), 151–165.

Gallo, E. F., Meszaros, J., Sherman, J. D., Chohan, M. O., Teboul, E., Choi, C. S., Moore, H., Javitch, J. A., & Kellendonk, C. (2018). Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nature Communications*, *9*, 1086. https://doi.org/10.1038/s41467-018-03272-2

Garcia-Garcia, A., Tancredi, A. N.-, & Leonardo, E. D. (2014). 5-HT1A receptors in mood and anxiety: Recent insights into autoreceptor versus heteroreceptor function. *Psychopharmacology*, *231*(4), 623–636. https://doi.org/10.1007/s00213-013-3389-x

Gerfen, C. R. (2023). Segregation of D1 and D2 dopamine receptors in the striatal direct and indirect pathways: An historical perspective. *Frontiers in Synaptic Neuroscience*, *14*, 1002960. https://doi.org/10.3389/fnsyn.2022.1002960

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D1 and D2 Dopamine Receptor-regulated Gene Expression of Striatonigral and Striatopallidal Neurons. *Science*, *250*(4986), 1429–1432. https://doi.org/10.1126/science.2147780

Goto, Y., & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, *8*(6), Article 6. https://doi.org/10.1038/nn1471

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences*, *30*(5), 220–227. https://doi.org/10.1016/j.tins.2007.03.003

Guillaumin, M. C. C., Viskaitis, P., Bracey, E., Burdakov, D., & Peleg-Raibstein, D. (2023). Disentangling the role of NAc D1 and D2 cells in hedonic eating. *Molecular Psychiatry*, *28*(8), 3531–3547. https://doi.org/10.1038/s41380-023-02131-x

Guitart-Masip, M., Beierholm, U. R., Dolan, R., Duzel, E., & Dayan, P. (2011). Vigor in the Face of Fluctuating Rates of Reward: An Experimental Examination. *Journal of Cognitive Neuroscience*, *23*(12), 3933–3938. https://doi.org/10.1162/jocn_a_00090

Gustafson, N. J., & Daw, N. D. (2011). Grid Cells, Place Cells, and Geodesic Generalization for Spatial Reinforcement Learning. *PLOS Computational Biology*, *7*(10), e1002235. https://doi.org/10.1371/journal.pcbi.1002235

Hamel, L., Cavdaroglu, B., Yeates, D., Nguyen, D., Riaz, S., Patterson, D., Khan, N., Kirolos, N., Roper, K., Ha, Q. A., & Ito, R. (2022). Cortico-Striatal Control over Adaptive Goal-Directed Responding Elicited by Cues Signaling Sucrose Reward or Punishment. *Journal of Neuroscience*, *42*(18), 3811–3822. https://doi.org/10.1523/JNEUROSCI.2175-21.2022

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2016). Mesolimbic Dopamine Signals the Value of Work. *Nature Neuroscience*, *19*(1), 117–126. https://doi.org/10.1038/nn.4173

Han, X., Jing, M., Zhao, T., Wu, N., Song, R., & Li, J. (2017). Role of dopamine projections from ventral tegmental area to nucleus accumbens and medial prefrontal cortex in reinforcement behaviors assessed using optogenetic manipulation. *Metabolic Brain Disease*, *32*(5), 1491–1502. https://doi.org/10.1007/s11011-017-0023-3

Hanson, E., Brandel-Ankrapp, K. L., & Arenkiel, B. R. (2021). Dynamic Cholinergic Tone in the Basal Forebrain Reflects Reward-Seeking and Reinforcement During Olfactory Behavior. *Frontiers in Cellular Neuroscience*, *15*, 635837. https://doi.org/10.3389/fncel.2021.635837

He, Y., Huang, Y. H., Schlüter, O. M., & Dong, Y. (2023). Cue- versus reward-encoding basolateral amygdala projections to nucleus accumbens. *eLife*, *12*, e89766. https://doi.org/10.7554/eLife.89766

Hernandes, P. M., Batistela, M. F., Vilela-Costa, H. H., Sant'Ana, A. B., Kumpel, V. D., Tirapelle, M. C., Bom, A. de O. P., de Andrade, T. G. C. S., & Zangrossi, H. (2021). Role

of 5-HT1A receptors in the ventral hippocampus in the regulation of anxiety- and panic-related defensive behaviors in rats. *Behavioural Brain Research*, *408*, 113296. https://doi.org/10.1016/j.bbr.2021.113296

Herrnstein, R. J. (1997). *The matching law: Papers in psychology and economics*. Harvard University Press.

Hofmeister, J., & Sterpenich, V. (2015). A Role for the Locus Ceruleus in Reward Processing: Encoding Behavioral Energy Required for Goal-Directed Actions. *Journal of Neuroscience*, *35*(29), 10387–10389. https://doi.org/10.1523/JNEUROSCI.1734-15.2015

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, *500*(7464), Article 7464. https://doi.org/10.1038/nature12475

Huang, S., Borgland, S. L., & Zamponi, G. W. (2019). Dopaminergic modulation of pain signals in the medial prefrontal cortex: Challenges and perspectives. *Neuroscience Letters*, *702*, 71–76. https://doi.org/10.1016/j.neulet.2018.11.043

Ikemoto, S., Glazier, B. S., Murphy, J. M., & McBride, W. J. (1997). Role of Dopamine D1 and D2 Receptors in the Nucleus Accumbens in Mediating Reward. *The Journal of Neuroscience*, *17*(21), 8580–8587. https://doi.org/10.1523/JNEUROSCI.17-21-08580.1997

Iyer, E. S., Vitaro, P., Wu, S., Muir, J., Tse, Y. C., Cvetkovska, V., & Bagot, R. C. (2024). *Reward integration in prefrontal-cortical and ventral-hippocampal nucleus accumbens*

inputs cooperatively modulates engagement (p. 2024.06.27.601063). bioRxiv. https://doi.org/10.1101/2024.06.27.601063

Jeong, H., Kim, D., Song, M., Paik, S.-B., & Jung, M. W. (2020). Distinct roles of parvalbumin- and somatostatin-expressing neurons in flexible representation of task variables in the prefrontal cortex. *Progress in Neurobiology*, *187*, 101773. https://doi.org/10.1016/j.pneurobio.2020.101773

Jeong, H., Taylor, A., Floeder, J. R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D. A., & Namboodiri, V. M. K. (2022). Mesolimbic dopamine release conveys causal associations. *Science*, *378*(6626), eabq6740. https://doi.org/10.1126/science.abq6740

Jing, M., Zhang, P., Wang, G., Feng, J., Mesik, L., Zeng, J., Jiang, H., Wang, S., Looby, J. C., Guagliardo, N. A., Langma, L. W., Lu, J., Zuo, Y., Talmage, D. A., Role, L. W., Barrett, P. Q., Zhang, L. I., Luo, M., Song, Y., … Li, Y. (2018). A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nature Biotechnology*, *36*(8), 726–737. https://doi.org/10.1038/nbt.4184

Kagiampaki, Z., Rohner, V., Kiss, C., Curreli, S., Dieter, A., Wilhelm, M., Harada, M., Duss, S. N., Dernic, J., Bhat, M. A., Zhou, X., Ravotto, L., Ziebarth, T., Wasielewski, L. M., Sönmez, L., Benke, D., Weber, B., Bohacek, J., Reiner, A., … Patriarchi, T. (2023). Sensitive multicolor indicators for monitoring norepinephrine in vivo. *Nature Methods*, *20*(9), 1426–1436. https://doi.org/10.1038/s41592-023-01959-z

Kato, T., Mitsukura, Y., Yoshida, K., Mimura, M., Takata, N., & Tanaka, K. F. (2022). Oscillatory Population-Level Activity of Dorsal Raphe Serotonergic Neurons Is

Inscribed in Sleep Structure. *The Journal of Neuroscience*, *42*(38), 7244–7255. https://doi.org/10.1523/JNEUROSCI.2288-21.2022

Kebabian, J. W., & Calne, D. B. (1979). Multiple receptors for dopamine. *Nature*, *277*(5692), 93–96. https://doi.org/10.1038/277093a0

Keiflin, R., & Janak, P. H. (2015). Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron*, *88*(2), 247–263. https://doi.org/10.1016/j.neuron.2015.08.037

Keistler, C., Barker, J. M., & Taylor, J. R. (2015). Infralimbic prefrontal cortex interacts with nucleus accumbens shell to unmask expression of outcome-selective Pavlovian-to-instrumental transfer. *Learning & Memory*, *22*(10), 509–513. https://doi.org/10.1101/lm.038810.115

Klausberger, T., & Somogyi, P. (2008). Neuronal Diversity and Temporal Dynamics: The Unity of Hippocampal Circuit Operations. *Science (New York, N.Y.)*, *321*(5885), 53–57. https://doi.org/10.1126/science.1149381

Koo, J. W., Lobo, M. K., Chaudhury, D., Labonté, B., Friedman, A., Heller, E., Peña, C. J., Han, M.-H., & Nestler, E. J. (2014). Loss of BDNF Signaling in D1R-Expressing NAc Neurons Enhances Morphine Reward by Reducing GABA Inhibition. *Neuropsychopharmacology*, *39*(11), 2646–2653. https://doi.org/10.1038/npp.2014.118

Köther, A. K., Alpers, G. W., Büdenbender, B., Lenhart, M., Michel, M. S., & Kriegmair, M. C. (2021). Predicting decisional conflict: Anxiety and depression in shared decision

making. *Patient Education and Counseling*, *104*(5), 1229–1236. https://doi.org/10.1016/j.pec.2020.10.037

Kravitz, A. V., & Kreitzer, A. C. (2012). Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. *Physiology (Bethesda, Md.)*, *27*(3), 10.1152/physiol.00004.2012. https://doi.org/10.1152/physiol.00004.2012

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818. https://doi.org/10.1038/nn.3100

Kuchibhotla, K. V., Hindmarsh Sten, T., Papadoyannis, E. S., Elnozahy, S., Fogelson, K. A., Kumar, R., Boubenec, Y., Holland, P. C., Ostojic, S., & Froemke, R. C. (2019). Dissociating task acquisition from expression during learning reveals latent knowledge. *Nature Communications*, *10*(1), 2151. https://doi.org/10.1038/s41467-019-10089-0

Kupchik, Y. M., Brown, R. M., Heinsbroek, J. A., Lobo, M. K., Schwartz, D. J., & Kalivas, P. W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nature Neuroscience*, *18*(9), 1230–1232. https://doi.org/10.1038/nn.4068

Kutlu, M. G., Zachry, J. E., Melugin, P. R., Cajigas, S. A., Chevee, M. F., Kelly, S. J., Kutlu, B., Tian, L., Siciliano, C. A., & Calipari, E. S. (2021). Dopamine release in the nucleus accumbens core signals perceived saliency. *Current Biology*, *31*(21), 4748-4761.e8. https://doi.org/10.1016/j.cub.2021.08.052

Lafferty, C. K., Yang, A. K., Mendoza, J. A., & Britt, J. P. (2020). Nucleus Accumbens Cell Type- and Input-Specific Suppression of Unproductive Reward Seeking. *Cell Reports*, *30*(11), 3729-3742.e3. https://doi.org/10.1016/j.celrep.2020.02.095

Langdon, A. J., Hathaway, B. A., Zorowitz, S., Harris, C. B. W., & Winstanley, C. A. (2019). Relative insensitivity to time-out punishments induced by win-paired cues in a rat gambling task. *Psychopharmacology*, *236*(8), 2543–2556. https://doi.org/10.1007/s00213-019-05308-x

Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., & Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological Medicine*, *50*(10), 1613–1622. https://doi.org/10.1017/S0033291719001570

LeGates, T. A., Kvarta, M. D., Tooley, J. R., Francis, T. C., Lobo, M. K., Creed, M. C., & Thompson, S. M. (2018). Reward behavior is regulated by the strength of hippocampus-nucleus accumbens synapses. *Nature*, *564*(7735), 258–262. https://doi.org/10.1038/s41586-018-0740-8

Lenth, R. V., Buerkner, P., Herve, M., Love, J., Riebl, H., & Singmann, H. (2021). *Emmeans: Estimated marginal means, aka least-squares means*. *CRAN*.

Lind, E. B., Sweis, B. M., Asp, A. J., Esguerra, M., Silvis, K. A., David Redish, A., & Thomas, M. J. (2023). A quadruple dissociation of reward-related behaviour in mice across excitatory inputs to the nucleus accumbens shell. *Communications Biology*, *6*(1), Article 1. https://doi.org/10.1038/s42003-023-04429-6

Lindenbach, D., Vacca, G., Ahn, S., Seamans, J. K., & Phillips, A. G. (2022). Optogenetic modulation of glutamatergic afferents from the ventral subiculum to the nucleus

accumbens: Effects on dopamine function, response vigor and locomotor activity. *Behavioural Brain Research*, *434*, 114028. https://doi.org/10.1016/j.bbr.2022.114028

Lipski, W. J., & Grace, A. A. (2013). Activation and inhibition of neurons in the hippocampal ventral subiculum by norepinephrine and locus coeruleus stimulation. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *38*(2), 285–292. https://doi.org/10.1038/npp.2012.157

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703–713. https://doi.org/10.1016/j.neuron.2005.05.002

Liu, X., & Carter, A. G. (2018). Ventral Hippocampal Inputs Preferentially Drive Corticocortical Neurons in the Infralimbic Prefrontal Cortex. *The Journal of Neuroscience*, *38*(33), 7351–7363. https://doi.org/10.1523/JNEUROSCI.0378-18.2018

Lladó-Pelfort, L., Santana, N., Ghisi, V., Artigas, F., & Celada, P. (2012). 5-HT1A receptor agonists enhance pyramidal cell firing in prefrontal cortex through a preferential action on GABA interneurons. *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(7), 1487–1497. https://doi.org/10.1093/cercor/bhr220

Lloyd, A., Roiser, J. P., Skeen, S., Freeman, Z., Badalova, A., Agunbiade, A., Busakhwe, C., DeFlorio, C., Marcu, A., Pirie, H., Saleh, R., Snyder, T., Fearon, P., & Viding, E. (2024). Reviewing explore/exploit decision-making as a transdiagnostic target for psychosis,

depression, and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *24*(5), 793–815. https://doi.org/10.3758/s13415-024-01186-9

Lobo, M. K., Covington, H. E., Chaudhury, D., Friedman, A. K., Sun, H., Damez-Werno, D., Dietz, D., Zaman, S., Koo, J. W., Kennedy, P. J., Mouzon, E., Mogri, M., Neve, R. L., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2010). Cell Type Specific Loss of BDNF Signaling Mimics Optogenetic Control of Cocaine Reward. *Science (New York, N.Y.)*, *330*(6002), 385–390. https://doi.org/10.1126/science.1188472

Loriaux, A. L., Roitman, J. D., & Roitman, M. F. (2011). Nucleus accumbens shell, but not core, tracks motivational value of salt. *Journal of Neurophysiology*, *106*(3), 1537–1544. https://doi.org/10.1152/jn.00153.2011

McGarry, L. M., & Carter, A. G. (2017). Prefrontal cortex drives distinct projection neurons in the basolateral amygdala. *Cell Reports*, *21*(6), 1426–1433. https://doi.org/10.1016/j.celrep.2017.10.046

Mingote, S., Amsellem, A., Kempf, A., Rayport, S., & Chuhma, N. (2019). Dopamine-Glutamate Neuron Projections to the Nucleus Accumbens Medial Shell and Behavioral Switching. *Neurochemistry International*, *129*, 104482. https://doi.org/10.1016/j.neuint.2019.104482

Mishchanchuk, K., Gregoriou, G., Qü, A., Kastler, A., Huys, Q. J. M., Wilbrecht, L., & MacAskill, A. F. (2024). Hidden state inference requires abstract contextual representations in the ventral hippocampus. *Science*, *386*(6724), 926–932. https://doi.org/10.1126/science.adq5874

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), Article 7759. https://doi.org/10.1038/s41586-019-1235-y

Moore, D. G., Valentini, G., Walker, S. I., & Levin, M. (2018). Inform: Efficient information-theoretic analysis of collective behaviors. *Frontiers in Robotics and AI*, *5*, 60.

Muir, J., Iyer, E. S., Tse, Y.-C., Sorensen, J., Wu, S., Eid, R. S., Cvetkovska, V., Wassef, K., Gostlin, S., Vitaro, P., Spencer, N. J., & Bagot, R. C. (2024). Sex-biased neural encoding of threat discrimination in nucleus accumbens afferents drives suppression of reward behavior. *Nature Neuroscience*, *27*(10), 1966–1976. https://doi.org/10.1038/s41593-024-01748-7

Muir, J., Tse, Y. C., Iyer, E. S., Biris, J., Cvetkovska, V., Lopez, J., & Bagot, R. C. (2020). Ventral Hippocampal Afferents to Nucleus Accumbens Encode Both Latent Vulnerability and Stress-Induced Susceptibility. *Biological Psychiatry*, *88*(11), 843–854. https://doi.org/10.1016/j.biopsych.2020.05.021

Nestler, E. J., & Carlezon, W. A. (2006). The Mesolimbic Dopamine Reward Circuit in Depression. *Biological Psychiatry*, *59*(12), 1151–1159. https://doi.org/10.1016/j.biopsych.2005.09.018

Nevin, J. A., & Mandell, C. (2017). Comparing positive and negative reinforcement: A fantasy experiment. *Journal of the Experimental Analysis of Behavior*, *107*(1), 34–38. https://doi.org/10.1002/jeab.237

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Nowend, K. L., Arizzi, M., Carlson, B. B., & Salamone, J. D. (2001). D1 or D2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. *Pharmacology Biochemistry and Behavior*, *69*(3), 373–382. https://doi.org/10.1016/S0091-3057(01)00524-X

Noworyta-Sokolowska, K., Kozub, A., Jablonska, J., Rodriguez Parkitna, J., Drozd, R., & Rygula, R. (2019). Sensitivity to negative and positive feedback as a stable and enduring behavioural trait in rats. *Psychopharmacology*, *236*(8), 2389–2403. https://doi.org/10.1007/s00213-019-05333-w

O'Donnell, P., & Grace, A. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: Hippocampal gating of prefrontal cortical input. *The Journal of Neuroscience*, *15*(5), 3622–3639. https://doi.org/10.1523/JNEUROSCI.15-05-03622.1995

O'Hearn, E., & Molliver, M. E. (1984). Organization of raphe-cortical projections in rat: A quantitative retrograde study. *Brain Research Bulletin*, *13*(6), 709–726. https://doi.org/10.1016/0361-9230(84)90232-6

Onge, J. R. S., Abhari, H., & Floresco, S. B. (2011). Dissociable Contributions by Prefrontal D1 and D2 Receptors to Risk-Based Decision Making. *Journal of Neuroscience*, *31*(23), 8625–8633. https://doi.org/10.1523/JNEUROSCI.1020-11.2011

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., & Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), Article 7643. https://doi.org/10.1038/nature21376

Otis, J. M., Zhu, M., Namboodiri, V. M. K., Cook, C. A., Kosyk, O., Matan, A. M., Ying, R., Hashikawa, Y., Hashikawa, K., Trujillo-Pisanty, I., Guo, J., Ung, R. L., Rodriguez-Romaguera, J., Anton, E. S., & Stuber, G. D. (2019). Paraventricular Thalamus Projection Neurons Integrate Cortical and Hypothalamic Signals for Cue-Reward Processing. *Neuron*, *103*(3), 423-431.e4. https://doi.org/10.1016/j.neuron.2019.05.018

Pan, W.-X., Brown, J., & Dudman, J. T. (2013). Neural signals of extinction in the inhibitory microcircuit of the ventral midbrain. *Nature Neuroscience*, *16*(1), 71–78. https://doi.org/10.1038/nn.3283

Pardo-Garcia, T. R., Garcia-Keller, C., Penaloza, T., Richie, C. T., Pickel, J., Hope, B. T., Harvey, B. K., Kalivas, P. W., & Heinsbroek, J. A. (2019). Ventral Pallidum Is the Primary Target for Accumbens D1 Projections Driving Cocaine Seeking. *The Journal of Neuroscience*, *39*(11), 2041–2051. https://doi.org/10.1523/JNEUROSCI.2822-18.2018

Parker, N. F., Baidya, A., Cox, J., Haetzel, L. M., Zhukovskaya, A., Murugan, M., Engelhard, B., Goldman, M. S., & Witten, I. B. (2022). Choice-selective sequences dominate in cortical relative to thalamic inputs to NAc to support reinforcement learning. *Cell Reports*, *39*(7). https://doi.org/10.1016/j.celrep.2022.110756

Parsons, M. P., Li, S., & Kirouac, G. J. (2007). Functional and anatomical connection between the paraventricular nucleus of the thalamus and dopamine fibers of the nucleus accumbens. *The Journal of Comparative Neurology*, *500*(6), 1050–1063. https://doi.org/10.1002/cne.21224

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: Affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476–483. https://doi.org/10.1016/j.tics.2012.07.009

Piantadosi, P. T., Yeates, D. C. M., & Floresco, S. B. (2020). Prefrontal cortical and nucleus accumbens contributions to discriminative conditioned suppression of reward-seeking. *Learning & Memory*, *27*(10), 429–440. https://doi.org/10.1101/lm.051912.120

Qi, J., Zhang, S., Wang, H.-L., Barker, D. J., Miranda-Barrientos, J., & Morales, M. (2016). VTA glutamatergic inputs to nucleus accumbens drive aversion by acting on GABAergic interneurons. *Nature Neuroscience*, *19*(5), 725–733. https://doi.org/10.1038/nn.4281

Rapanelli, M., Wang, W., Hurley, E., Feltri, M. L., Pittenger, C., Frick, L. R., & Yan, Z. (2023). Cholinergic neurons in the basal forebrain are involved in behavioral abnormalities associated with Cul3 deficiency: Role of prefrontal cortex projections in cognitive deficits. *Translational Psychiatry*, *13*(1), 1–11. https://doi.org/10.1038/s41398-023-02306-8

Reed, S. J., Lafferty, C. K., Mendoza, J. A., Yang, A. K., Davidson, T. J., Grosenick, L., Deisseroth, K., & Britt, J. P. (2018). Coordinated Reductions in Excitatory Input to the

Nucleus Accumbens Underlie Food Consumption. *Neuron*, *99*(6), 1260-1273.e4. https://doi.org/10.1016/j.neuron.2018.07.051

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory: Vol. Vol. 2*.

Richard, J. M., & Berridge, Kent. C. (2013). Prefrontal cortex modulates desire and dread generated by nucleus accumbens glutamate disruption. *Biological Psychiatry*, *73*(4), 360–370. https://doi.org/10.1016/j.biopsych.2012.08.009

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *The Journal of Neuroscience*, *29*(42), 13365–13376. https://doi.org/10.1523/JNEUROSCI.2572-09.2009

Sackett, D. A., Saddoris, M. P., & Carelli, R. M. (2017). Nucleus Accumbens Shell Dopamine Preferentially Tracks Information Related to Outcome Value of Reward. *eNeuro*, *4*(3). https://doi.org/10.1523/ENEURO.0058-17.2017

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015a). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. *The Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015b). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal

Complementary Signals for Error Prediction and Incentive Motivation. *Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Salamone, J. D., Correa, M., Yang, J.-H., Rotolo, R., & Presby, R. (2018). Dopamine, Effort-Based Choice, and Behavioral Economics: Basic and Translational Research. *Frontiers in Behavioral Neuroscience*, *12*, 52. https://doi.org/10.3389/fnbeh.2018.00052

Santana, N., & Artigas, F. (2017). Laminar and Cellular Distribution of Monoamine Receptors in Rat Medial Prefrontal Cortex. *Frontiers in Neuroanatomy*, *11*, 87. https://doi.org/10.3389/fnana.2017.00087

Sayegh, F. J. P., Mouledous, L., Macri, C., Pi Macedo, J., Lejards, C., Rampon, C., Verret, L., & Dahan, L. (2024). Ventral tegmental area dopamine projections to the hippocampus trigger long-term potentiation and contextual learning. *Nature Communications*, *15*(1), 4100. https://doi.org/10.1038/s41467-024-47481-4

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Scott, A., Palmer, D., Newell, B., Lin, I., Cayton, C. A., Paulson, A., Remde, P., & Richard, J. M. (2023). Ventral Pallidal GABAergic Neuron Calcium Activity Encodes Cue-Driven

Reward Seeking and Persists in the Absence of Reward Delivery. *Journal of Neuroscience*, *43*(28), 5191–5203. https://doi.org/10.1523/JNEUROSCI.0013-23.2023

Segal, M. (1990). Serotonin attenuates a slow inhibitory postsynaptic potential in rat hippocampal neurons. *Neuroscience*, *36*(3), 631–641. https://doi.org/10.1016/0306-4522(90)90006-P

Smith, R. J., Lobo, M. K., Spencer, S., & Kalivas, P. W. (2013). Cocaine-induced adaptations in D1 and D2 accumbens projection neurons (a dichotomy not necessarily synonymous with direct and indirect pathways). *Current Opinion in Neurobiology*, *23*(4), 546–552. https://doi.org/10.1016/j.conb.2013.01.026

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., Sousa, N., & Rodrigues, A. J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. *Nature Communications*, *7*(1), 11829. https://doi.org/10.1038/ncomms11829

Soares-Cunha, C., Coimbra, B., Sousa, N., & Rodrigues, A. J. (2016). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neuroscience & Biobehavioral Reviews*, *68*, 370–386. https://doi.org/10.1016/j.neubiorev.2016.05.021

Soares-Cunha, C., de Vasconcelos, N. A. P., Coimbra, B., Domingues, A. V., Silva, J. M., Loureiro-Campos, E., Gaspar, R., Sotiropoulos, I., Sousa, N., & Rodrigues, A. J. (2020). Nucleus accumbens medium spiny neurons subtypes signal both reward and aversion. *Molecular Psychiatry*, *25*(12), 3241–3255. https://doi.org/10.1038/s41380-019-0484-3

Soares-Cunha, C., & Heinsbroek, J. A. (2023). Ventral pallidal regulation of motivated behaviors and reinforcement. *Frontiers in Neural Circuits*, *17*, 1086053. https://doi.org/10.3389/fncir.2023.1086053

Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., & Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell*, *184*(10), 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

St-Amand, D., Sheldon, S., & Otto, A. R. (2018). Modulating Episodic Memory Alters Risk Preference during Decision-making. *Journal of Cognitive Neuroscience*, *30*(10), 1433–1441. https://doi.org/10.1162/jocn_a_01253

Stern, E. A., Kincaid, A. E., & Wilson, C. J. (1997). Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons in vivo. *Journal of Neurophysiology*, *77*(4), 1697–1715. https://doi.org/10.1152/jn.1997.77.4.1697

Stopper, C. M., & Floresco, S. B. (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(1), 97–112. https://doi.org/10.3758/s13415-010-0015-9

Stuber, G. D., Sparta, D. R., Stamatakis, A. M., van Leeuwen, W. A., Hardjoprajitno, J. E., Cho, S., Tye, K. M., Kempadoo, K. A., Zhang, F., Deisseroth, K., & Bonci, A. (2011). Amygdala to nucleus accumbens excitatory transmission facilitates reward seeking. *Nature*, *475*(7356), 377–380. https://doi.org/10.1038/nature10194

Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making. *Neuron*, *66*(3), 449–460. https://doi.org/10.1016/j.neuron.2010.03.033

Syed, E. C. J., Grima, L. L., Magill, P. J., Bogacz, R., Brown, P., & Walton, M. E. (2016). Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nature Neuroscience*, *19*(1), Article 1. https://doi.org/10.1038/nn.4187

Szczypka, M. S., Kwok, K., Brot, M. D., Marck, B. T., Matsumoto, A. M., Donahue, B. A., & Palmiter, R. D. (2001). Dopamine Production in the Caudate Putamen Restores Feeding in Dopamine-Deficient Mice. *Neuron*, *30*(3), 819–828. https://doi.org/10.1016/S0896-6273(01)00319-1

Tecuapetla, F., Patel, J. C., Xenias, H., English, D., Tadros, I., Shah, F., Berlin, J., Deisseroth, K., Rice, M. E., Tepper, J. M., & Koos, T. (2010). Glutamatergic Signaling by Mesolimbic Dopamine Neurons in the Nucleus Accumbens. *The Journal of Neuroscience*, *30*(20), 7105–7110. https://doi.org/10.1523/JNEUROSCI.0265-10.2010

Thibeault, K. C., Gunes Kutlu, M., Sanders, C., & Calipari, E. S. (2019). Cell-type and projection-specific dopaminergic encoding of aversive stimuli in addiction. *Brain Research*, *1713*, 1–15. https://doi.org/10.1016/j.brainres.2018.12.024

Thierry, A. M., Gioanni, Y., Dégénétais, E., & Glowinski, J. (2000). Hippocampo-prefrontal cortex pathway: Anatomical and electrophysiological characteristics. *Hippocampus*, *10*(4), 411–419. https://doi.org/10.1002/1098-1063(2000)10:4<411::AID-HIPO7>3.0.CO;2-A

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*, 212–222. https://doi.org/10.2307/1415413

Titulaer, J., Björkholm, C., Feltmann, K., Malmlöf, T., Mishra, D., Bengtsson Gonzales, C., Schilström, B., & Konradsson-Geuken, Å. (2021). The Importance of Ventral Hippocampal Dopamine and Norepinephrine in Recognition Memory. *Frontiers in Behavioral Neuroscience*, *15*, 667244. https://doi.org/10.3389/fnbeh.2021.667244

Tremblay, R., Lee, S., & Rudy, B. (2016). GABAergic interneurons in the neocortex: From cellular properties to circuits. *Neuron*, *91*(2), 260–292. https://doi.org/10.1016/j.neuron.2016.06.033

Tzschentke, T. M., & Schmidt, W. J. (2000). Functional relationship among medial prefrontal cortex, nucleus accumbens, and ventral tegmental area in locomotion and reward. *Critical Reviews in Neurobiology*, *14*(2), 131–142.

van Holstein, M., MacLeod, P. E., & Floresco, S. B. (2020). Basolateral amygdala—Nucleus accumbens circuitry regulates optimal cue-guided risk/reward decision making. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *98*, 109830. https://doi.org/10.1016/j.pnpbp.2019.109830

Verharen, J. P. H., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2019). Differential contributions of striatal dopamine D1 and D2 receptors to component processes of

value-based decision making. *Neuropsychopharmacology*, *44*(13), 2195–2204. https://doi.org/10.1038/s41386-019-0454-0

Vierock, J., Rodriguez-Rozada, S., Dieter, A., Pieper, F., Sims, R., Tenedini, F., Bergs, A. C. F., Bendifallah, I., Zhou, F., Zeitzschel, N., Ahlbeck, J., Augustin, S., Sauter, K., Papagiakoumou, E., Gottschalk, A., Soba, P., Emiliani, V., Engel, A. K., Hegemann, P., & Wiegert, J. S. (2021). BiPOLES is an optogenetic tool developed for bidirectional dual-color control of neurons. *Nature Communications*, *12*(1), 4527. https://doi.org/10.1038/s41467-021-24759-5

Voorn, P., Vanderschuren, L. J. M. J., Groenewegen, H. J., Robbins, T. W., & Pennartz, C. M. A. (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends in Neurosciences*, *27*(8), 468–474. https://doi.org/10.1016/j.tins.2004.06.006

Wan, J., Peng, W., Li, X., Qian, T., Song, K., Zeng, J., Deng, F., Hao, S., Feng, J., Zhang, P., Zhang, Y., Zou, J., Pan, S., Shin, M., Venton, B. J., Zhu, J. J., Jing, M., Xu, M., & Li, Y. (2021). A genetically encoded sensor for measuring serotonin dynamics. *Nature Neuroscience*, *24*(5), 746–752. https://doi.org/10.1038/s41593-021-00823-7

Wang, A. Y., Miura, K., & Uchida, N. (2013). The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nature Neuroscience*, *16*(5), 639–647. https://doi.org/10.1038/nn.3377

Warlow, S. M., Singhal, S. M., Hollon, N. G., Faget, L., Dowlat, D. S., Zell, V., Hunker, A. C., Zweifel, L. S., & Hnasko, T. S. (2024). Mesoaccumbal glutamate neurons drive reward via glutamate release but aversion via dopamine co-release. *Neuron*, *112*(3), 488-499.e5. https://doi.org/10.1016/j.neuron.2023.11.002

Wassum, K. M., Ostlund, S. B., & Maidment, N. T. (2012). Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task. *Biological Psychiatry*, *71*(10), 846–854. https://doi.org/10.1016/j.biopsych.2011.12.019

Wenzel, J. M., Zlebnik, N. E., Patton, M. H., Smethells, J. R., Ayvazian, V. M., Dantrassy, H. M., Zhang, L.-Y., Mathur, B. N., & Cheer, J. F. (2023). Selective chemogenetic inactivation of corticoaccumbal projections disrupts trait choice impulsivity. *Neuropsychopharmacology*, *48*(12), Article 12. https://doi.org/10.1038/s41386-023-01604-5

West, E. A., & Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after Reinforcer Devaluation. *The Journal of Neuroscience*, *36*(4), 1128–1139. https://doi.org/10.1523/JNEUROSCI.2976-15.2016

White, N. M. (1989). Reward or reinforcement: What's the difference? *Neuroscience & Biobehavioral Reviews*, *13*(2), 181–186. https://doi.org/10.1016/S0149-7634(89)80028-4

Willmore, L., Cameron, C., Yang, J., Witten, I. B., & Falkner, A. L. (2022). Behavioural and dopaminergic signatures of resilience. *Nature*, *611*(7934), 124–132. https://doi.org/10.1038/s41586-022-05328-2

Wilson, C. (2008). Up and down states. *Scholarpedia Journal*, *3*(6), 1410.

Wilson, C. J., & Groves, P. M. (1981). Spontaneous firing patterns of identified spiny neurons in the rat neostriatum. *Brain Research*, *220*(1), 67–80. https://doi.org/10.1016/0006-8993(81)90211-0

Wilson, C., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *The Journal of Neuroscience*, *16*(7), 2397–2410. https://doi.org/10.1523/JNEUROSCI.16-07-02397.1996

Wu, J.-L., Li, Z.-M., Chen, H., Chen, W.-J., Hu, N.-Y., Jin, S.-Y., Li, X.-W., Chen, Y.-H., Yang, J.-M., & Gao, T.-M. (2024). Distinct septo-hippocampal cholinergic projections separately mediate stress-induced emotional and cognitive deficits. *Science Advances*, *10*(45), eado1508. https://doi.org/10.1126/sciadv.ado1508

Yamaguchi, T., Wang, H.-L., Li, X., Ng, T. H., & Morales, M. (2011). Mesocorticolimbic glutamatergic pathway. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(23), 8476–8490. https://doi.org/10.1523/JNEUROSCI.1598-11.2011

Yang, H., Jong, J. W. de, Tak, Y., Peck, J., Bateup, H. S., & Lammel, S. (2018). Nucleus Accumbens Subnuclei Regulate Motivated Behavior via Direct Inhibition and Disinhibition of VTA Dopamine Subpopulations. *Neuron*, *97*(2), 434-449.e4. https://doi.org/10.1016/j.neuron.2017.12.022

Yoo, J. H., Zell, V., Gutierrez-Reed, N., Wu, J., Ressler, R., Shenasa, M. A., Johnson, A. B., Fife, K. H., Faget, L., & Hnasko, T. S. (2016). Ventral tegmental area glutamate neurons co-release GABA and promote positive reinforcement. *Nature Communications*, *7*(1), 13697. https://doi.org/10.1038/ncomms13697

Yoshida, K., Drew, M. R., Kono, A., Mimura, M., Takata, N., & Tanaka, K. F. (2021). Chronic social defeat stress impairs goal-directed behavior through dysregulation of ventral

hippocampal activity in male mice. *Neuropsychopharmacology*, *46*(9), 1606–1616. https://doi.org/10.1038/s41386-021-00990-y

Yoshida, K., Drew, M. R., Mimura, M., & Tanaka, K. F. (2019). Serotonin-mediated inhibition of ventral hippocampus is required for sustained goal-directed behavior. *Nature Neuroscience*, *22*(5), Article 5. https://doi.org/10.1038/s41593-019-0376-5

Yoshida, K., Tsutsui-Kimura, I., Kono, A., Yamanaka, A., Kobayashi, K., Watanabe, M., Mimura, M., & Tanaka, K. F. (2020). Opposing Ventral Striatal Medium Spiny Neuron Activities Shaped by Striatal Parvalbumin-Expressing Interneurons during Goal-Directed Behaviors. *Cell Reports*, *31*(13), 107829. https://doi.org/10.1016/j.celrep.2020.107829

Yu, Z., Guindani, M., Grieco, S. F., Chen, L., Holmes, T. C., & Xu, X. (2022). Beyond t test and ANOVA: Applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron*, *110*(1), 21–35. https://doi.org/10.1016/j.neuron.2021.10.030

Zhu, Y., Wienecke, C. F. R., Nachtrab, G., & Chen, X. (2016). A thalamic input to the nucleus accumbens mediates opiate dependence. *Nature*, *530*(7589), 219–222. https://doi.org/10.1038/nature16954

Zhuo, Y., Luo, B., Yi, X., Dong, H., Miao, X., Wan, J., Williams, J. T., Campbell, M. G., Cai, R., Qian, T., Li, F., Weber, S. J., Wang, L., Li, B., Wei, Y., Li, G., Wang, H., Zheng, Y., Zhao, Y., … Li, Y. (2024). Improved green and red GRAB sensors for monitoring dopaminergic activity in vivo. *Nature Methods*, *21*(4), 680–691. https://doi.org/10.1038/s41592-023-02100-w

A role for GABAergic neurons in outcome integration?

The previous chapter describes a novel mechanism by which mPFC and vHip projections to the NAc integrate outcomes over multiple trials to tune engagement to recent history of reward. It further establishes that uncertainty gates the expression of this mechanisms in vHip, thus defining a process by which uncertainty interacts with outcome in order to continuously modulate behavior. The defining feature of this mechanism is reward-mediated suppression, a surprising finding and counterintuitive based on normative descriptions of neural responses to reward. One potential explanation is that a population of inhibitory cells may be activated to inhibit projection neurons originating in mPFC and vHip thus enabling the reward-associated suppression, and by extension, reward integration described in the previous chapter. To investigate this possibility, the following chapter examines how inhibitory cell populations in mPFC and vHip integrate reward with an examination of how behavioral state modulates these patterns of encoding. This work further contextualizes the patterns of outcome integration observed in mPFC and vHip projections to NAc and suggests the existence of a common mechanism for the modulation of both local and NAc-projecting populations in mPFC and vHip, gated by behavioral state.

Chapter 3

Behavioral state gates reward integration in GABAergic neurons in medial prefrontal cortex and ventral hippocampus in male mice

Eshaan S. Iyer[1], Vedrana Cvetkovska[2], Joëlle Lopez[2], Becket Ebitz[3], Rosemary C. Bagot[2,4]

1. Integrated Program in Neuroscience, McGill University, Montréal, Canada

2. Department of Psychology, McGill University, Montréal, QC, Canada

3. Département de neurosciences, Université de Montréal, Montréal, QC, Canada

4. Ludmer Centre for Neuroinformatics and Mental Health, Montréal, QC, Canada

Abstract

Outcome encoding in reward-motivated behavior is observed in principal neurons across various brain regions. Prevailing views of neural encoding of reward focus on excitatory neurons as the primary players with GABAergic interneurons relegated to a supporting role, overlooking the possibility that GABAergic interneurons might independently encode information. This could have important implications for understanding neural processing of reward. Here we examined reward encoding in GABAergic interneurons in medial prefrontal cortex (mPFC) and ventral hippocampus (vHip), regions in which glutamatergic neurons integrate outcomes to modulate task engagement. We used dual-site fiber photometry in an operant reward task to simultaneously record from GABAergic interneuron populations in mPFC and vHip in male mice. We identify a common motif of reward integration in GABAergic neurons in both regions, whereby relative suppression of neural activity maps onto outcome history, and this is sensitive to behavioral state. Using a hidden Markov model to decompose behavior into bouts of exploration and exploitation, we find that exploration gates outcome integration in both mPFC and vHip GABAergic neurons. Globally, interneuron activity in both regions predicted engagement independent of state, but that choice prediction was state-dependent and distributed between mPFC and vHip. Overall, these findings establish a distinct role for GABAergic neurons in outcome encoding and state-sensitive modulation of reward-motivated behavior.

Introduction

The ability to integrate information about outcomes is essential to adaptive behavior. Information must be continuously processed across multiple timescales, integrating outcome history to tune motivated behavior. Despite considerable interest in the role of excitatory neuron populations in outcome encoding and reward-motivated behavior, GABAergic interneurons remain largely unexplored (Otis et al., 2017; Parker et al., 2022; Spellman et al., 2021; Sul et al., 2010; Yoshida et al., 2019, 2021). While glutamatergic neurons encode and propagate information across brain circuits, GABAergic interneurons are considered to play a supporting role, tuning the dynamics of these excitatory neurons and shaping local network activity (Klausberger & Somogyi, 2008; Tremblay et al., 2016). However, several recent studies challenge this view, finding that GABAergic neurons in reward-processing regions such the ventral tegmental area (VTA) and ventral pallidum (VP) also encode reward-relevant information (Bouarab et al., 2019; Pan et al., 2013; Scott et al., 2023).

We recently found that glutamatergic neurons in both the medial prefrontal cortex (mPFC) and ventral hippocampus (vHip) integrate outcome histories to modulate reward-motived behavior (Iyer et al., 2024). However, whether GABAergic interneurons in these regions also play a role in encoding and integrating reward-relevant information remains an outstanding question. To address this, we used *in vivo* fiber photometry to record activity in GABAergic populations simultaneously in mPFC and vHip during reward-guided choice in a two-armed bandit task. We considered that GABAergic

interneurons might also encode reward-relevant information, or alternatively, that they might provide only broader modulation of local activity, in which case, we expected to observe general modulation of activity by higher level features such as task structure or behavioral state.

Surprisingly, we identify a unique role of GABAergic neurons in supporting outcome integration in mPFC and vHip. Using trial-by-trial modeling of neural activity, we identified reward-mediated suppression as a mechanism for integrating outcome information across trials in both mPFC and vHip, revealing a fundamental role of GABAergic neurons in reward encoding. To determine whether these neurons may also play a broader role in modulation by higher-order features, we then explored how the nature of outcome encoding in GABAergic neurons is influenced by task structure, finding that the fidelity of outcome encoding is sensitive to varying task demands. Using hidden Markov modeling to estimate latent behavioral states, we find that exploration gates outcome integration in both mPFC and vHip. While both mPFC and vHip interneuron activity predict behavioral engagement across behavioral state, effects on choice are state-dependent and distinct between regions. Activity in mPFC (but not vHip) GABAergic interneurons predicted choice under exploration states while activity in vHip (but not mPFC) GABAergic neurons predicted choice under exploitation. These findings reveal a previously unappreciated role for GABAergic neurons in outcome integration for state-sensitive tuning of reward-motivated behavior with this functional segregation providing a potential mechanism for more efficient and specialized behavioral control.

Main Text

GABAergic mPFC and vHip neurons encode outcomes in a probabilistically rewarded environment

To assess population-level neural encoding by GABAergic neurons in mPFC and vHip under matched conditions and trial histories, we injected AAV-mDlx-GCaMP6f to target GCaMP6f expression to GABAergic neurons (Fig. 1D) and implanted optic fibers in mPFC and vHip to record Ca2+-associated fluorescence (Fig. 1A-C) while mice engaged in reward-guided choice (Fig 1E). We trained mice in a two-lever probabilistic reward learning task (i.e. a two-armed bandit task) in which lever pressing probabilistically earns a chocolate milk reward (Fig. 1E). Following each lever press, one of two different auditory cues signaled trial outcome (rewarded, unrewarded) and start of the inter-trial interval (ITI). To maintain a dynamic environment with robustly encountered rewarded and unrewarded outcomes, levers were probabilistically rewarded on 80% or 20% of presses with probabilities switched after five consecutive responses on the high probability lever. Mice robustly engaged with this task, encountering high numbers of rewarded and unrewarded trials, and failing to respond (omitting) on a low number of trials (Fig. 1F).

Reward strongly suppressed activity in mPFC and vHip GABAergic neurons, with some differences in the evolution of signal across the ITI (Fig. 1G, I). In mPFC, a peak following the lever press and outcome delivery is followed by gradually emerging reward-associated suppression across the ITI. In vHip, suppression emerges after lever press and outcome delivery and is sustained following reward with activity gradually

increasing following unrewarded outcomes. We focused analysis on the end of the ITI (8-10 sec after lever press) when trial outcome has been integrated prior to next trial start. By ITI end, reward robustly suppressed activity in mPFC GABAergic neurons (Fig. 1H). Reward also robustly suppressed activity in vHip GABAergic neurons (Fig. 1J). This indicates that outcome encoding emerges across the ITI with reward suppressing activity of GABAergic neurons in mPFC and vHip.

To explore modulation by other task factors, we examined neural encoding time-locked to decision-relevant behaviors. First, we examined neural encoding time-locked to licking and did not observe clear encoding of licks (Fig S1 A,B). We then time-locked neural activity to the lever press, splitting neural traces by the identity of the upcoming choice relative to the prior choice (same as the previous choice, or a shift from the previous choice) (Fig. S1 C, D). Given dramatic reward-induced suppression in GABAergic activity, we included the previous trial as a covariate to account for shifting baselines. vHip GABAergic activity does not encode upcoming choice: the primary source of modulation prior to the lever-press is the previous outcome. However, following a rewarded, but not an unrewarded outcome, lower mPFC choice shifting associates with lower mPFC activity. We also time-locked neural activity to lever press, splitting neural traces by the fixed identity of the lever as ipsilateral or contralateral to the fiber implant (Fig. S1 E,F). We find that mPFC GABAergic activity does not encode the identity of the upcoming lever choice: the primary source of modulation prior to the lever-press is the previous outcome. However, following a rewarded choice, increased vHip GABAergic activity precedes a contralateral choice, whereas, following an unrewarded choice, increased vHip GABAergic activity predicts an ipsilateral choice.

This points to partially dissociable roles of mPFC and vHip in choice behavior with mPFC tracking choice identity relative to prior choice and vHip tracking choice relative to fixed lever identity.

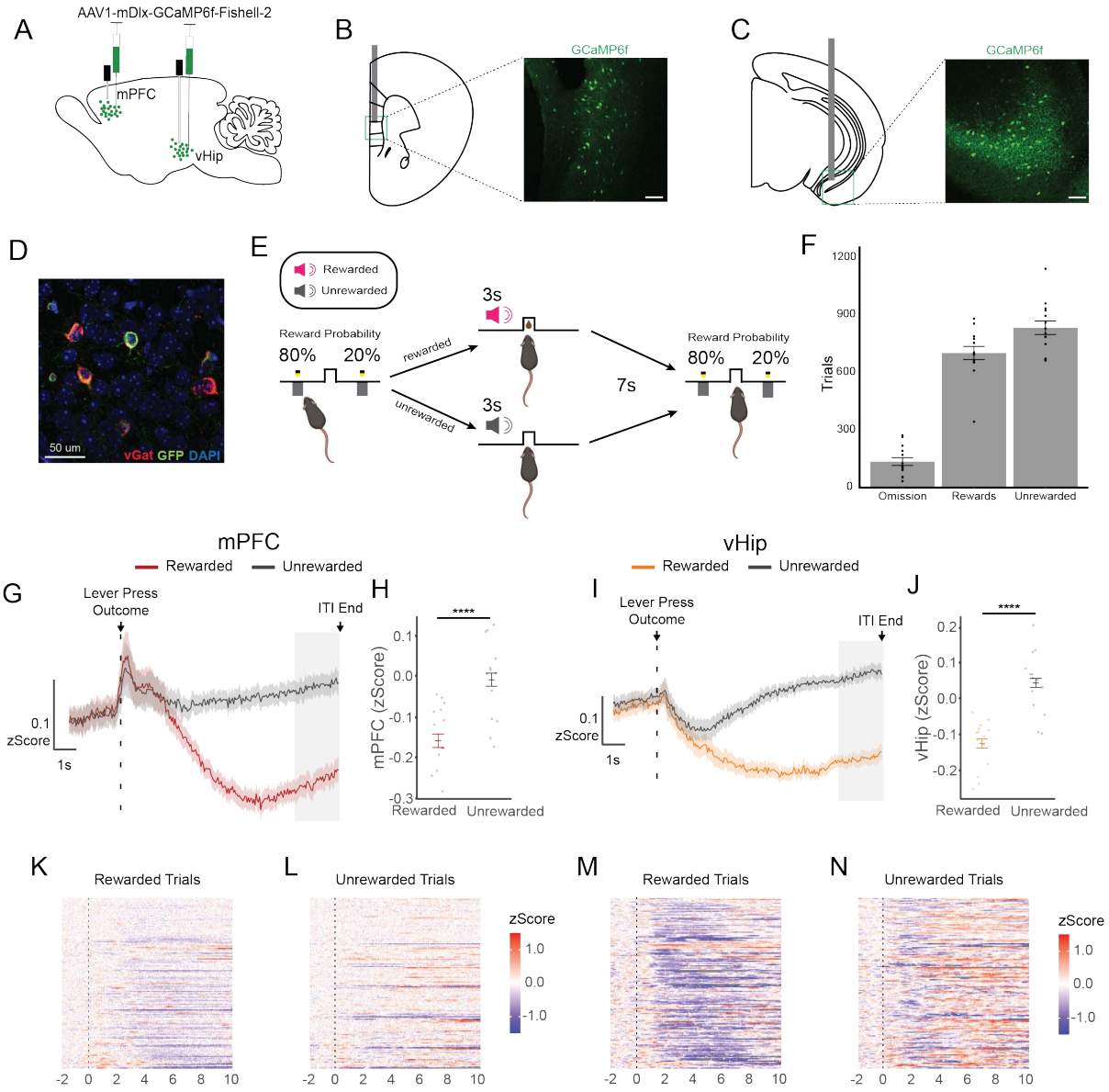


Figure 1. mPFC and vHip similarly encode reward in a probabilistically rewarded environment. *(A) jGCaMP6f under a mDlx promoter is injected into the medial prefrontal cortex (mPFC) and ventral hippocampus (vHip) and optic fibers implanted above each injection site to simultaneously probe neural activity indicated by Ca^{2+}-associated fluorescence changes in (B) mPFC and (C) vHip GABAergic populations as perform a two-armed bandit task. (D) Examining the expression of vGat and gCaMP shows that gCaMP-expressing cells (green) were also vGat+ (red) confirming targeting*

*to GABAergic neurons. (E) Schematic of two-armed bandit task. Mice (n=14) lever press in a two-lever task in which one lever is rewarded with chocolate milk on 80% of trials, and the other on 20%. Following a lever press, levers retract, and auditory cues signal outcome and start of a 10 sec inter-trial interval (ITI). Contingencies switch after five consecutive responses on the high probability lever. (F) Mice robustly engage with the task, experiencing high numbers of rewarded and unrewarded trials, and low numbers of omissions. Estimated mean GABAergic (G) mPFC and (I) vHip activity across all rewarded and unrewarded trials. Analysis focused on 8-10 sec after lever press (ITI end). (H) At ITI end, mPFC activity is suppressed by rewarded outcomes ($Z=13.033$, $p<0.0001$). (J) At ITI end, vHip activity is suppressed by rewarded outcomes ($Z=15.757$, $p<0.0001$). Heatmap of mPFC activity to (K) rewarded outcomes and (L) unrewarded outcomes in a representative animal across one session. Heatmap of vHip activity to (M) rewarded outcomes and (N) unrewarded outcomes in a representative animal across one session. Error bars represent SEM around the estimated mean. **** $p<0.0001$*

mPFC and vHip GABAergic neurons similarly integrate outcome history

Given that reward-induced suppression lasted throughout the ITI, and that previous reward most robustly modulates neural activity into upcoming trials, we speculated that this enduring modulation integrates reward information across successive trials. To test this, we sorted trials by both prior and current outcome, identifying trial sequences that were rewarded then rewarded (R→R), rewarded then unrewarded (R→U), unrewarded then rewarded (U→R), and unrewarded then unrewarded (U→U). We then compared neural activity across the ITI on the most recent trial to determine how prior outcome modulates outcome encoding on the current trial. Both previous and current outcome modulate mPFC GABAergic neuron activity (Fig. 2A). Following a given trial (t -1), reward suppresses mPFC GABAergic activity (Fig. 2B) effectively resetting the baseline for the next trial. Reward on the subsequent trial (t0) similarly suppresses mPFC by ITI end, compounding the effect of the previous outcome. When mice are unrewarded on the subsequent trial (t0), suppression of mPFC by prior

reward is maintained through ITI end. Similarly, when mice are rewarded on trial t0, suppression of mPFC by prior reward is maintained through ITI end (Fig 2C). This suggests that reward enduringly suppresses mPFC GABAergic neuron activity and that, in the absence of subsequent reward, this suppression slowly dissipates.

We then examined if vHip GABAergic neurons similarly integrates outcomes (Fig. 2D). Following a given trial (t-1), reward suppresses vHip activity (Fig. 2E). As with mPFC, this resets the baseline for the next trial (t0), wherein reward suppresses vHip, compounding the effect of prior outcome. When the subsequent trial (t0) is unrewarded, suppression of vHip GABAergic activity by prior reward is maintained through ITI end. Similarly, when mice are rewarded on trial t0, suppression of vHip by prior reward is maintained through ITI end (Fig. 2F). Together, this shows that mPFC and vHip GABAergic neurons similarly integrate outcomes across trials. In both circuits, reward suppresses neural activity and activity gradually increases following subsequent unrewarded outcomes, such that, by ITI end, the relative degree of suppression represents an integrated reward outcome history and describes the current reward statistics of the environment.

To examine if GABAergic mPFC and vHip population outcome encoding is also sensitive to higher order features, such as task demands, we recorded neural activity while degrading task requirements to sequentially eliminate choice and action. We first eliminated choice, extending only a single lever while maintaining the requirement to press to elicit an outcome. To hold outcome experience constant, the specific sequence of rewarded and unrewarded outcomes was yoked to each animal's prior performance on the two-lever task (Fig. 3A). In the absence of choice, GABAergic mPFC neurons

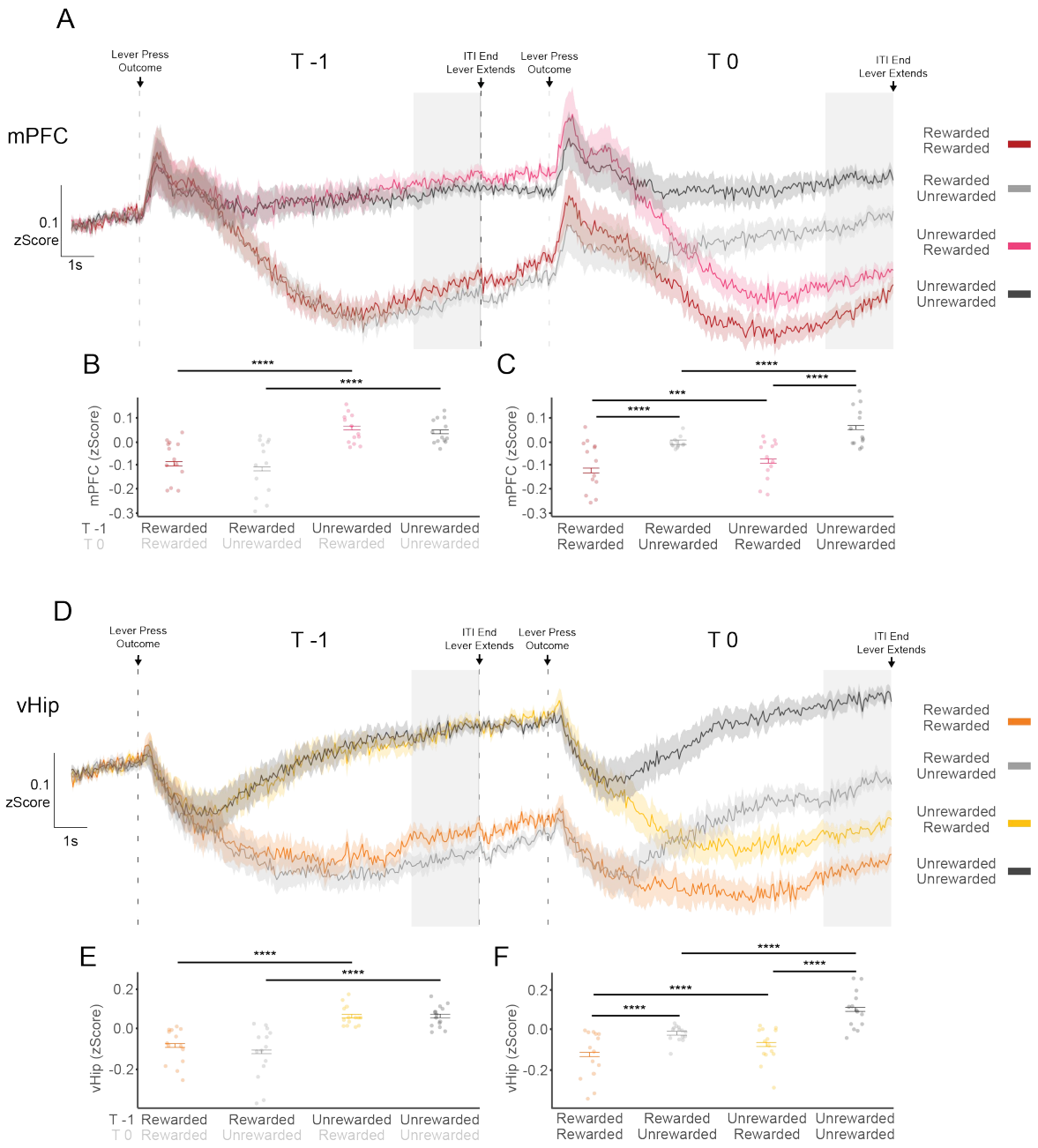


Figure 2. mPFC and vHip similarly integrate reward history. *(A) Estimated mean mPFC activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end). (B) On trial t-1, mPFC activity is significantly suppressed by reward (U→U vs R→U: Z=17.441, p<0.0001; U→R vs R→R: Z=13.442, p<0.0001). (C) On the subsequent trial, t0, mPFC activity is significantly suppressed by current reward (U→U vs U→R: Z=-14.554, p<0.0001; R→U vs R→R: Z=-11.182, p<0.0001). When trial t0 is unrewarded, mPFC activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 6.405, p<0.0001). When trial t0 is rewarded, mPFC activity remains significantly suppressed by reward experienced on the previous*

*trial, t-1, (U→R vs R→R: Z= 3.747, p=0.0003) (D) Estimated mean vHip activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pail (E) On trial t-1, vHip activity is significantly suppressed by reward (U→U vs R→U: Z=20.942, p<0.0001; U→R vs R→R: Z=15.901, p<0.0001). (F) On the subsequent trial, t0, vHip activity is significantly suppressed by current reward (U→U vs U→R: Z=-18.885, p<0.0001; U→U vs R→R: Z=-10.623, p<0.0001). When trial t0 is unrewarded, vHip activity remains suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 12.806, p<0.0001). When trial t0 is rewarded, vHip activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→R vs R→R: Z= 4.990, p<0.0001). Error bars represent SEM around the estimated mean. *** p<0.001 **** p<0.0001*

continued to encode previous and current outcome (Fig. 3B). On trial t0, by ITI end, current and prior outcomes were encoded, as in the two-lever task (Fig. 3C, Fig. 2C). Examining activity in GABAergic vHip neurons in the one-lever task also revealed largely similar outcome-mediated modulation (Fig. 3D, Fig. 2D). At ITI end, prior and current outcomes were integrated, similar to the two-lever task (Fig. 3E, Fig. 2F). This reveals that GABAergic neurons in both mPFC and vHip maintain similarly graded representations of reward history independent of choice requirements.

We then eliminated response requirements, removing both levers with rewards delivered in a choice-free response-free task. Trials continued to be signaled by cue-lights, but without lever extension and outcomes were passively delivered yoked to each animal's individual performance on the full two-lever task (Fig. 3F). In the absence of response requirements, traces from GABAergic neurons in mPFC (Fig. 3G) and vHip (Fig. 3I) are noisier yet appear to continue to integrate outcome history, albeit with lower fidelity. On trial t0, by ITI end, GABAergic mPFC neurons encode current outcome and trend towards encoding previous outcome (Fig. 3H). Examining activity in GABAergic vHip neurons revealed encoding of current, but not previous outcome, pointing to increased sensitivity of vHip to task demands compared to mPFC (Fig. 3J). This

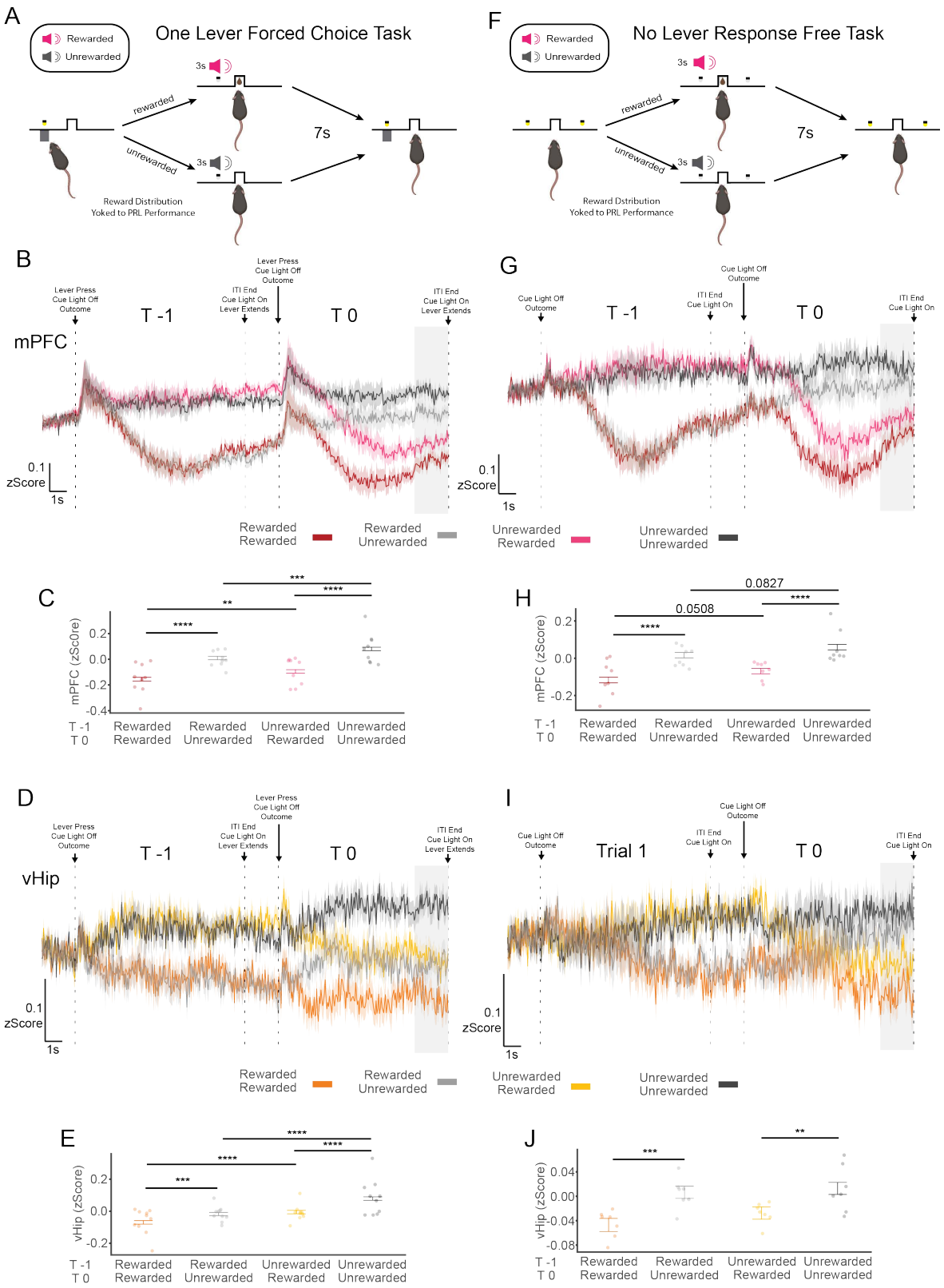


Figure 3. Outcome integration in mPFC and vHip is independent of task requirements. *(A) Schematic of one-lever forced choice task in which lever presses are*

rewarded on a schedule yoked to each animal's individual performance in the final three days of the two-armed bandit task. Following a lever press, levers retract, and auditory cues signal outcome and start of a 10 sec inter-trial interval (ITI). (B) Estimated mean mPFC activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end). (C) On trial t0, mPFC activity is significantly suppressed by current reward (U→U vs U→R: Z=-9.927, p<0.0001; R→U vs R→R: Z=-9.109, p<0.0001). When trial t0 is unrewarded, mPFC activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 4.033, p=0.0002). When trial t0 is rewarded, mPFC activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→R vs R→R: Z= 4.033, p=0.0016) (D) Estimated mean vHip activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end). (E) On trial t0, vHip activity is significantly suppressed by current reward (U→U vs U→R: Z=-6.452, p<0.0001; R→U vs R→R: Z=-3.924, p=0.0002). When trial t0 is unrewarded, vHip activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 7.406, p<0.0001). When trial t0 is rewarded, vHip activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→R vs R→R: Z= 4.862, p<0.0001). (F) Schematic of no lever response free task. Mice are allowed to collect rewards delivered on a schedule yoked to each animal's individual trial statistics (latency and outcome) of the two-armed bandit task. Trial structure is signaled by cue-light illumination and after a predetermined delay auditory cues signal outcome and start of a 10 sec ITI. (G) Estimated mean mPFC activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end). (H) On trial t0, mPFC activity is significantly suppressed by current reward (U→U vs U→R: Z=-6.063, p<0.0001; R→U vs R→R: Z=-5.865, p<0.0001). When trial t0 is unrewarded, mPFC activity trends towards remaining suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 2.031, p=0.083). When trial t0 is rewarded, mPFC activity trends towards remaining suppressed by reward experienced on the previous trial, t-1, (U→R vs R→R: Z= 2.230, p=0.0508) (I) Estimated mean vHip activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end). (J) On trial t0, vHip activity is significantly suppressed by current reward (U→U vs U→R: Z=-2.861, p=0.0084; R→U vs R→R: Z=-3.668, p=0.0005).

suggests that outcome integration is a core function of GABAergic populations that continues across changing demands but that the fidelity of this integration diminishes as the utility of outcome tracking diminishes with simplified task demands.

Exploration gates outcome integration in GABAergic mPFC and vHip populations

We observed that degrading task requirements injected noise into the signal, reducing the fidelity of outcome encoding, yet did not fundamentally alter the information encoded by these signals. We then considered if GABAergic neurons might be sensitive to modulation by other higher-order factors. Previous studies suggest that in probabilistically rewarded learning tasks, animals often switch between two behavioral states. In an exploration state animals sample the different options, whereas in an exploitation state, animals perseverate, or *exploit*, one option. These studies report differences in the rate of reward learning between states, but the neural bases of these differences remain unknown (C. S. Chen et al., 2021, 2024; Ebitz et al., 2018). To examine how behavioral state interacts with outcome integration in mPFC and vHip, we implemented an explore-exploit hidden Markov model (C. S. Chen et al., 2021, 2024; Ebitz et al., 2018). This models each trial as falling into one of three different latent states: exploitation of the left lever, exploitation of the right lever, and exploration where the animals sample both levers (Fig. 4A). Using this approach to classify trials into either explore or exploit, we examined how outcome is integrated in explore versus exploit trials in GABAergic populations in mPFC (Fig. 4B,D) and vHip (Fig. 3F,H). Under exploration, defined as when animals explore on trial t0, we observed a full encoding of reward history in mPFC (Fig. 4C) and vHip (Fig. 4G). Reward suppressed mPFC and vHip by ITI end, integrating with the effects of the prior outcomes via graded suppression. When mice were rewarded on trial t0, suppression of mPFC and vHip by prior reward on t-1 was maintained through ITI end. When mice were unrewarded on

trial t0, suppression of mPFC and vHip by prior reward on t-1 was maintained through ITI end, relative to trials with two consecutive unrewarded outcomes. This describes a pattern of activity wherein activity at ITI end accurately reflects the recent integrated reward statistics of the environment.

We then examined encoding under a state of exploitation, defined as when animals exploit on trial t0, and observed that mPFC (Fig. 4E) and vHip (Fig. 4I) GABAergic encoding is degraded and anchors more to loss. Reward on a trial (t0) suppresses mPFC and vHip GABAergic activity by ITI end to a similar degree, regardless of prior outcome. Suppression of mPFC and vHip by prior reward on t-1 is only maintained through the next trial when mice are unrewarded on the current trial (t0). This effect is marginal in mPFC and robust in vHip. This suggests that, during exploitation, a single reward maximally and enduringly suppresses mPFC or vHip GABAergic activity and that, in the absence of subsequent reward, this suppression slowly dissipates. This describes a pattern of activity wherein activity at ITI end reflects the recent history of loss in the environment.

Animals tend to spend more time exploring than exploiting (Fig S2B) resulting in fewer trials which could impact signal quality. We confirmed that this alone does not explain the difference in neural encoding between exploration and exploitation. Taking sub-samples of exploration trials matched to the number of exploitation trials and repeated 100 times to compare the average significance across all comparisons, we confirmed full encoding of reward history in both mPFC (Fig S2C) and vHip (Fig S2D). Together this suggests that outcome integration in mPFC and vHip GABAergic neurons is dependent on behavioral state. Exploration elicits enhanced integration of information

to drive reward-motivated behavior. However, once exploiting, the neural representation of outcomes becomes anchored to loss, likely driven by the decreased utility of reward-oriented information during exploitation.

Behavioral state interacts with GABAergic mPFC and vHip populations to modulate choice but not engagement.

Examining neural representation of outcomes identified reward-mediated suppression as a mechanism for outcome encoding in mPFC and vHip GABAergic neurons. We also find that these neural populations utilize reward-mediated suppression to integrate outcomes across multiple trials when animals are exploring, but not exploiting. Given the interplay between neural encoding and behavioral state, we then asked how state-dependent neural processing may modulate behavior.

Glutamatergic populations in mPFC and vHip co-operatively integrate outcome information to modulate task engagement, but whether this is a behavioral function unique to excitatory cells or shared by GABAergic neurons is unknown. To determine if GABAergic populations in mPFC and vHip also contribute to modulating task engagement in a state-dependent manner, we examined if the interaction between behavioral state and neural activity at ITI end predicts task engagement, operationalized as latency to lever press on the subsequent trial. A linear mixed effects model revealed significant relationships between latency to lever press and GABAergic neuron activity in mPFC and vHip independent of behavioral state (Fig 4J, L). Increased activity during outcome integration in either mPFC or vHip increases latency to lever press, indicating

reduced behavioral engagement. This suggests that GABAergic neuron activity in both mPFC and vHip may modulate task engagement irrespective of behavioral state.

Examining neural activity time-locked to choice behavior, we identified differences in mPFC, but not vHip, GABAergic neuron activity preceding shifts in choice behavior following a rewarded outcome (Fig S1). Given state-dependent differences in outcome encoding, each region may also be differentially recruited in behavioral control over choice in a state-dependent manner, distinct from a role in modulating task engagement. To test this, we examined how neural activity predicts choice in periods when animals are exploring versus when animals are exploiting. Behaviorally, as expected, we find that animals are generally more likely to stay with their previous choice when exploiting than exploring (Fig. S2A). Yet, during exploration, increased activity in GABAergic mPFC, but not vHip, neurons predicts higher staying probabilities (Fig. 4K). In contrast, during exploitation, increased activity in GABAergic vHip, but not mPFC, neurons predicts reduced staying probabilities (Fig. 4M). This suggests that mPFC GABAergic neurons preferentially modulate staying under exploration whereas vHip GABAergic neurons preferentially modulate staying under exploitation. Together our results demonstrate that mPFC and vHip GABAergic neurons dynamically track outcome information in a state-dependent manner. Behaviorally, we observe that activity in both regions similarly reflects behavioral engagement in a state-independent manner but that each region is specialized to modulate choice behavior according to recent history of reward under distinct behavioral states pointing to complementary roles in control of reward-motivated behavior.

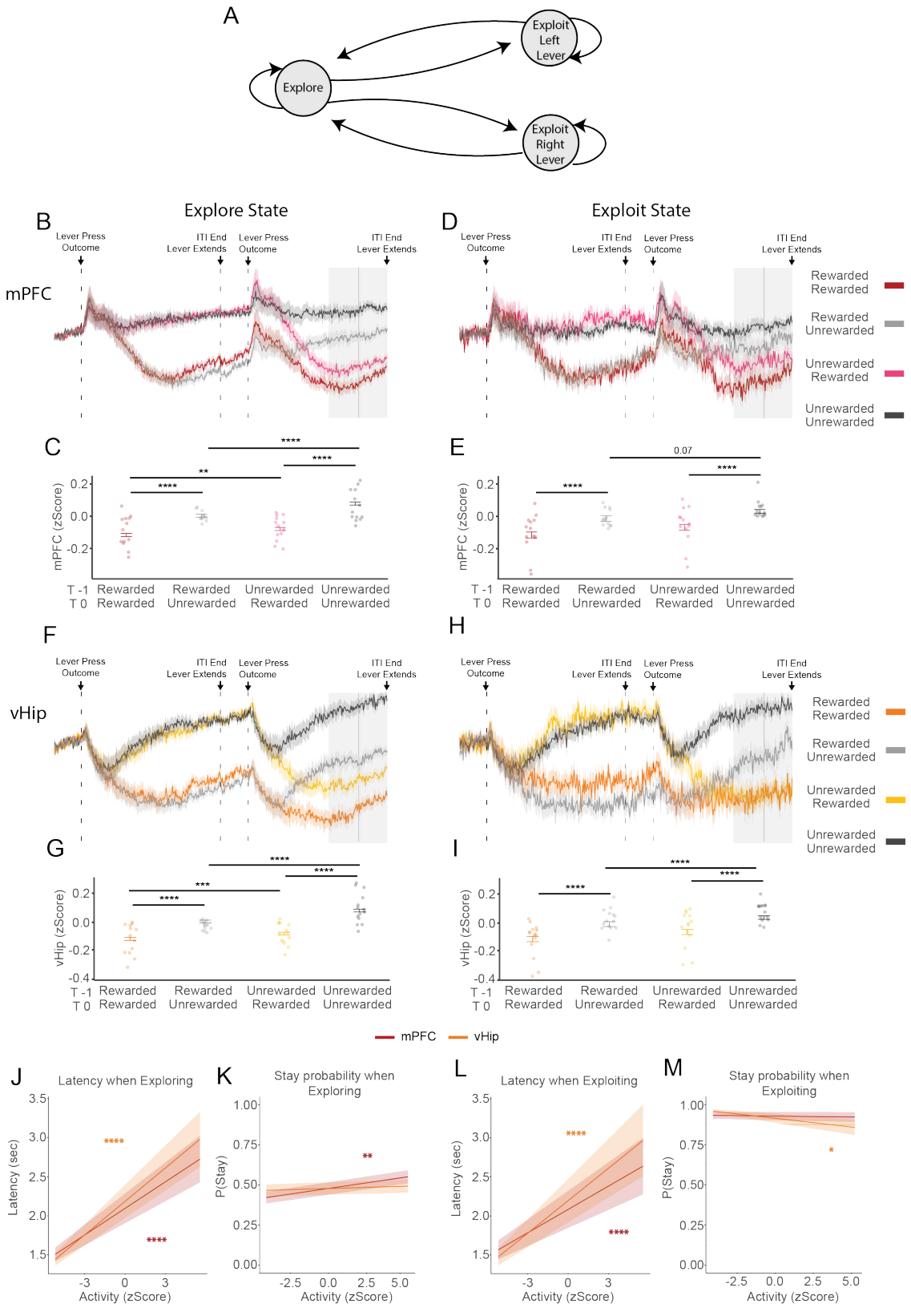


Figure 4. Exploration gates mPFC and vHip integration. *(A) Hidden Markov model identifies latent exploration or exploitation states that underlie a given sequence of*

choices. Choices are categorized as one of three states: exploitation of the left lever, exploitation of the right lever, or exploration where choice between levers is random. The two exploitation states are collapsed to a single state. (B) Estimated mean mPFC activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end) when animals are exploring. (C) On trial t0, mPFC activity is significantly suppressed by current reward (U→U vs U→R: Z=-14.087, p<0.0001; R→U vs R→R: Z=-10.300, p<0.0001). When trial t0 is unrewarded, mPFC activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 6.618, p<0.0001). When trial t0 is rewarded, mPFC activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→R vs R→R: Z= 3.173, p=0.0062). (D) Estimated mean mPFC activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end) when animals are exploIng. (E) On trial t0, mPFC activity is significantly suppressed by current reward (U→U vs U→R: Z=-4.920, p<0.0001; R→U vs R→R: Z=-4.260, p<0.0001). When trial t0 is unrewarded, mPFC activity trends towards remaining suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 2.366, p<0.0700). (F) Estimated mean vHip activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end) when animals are exploring. (G) On trial t0, vHip activity is significantly suppressed by current reward (U→U vs U→R: Z=-15.961, p<0.0001; R→U vs R→R: Z=-9.605, p<0.0001). When trial t0 is unrewarded, vHip activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 11.433, p<0.0001). When trial t0 is rewarded, vHip activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→R vs R→R: Z= 5.374, p<0.0001). (H) Estimated mean vHip activity across pairs of consecutive trials (t-1→t0) showing rewarded+rewarded (R→R), rewarded+unrewarded (R→U), unrewarded+rewarded (U→R), and unrewarded+unrewarded (U→U) trial pairs. Analysis focused on 8-10 sec after lever press (ITI end) when animals are exploiting. (I) On trial t0, vHip activity is significantly suppressed by current reward (U→U vs U→R: Z=-9.955, p<0.0001; R→U vs R→R: Z=-4.581, p<0.0001). When trial t0 is unrewarded, vHip activity remains significantly suppressed by reward experienced on the previous trial, t-1, (U→U vs R→U: Z= 5.334, p<0.0001). (J) Plot of estimated latency to respond on the subsequent trial given mPFC and vHip activity at ITI end shows that increased activity associates with longer latency under exploration (mPFC: trend= -0.028, Z=-6.454, p<0.0001; vHip: trend= -0.033, Z=-8.024, p<0.0001). (K) Plot of estimated staying probability on a given trial given mPFC and vHip activity before lever pressing shows that increased mPFC but not vHip activity increases staying probability when exploring (mPFC: trend= 0.056, Z=2.661, p=0.0078). (L) Plot of estimated latency to respond on the subsequent trial given mPFC and vHip activity at ITI end shows that increased activity associates with longer latency under exploitation (mPFC: trend= -0.024, Z=-3.312, p<0.0001; vHip: trend= -0.032, Z=-4.332, p<0.0001). (M) Plot of estimated staying probability on a given trial given mPFC and vHip activity before lever

*pressing shows that increased vHip but not mPFC activity decreases staying probability when exploiting (vHip: trend= -0.147, Z=-2.049, p=0.0405). Error bars represent SEM around the estimated mean. *p<0.05, **p<0.01 *** p<0.001 **** p<0.0001*

Discussion

We examined the role of inhibitory populations in reward processing in male mice in two regions implicated in motivated behavior. Using dual-site fiber photometry to probe trial-by-trial outcome encoding simultaneously in two circuits in the same animal during reward-guided choice, we find that inhibitory GABAergic neurons in mPFC and vHip similarly integrate reward via suppression of neural activity and that the fidelity of this encoding is sensitive to varying task demands. We then show that the richness of this encoding is also sensitive to behavioral state, with exploration between options eliciting a richer integration of outcome history, and exploitation associated with relatively impoverished outcome integration that is tuned to loss. Exploring how these patterns of neural activity associate with behavior, we identify associations with task engagement that are independent of behavioral state. Beyond this, we reveal that GABAergic populations in mPFC and vHip associate with dissociable roles in organizing choice behavior, with mPFC preferentially implicated in choice during exploration and vHip preferentially implicated choice during exploitation. Taken together, we identify a role for GABAergic populations in mPFC and vHip in integrating outcomes to support engagement and show that behavioral modulation by these regions is state-dependent. This suggests an important, yet often overlooked role for GABAergic neurons in reward-motivated behavior. While the present work is limited to male mice, in previous work in glutamatergic populations in mPFC and vHip we did not observe sex differences in neural integration of outcome nor in control of behavioral engagement (Iyer et al., 2024).

Indeed, our work suggests that sex differences in these circuits are likely specific to threat processing (Muir et al., 2024). Nevertheless, it will be important to determine if the present findings in GABAergic neurons also extend to female mice.

Our finding that mPFC GABAergic populations integrate information about outcomes of reward-motivated actions is consistent with the well-established role of mPFC in reward processing (Jeong et al., 2020; Sul et al., 2010). Our present and recent findings suggest that these functions are not unique to mPFC and are shared by other regions, such as vHip (Iyer et al., 2024). We recently demonstrated that glutamatergic projections from mPFC and vHip to the NAc encode outcome information through reward-mediated suppression, this suppression slowly dissipating after loss (Iyer et al., 2024). Here we show that this function is not unique to glutamatergic projection neurons but is shared by local GABAergic interneurons pointing to more global outcome-mediated modulation in mPFC and vHip. Neuromodulators are a likely candidate for this shared, region-level activity modulation. For example, this could be accomplished through dopaminergic modulation mediated by dopaminergic signaling at D2 receptors in mPFC and vHip to decrease excitability (Gerfen, 2023; Kebabian & Calne, 1979; Thibeault et al., 2019). Both mPFC and vHip are innervated by dopaminergic inputs from the ventral tegmental area that convey information about reward as well as associated cues and actions (Berridge, 2007; Han et al., 2017; Howe et al., 2013; Keiflin & Janak, 2015; Sayegh et al., 2024; Syed et al., 2016; Wassum et al., 2012). Another candidate is serotonergic modulation at 5-HT1A receptors in mPFC and vHip to decrease excitability (Hernandes et al., 2021; Lladó-Pelfort et al., 2012). Serotonin neurons in the dorsal raphe nucleus respond to reward and modulate their

baseline firing rate according to the average level of reward, an encoding feature that generally maps onto patterns of outcome encoding we observe in mPFC and vHip (Cohen et al., 2015). Noradrenergic and cholinergic signaling could also a play a role (C. S. Chen et al., 2024). Exploring the contributions and interactions of neuromodulators in supporting reward-mediated suppression represents a rich and exciting direction for future research.

Our findings also reveal intriguing distinctions between encoding in GABAergic populations and glutamatergic projections from these regions to the NAc. For example, GABAergic, but not glutamatergic neurons encode choice in a state-dependent manner. Furthermore, while outcome encoding in glutamatergic projection neurons is modulated by task demands, GABAergic neurons appear relatively insensitive to varying task requirements, with the noise in encoding increasing as task demands decrease without varying the informational content encoded. Instead, GABAergic neurons show sensitivity to modulation by behavioral state. These dissociations suggest that task demands and behavioral state independently impact choice and that distinct neuronal populations are tuned to distinct factors hinting at multiplexed control of encoding by mPFC and vHip GABAergic and glutamatergic projection neurons.

Overall, our findings point to unified, state-invariant control of behavioral engagement by mPFC and vHip with distributed control over choice behavior that is gated by behavioral state. While exploring (i.e. sampling between different options) mPFC GABAergic activity predicts choice behavior, consistent with previous findings in mPFC control of reward-motivated behavior. Curiously, we find increased mPFC GABAergic activity, associates with increased staying probability, a finding that is

seemingly incongruent with our observation that unrewarded outcomes increase mPFC activity. This could point to downstream mechanisms that transform mPFC signaling of choice behavior, or the existence of a sub-population of GABAergic neurons in mPFC that specifically encodes choice. Future work to resolve likely heterogeneity in the signal with single cell resolution will lead to interesting insights. In contrast, vHip signaling of choice aligns with the observation that unrewarded outcomes increase vHip GABAergic activity. When exploiting, choice behavior associates with vHip GABAergic activity such that, as GABAergic activity increases in vHip, animals are less likely to stay with their previous option. This suggests a specialized role for vHip in transitioning out of exploitation states when no longer advantageous. This is consistent with findings from recent work proposing a privileged role for vHip in representing as well as transitioning between states (Mishchanchuk et al., 2024).

Here we examined simultaneous GABAergic neuron encoding in two key brain regions for motivated behavior in male mice. By situating outcome encoding within the context of recent outcome history and behavioral state we identified a common neural mechanism of sustained temporal integration of reward outcomes and reveal how behavioral state differentially shapes internal representations within mPFC and vHip. We also identify critical region-specific interactions with choice: while mPFC GABAergic neuron activity uniquely predicts choice during exploration, vHip GABAergic neuron activity uniquely predicts choice during exploitation. Overall, our findings demonstrate that GABAergic neurons integrate outcome information across time and illustrate the complex interplay between behavioral state and distributed encoding across brain regions in reward-seeking behavior.

Methods

Animals

Mice were maintained on a 12-h light-dark cycle (lights on at 7:00AM) at 22-25ºC, group-housed with 3-4 same-sex cage-mates with *ad libitum* access to food and water. All experimental manipulations occurred during the light cycle, in accordance with guidelines of McGill University's Comparative Medicine and Animal Resources Center and approved by the McGill Animal Care Committee. 7-week-old male C57BL/6J mice were obtained from Jackson Laboratories and habituated to the colony room one week prior to start of manipulations. Mice were food restricted to 85% of their free-feeding body weight during experimentation.

Surgeries

Stereotaxic surgery was performed under ketamine (100 mg/kg)/xylazine (10 mg/kg) anesthesia. To achieve GABAergic neuron-specific GcaMP6f expression in mPFC and vHip, 0.5µl pAAV-mDlx-GcaMP6f-Fishell-2 virus (7×10^{12}VG/ml; Addgene) was infused into the infralimbic mPFC (A/P: 1.90, M/L: +/-0.3, D/V: -2.90) and ventral subiculum of the vHip (A/P: -3.40, M/L: +/-3.00, D/V: -4.90) at a rate of 0.1µl per min and allowed to diffuse for 10 min before withdrawing the needle. pAAV-mDlx-GcaMP6f-Fishell-2 was a gift from Gordon Fishell (Addgene plasmid # 83899 ; http://n2t.net/addgene:83899 ; RRID:Addgene_83899)(CITE). Chronically implantable optic fibers (Neurophotometrics) with 200µm core and 0.37 NA threaded through ceramic ferrules were implanted above the infralimbic mPFC (A/P: 1.90, M/L: +/-0.3, D/V: -2.80) and ventral subiculum of the vHip (A/P: -3.40, M/L: +/-3.00, D/V: -4.75)

Recordings began minimum 4 weeks after surgery to allow sufficient time for stable and robust retrograde virus expression.

Histology

After completion of all behavioral testing, mice were deeply anesthetized with ketamine/xylazine and transcardially perfused with phosphate buffered saline (PBS) and paraformaldehyde (4%). Brains were removed and post-fixed in paraformaldehyde for 24h and stored in PBS until sectioning on a vibratome (50 µm). Sections were mounted with Vectashield with DAPI (Vector Laboratories) and examined under a fluorescent microscope (Leica DM6000 B) to confirm viral expression and fiber placement. A confocal microscope (Zeiss LSM710) was used to obtain fluorescent images. Images were acquired with a 10x air objective using Zeiss Zen Blue imaging software. Images were collected in the McGill University Advanced BioImaging Facility (ABIF), *RRID:SCR_017697*. Mistargeted animals were excluded from analysis.

RNAscope

To confirm viral expression was restricted to GABAergic neurons, animals were injected with AAV1-mDLX-GcaMP6 targeting the mPFC and four weeks later transcardially perfused with formalin, and brains cryopreserved in a 10-30% sucrose gradient then flash-frozen in 2-methylbutane. Fixed-frozen brains were cryosectioned in 16um thick slices. Inhibitory neurons were identified by RNAScope in situ hybridization for Slc32a1 (vesicular GABA transporter; vGat (ACD Bio, Mm-Slc32a1-C2 probe) to identify GABAergic neurons followed by immunohistochemistry for GFP (Aves Labs,

chicken polyclonal anti-GFP antibody) to identify GcaMP-expressing neurons. Images were acquired on a Zeiss LSM700 confocal microscope using a 20x (NA 0.8) objective.

Apparatus

Behavioral experiments were performed in standard Med Associates operant boxes (15.24 x 13.34 x 12.7 cm) enclosed in sound attenuating chambers outfitted with a programmable audio generator, two retractable levers and cue lights either side of a food port for delivering a liquid chocolate milk reward (30µl, Nesquik) diluted with water 2:1. Boxes were controlled and data collected by a computer running MED-PC software (Med-Associates).

Lever Press Training

Training was completed in three stages, with all training sessions lasting 30 minutes. In the first stage, animals were presented with two levers, both of which delivered a chocolate milk reward with a 100% probability. To signal the start of the trial, both levers extended and the cue lights above the levers turned on, animals then had 60 seconds to make a response on either lever. A press on either lever resulted in lever retraction, immediate delivery of a 30 µL chocolate milk reward, and the start of a 3 second auditory cue (2kHz pure tone or white noise). Following either a lever press or 60 seconds with no press (i.e. an omission), a 10 second intertrial interval (ITI) was triggered. After one session with over 25 responses, animals progressed to the second stage. In this stage animals again were presented with two levers but reward was now delivered with a 50% probability on both levers. To signal the start of the trial, both levers extended and the cue lights above the levers turned on, animals then had 60

seconds to make a response. A lever press resulted in lever retraction and immediate delivery of the outcome, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise, counterbalanced across animals) or just a 3 second auditory cue (white noise or 2kHz pure tone). Following either a lever press or omission, a 10 second intertrial interval (ITI) was triggered. Following two consecutive sessions with over 40 responses, animals progressed to the third stage. This stage was the same as stage two except that animals now had only 10 seconds to make a response before an omission was registered. Following two consecutive sessions with over 100 responses animals achieved criterion to progress to the two-armed bandit task.

Two-armed bandit Task

The two-armed bandit task was performed over of 6 days with each session lasting one hour. In this task, animals were presented with two levers with one lever rewarded on 80% of trials, and the other lever rewarded on 20% of trials. To signal the start of the trial, both levers extended and the cue lights above the levers turned on, animals then had 10 seconds to make a response on either lever or an omission was registered. A lever press resulted in lever retraction and immediate delivery of the outcome, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise, counterbalanced across animals) or simply a different 3 second auditory cue (white noise or 2kHz pure tone) signaling non-reward. Following either a lever press or an omission, a 10 second intertrial interval (ITI) was triggered. To maintain a dynamic learning environment and high rates of rewarded and unrewarded outcomes, probability of reward was switched between levers after five consecutive

responses on the high probability lever. Four animals remained on the two-armed bandit task during the task degradation (data not shown).

One-Lever Forced Choice Task

The one lever forced choice task was performed over 3 days with each session lasting one hour. In this task, animals were presented with a single lever (counterbalanced across animals). Pressing this lever resulted in probabilistic reward on a predetermined schedule. The outcome schedule was matched to each animal's individual performance in the final three days of the two-armed bandit task, such that the first session in the one-lever task was yoked to the reward schedule experienced by the animal on day four in the two-armed bandit task, the second to day five, and the third to day six. To signal the start of the trial, the lever extended and the cue light above the lever turned on. Animals then had 10 seconds to make a response. A lever press resulted in lever retraction and immediate delivery of the outcome, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise) or simply a different 3 second auditory cue (white noise or 2kHz pure tone). Following either a lever press or an omission, a 10 second intertrial interval (ITI) was triggered.

No Lever Response Free Task

The no lever response free task was performed over the course of 3 days with each session lasting one hour. In this task, animals were able to retrieve non-contingently delivered rewards under a similar trial structure to both the two-armed bandit task and the one-lever forced choice task but with no levers available. To signal the start of the trial, cue lights above both levers turned on and remained illuminated for

a period of time matched to each animal's response time in the last three days of the two-armed bandit task. After cue lights turned off, outcomes were delivered, either a 30 µL chocolate milk reward and a 3 second auditory cue (2kHz pure tone or white noise) or simply a different 3 second auditory cue (white noise or 2kHz pure tone). As in the one-lever task, the outcome schedule was matched to each animal's performance in the final three days of the two-armed bandit task now also matching the latency to receive the outcome to the trial-by-trial latency to lever press on the two-armed bandit task with a 10 second intertrial interval (ITI).

Frame Independent Projected Fiber Photometry

To measure calcium-associated changes in fluorescence in real time, recordings were made from vHip and mPFC GABAergic neurons during the two-armed bandit task. Samples were collected at a frequency of 20 Hz using Neurophotometrics hardware through Bonsai and FlyCap software. Recordings were coupled to the start of behavioral analysis by interfacing Bonsai with MED-PC using a custom DAQ box (Neurophotometrics).

Photometry data extraction and normalization

Photometry data were extracted and analyzed using custom-written scripts in Python. To normalize the data, the control channel (415nm) was fitted to the raw (470nm). The fitted control was then subtracted from the raw trace. The resultant trace was divided by the fitted control giving the ΔF/F and converted to a Z-score. This calculation was performed over the entirety of the session to preserve dynamic fluctuation in population activity that persists beyond individual trials to allow

comparison across trials. For analyses of reward history, Z-scores were baseline subtracted from average activity in the two seconds prior to lever press on trial t-1 to account for shifted baselines in trial t0.

Data Analysis & Statistics

Linear Mixed Effects Regression

Linear Mixed Effects Regression Models are a powerful approach to probe variance attributable to variables of interest (e.g. trial outcome) while simultaneously controlling for random effects (e.g. session ID) (Fetcho et al., 2023; Kato et al., 2022; Yu et al., 2022). This is useful for modeling instances where there is nonindependence in the structure of data e.g. multiple trials recorded within multiple animals. Models were fit using the full interaction of the factors of interest (trial outcome, previous trial outcome, sex) and using animal ID and session ID as random effects using the *lme4* package in R (Bates et al., 2014). Where the dependent variable was latency, a Gamma link function was used to approximate the non-gaussian distribution. The fitted models were used to calculate estimated marginal means using the *emmeans* package in R (Lenth et al., 2021). The effect of variables of interest were then examined by comparing estimated marginal means. Given the large number of samples generated using this approach (all trials x all animals), comparisons of estimated marginal means were conducted using a Z-test and Sidak's method to adjust for multiple comparisons. Grubbs' test was used to exclude outliers.

Explore-Exploit Hidden Markov Model

We fit each animal's sequence of choices to a hidden Markov model to identify bouts of trials where animals explore their environment vs bouts of trials where animals exploit a given choice (C. S. Chen et al., 2021). This model is initialized with 3 hidden state: one where the animal exploits the left lever, one where the animal exploits the right lever, and one where the animal explores, being equally likely to select either lever. We defined the emissions matrix for the exploration state as a uniform distribution over choices, modeling behavior where animals sample between choices, and the emissions matrix for the exploitation states as a fixed choice, modeling behavior where animals perseverate on either the left lever or the right lever. We tied parameters across the two exploit states such that that the probability of transitioning from the explore state into either of the exploit states was the same. Similarly, we tied parameters such that the probability of transitioning from either of the exploit states into the explore state was the same. The exploit states also had the same probability of sustaining themselves. This defines an explore-exploit structure while allowing for 3 hidden states. This model also prevented transitions between the two exploit states, forcing entry to an exploration state first before moving into the other exploit state.

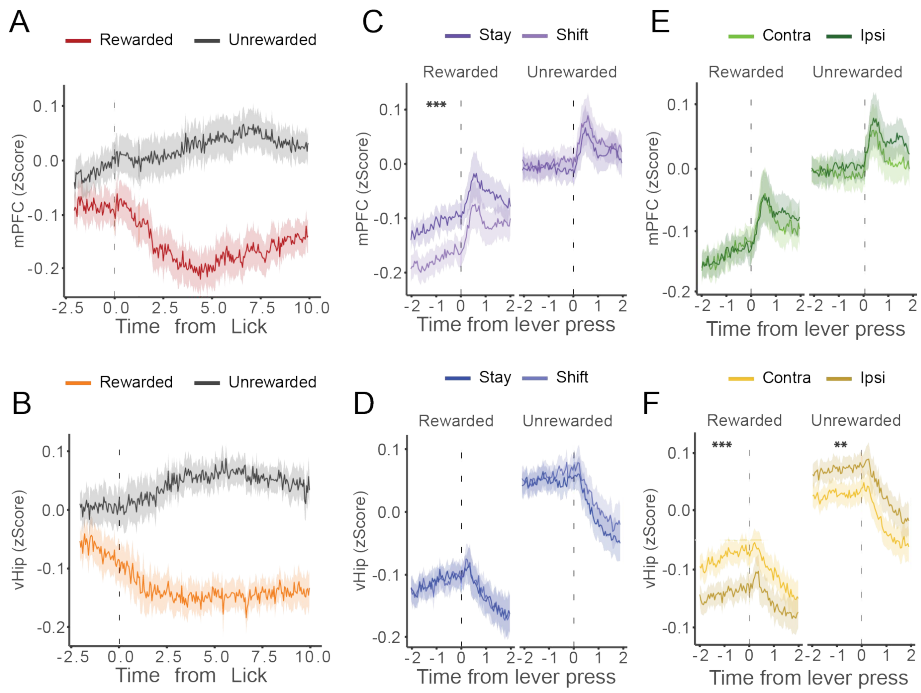
Having fixed emissions matrices and a constrained the transition matrix through parameter tying, we define an explore-exploit hidden Markov model with only two free parameters: the transition probability from exploration to exploitation and the transition probability from exploitation to exploration. We then fit this model to each animal's choice sequences using the Baum-Welch algorithm to find the maxima of the complete-data likelihood. We defined the initial distribution over states as p=1 for the explore state

and p=0 for the exploit states because animals should have no knowledge of the high-probability lever at the start of each session and thus needs to initially explore to discover the reward distribution of its environment. The algorithm was reinitialized 10 times and the model that maximized the observed log likelihood for each session was selected. The Viterbi algorithm was then used to discover a posteriori the most likely sequence of latent states.

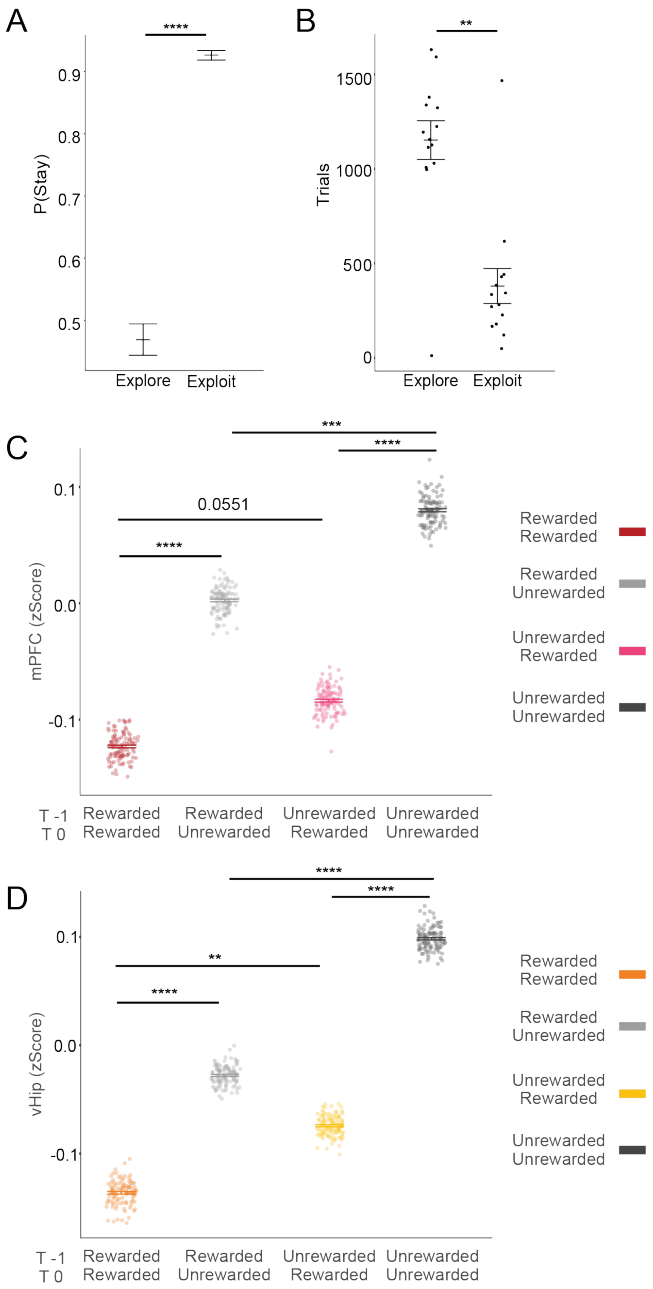
Subsampling analysis of explore trials

In order to account for the imbalance in sample size between exploration and exploitation trials, we randomly subsampled the exploration trials so that the total number of trials matched the total number of exploitation trials. Following this subsampling, we performed a linear mixed effects regression, as described above, to estimate the marginal mean activity and Z-value associated with each set of trial combinations. We then repeated this process 100 times, plotting the estimated means derived from each regression. p-values were then estimated from the average of the Z-values calculated from each sub-sample.

Supplementary Information



Supplemental Figure 1. mPFC and vHip GABAergic activity time-locked to different task-relevant features. *Estimated mean (A) mPFC and (B) vHip activity time-locked to the first rewarded and unrewarded lick during the ITI. Estimated mean (C) mPFC and (D) vHip activity time-locked to lever press on trials in which the previous choice is repeated (stay) and when a different choice is made (shift). Estimated mean (E) mPFC and (F) vHip activity time-locked to lever press on trials in which animals chose the lever contralateral or ipsilateral to their implant. **p<0.001, ***p<0.0001*



Supplemental Figure 2 Outcome encoding in mPFC and vHip under exploration is not explained by differences in trial number. *(A) Animals are more likely to stay when exploiting than exploring (Z=42.6288, p<0.0001) (B) Animals spend more time exploring than exploiting in the two-lever probabilistic reward learning task (t(13)=-4.0636, p=0.0013). Subsampling explore trials 100 times to match the number of exploit confirms reveal full outcome encoding is observed under comparable trial numbers in (C) mPFC (U→R vs R→R: Z=1.9182, p=0.0551; U→U vs R→U: Z=3.8397, p=0.0001; U→U vs U→R: Z=-6.0547, p<0.0001; R→U vs R→R: Z=-8.1729, p<0.0001) and (D) vHip (U→R vs R→R: Z=3.1813, p=0.0015; U→U vs R→U: Z=6.5774, p<0.0001; U→U vs U→R: Z=-5.5250, p<0.0001; R→U vs R→R: Z=-8.0869, p<0.0001).*

References

Abboud, S., & Sadeh, D. (1984). The use of cross-correlation function for the alignment of ECG waveforms and rejection of extrasystoles. *Computers and Biomedical Research*, *17*(3), 258–266. https://doi.org/10.1016/S0010-4809(84)80017-8

Abela, A. R., Duan, Y., & Chudasama, Y. (2015). Hippocampal interplay with the nucleus accumbens is critical for decisions about time. *The European Journal of Neuroscience*, *42*(5), 2224–2233. https://doi.org/10.1111/ejn.13009

Aberman, J. E., & Salamone, J. D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience*, *92*(2), 545–552. https://doi.org/10.1016/S0306-4522(99)00004-4

Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron*, *59*(4), 648–661. https://doi.org/10.1016/j.neuron.2008.07.004

Amir, N., Prouvost, C., & Kuckertz, J. M. (2012). Lack of a Benign Interpretation Bias in Social Anxiety Disorder. *Cognitive Behaviour Therapy*, *41*(2), 119–129. https://doi.org/10.1080/16506073.2012.662655

Bagot, R. C., Parise, E. M., Peña, C. J., Zhang, H.-X., Maze, I., Chaudhury, D., Persaud, B., Cachope, R., Bolaños-Guzmán, C. A., Cheer, J. F., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2015). Ventral hippocampal afferents to the nucleus accumbens regulate susceptibility to depression. *Nature Communications*, *6*, 7062. https://doi.org/10.1038/ncomms8062

Baldo, B. A., Sadeghian, K., Basso, A. M., & Kelley, A. E. (2002). Effects of selective dopamine D1 or D2 receptor blockade within nucleus accumbens subregions on ingestive behavior and associated motor activity. *Behavioural Brain Research*, *137*(1), 165–177. https://doi.org/10.1016/S0166-4328(02)00293-0

Barbano, M. F., Qi, J., Chen, E., Mohammad, U., Espinoza, O., Candido, M., Wang, H., Liu, B., Hahn, S., Vautier, F., & Morales, M. (2024). VTA glutamatergic projections to the nucleus accumbens suppress psychostimulant-seeking behavior. *Neuropsychopharmacology*, 1–11. https://doi.org/10.1038/s41386-024-01905-3

Bari, B. A., Grossman, C. D., Lubin, E. E., Rajagopalan, A. E., Cressy, J. I., & Cohen, J. Y. (2019). Stable representations of decision variables for flexible behavior. *Neuron*, *103*(5), 922-933.e7. https://doi.org/10.1016/j.neuron.2019.06.001

Barker, J. M., Bryant, K. G., & Chandler, L. J. (2019). Inactivation of ventral hippocampus projections promotes sensitivity to changes in contingency. *Learning & Memory*, *26*(1), 1–8.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv Preprint arXiv:1406.5823*.

Bathellier, B., Tee, S. P., Hrovat, C., & Rumpel, S. (2013). A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. *Proceedings of the National Academy of Sciences*, *110*(49), 19950–19955. https://doi.org/10.1073/pnas.1312125110

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., & Dayan, P. (2013). Dopamine Modulates Reward-Related Vigor. *Neuropsychopharmacology*, *38*(8), 1495–1503. https://doi.org/10.1038/npp.2013.48

Belujon, P., & Grace, A. A. (2008). Critical role of the prefrontal cortex in the regulation of hippocampus-accumbens information flow. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(39), 9797–9805. https://doi.org/10.1523/JNEUROSCI.2200-08.2008

Bercovici, D. A., Princz-Lebel, O., Tse, M. T., Moorman, D. E., & Floresco, S. B. (2018). Optogenetic Dissection of Temporal Dynamics of Amygdala-Striatal Interplay during Risk/Reward Decision Making. *eNeuro*, *5*(6), ENEURO.0422-18.2018. https://doi.org/10.1523/ENEURO.0422-18.2018

Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology*, *191*(3), 391–431. https://doi.org/10.1007/s00213-006-0578-x

Beyeler, A., Namburi, P., Glober, G. F., Simonnet, C., Calhoon, G. G., Conyers, G. F., Luck, R., Wildes, C. P., & Tye, K. M. (2016). Divergent routing of positive and negative information from the amygdala during memory retrieval. *Neuron*, *90*(2), 348–361. https://doi.org/10.1016/j.neuron.2016.03.004

Bishop, S. J., & Gagne, C. (2018). Anxiety, Depression, and Decision Making: A Computational Perspective. *Annual Review of Neuroscience*, *41*, 371–388. https://doi.org/10.1146/annurev-neuro-080317-062007

Bloem, B., Poorthuis, R. B., & Mansvelder, H. D. (2014). Cholinergic modulation of the medial prefrontal cortex: The role of nicotinic receptors in attention and regulation of neuronal activity. *Frontiers in Neural Circuits*, *8*, 17. https://doi.org/10.3389/fncir.2014.00017

Borrego, M. B., Grigsby, K. B., Townsley, K. G., Chan, A., Firsick, E. F., Tran, A., & Ozburn, A. R. (2022). Central nucleus of the amygdala projections onto the nucleus accumbens core regulate binge-like alcohol drinking in a CRF-dependent manner. *Neuropharmacology*, *203*, 108874. https://doi.org/10.1016/j.neuropharm.2021.108874

Bouarab, C., Thompson, B., & Polter, A. M. (2019). VTA GABA Neurons at the Interface of Stress and Reward. *Frontiers in Neural Circuits*, *13*. https://doi.org/10.3389/fncir.2019.00078

Bouras, N. N., Mack, N. R., & Gao, W.-J. (2023). Prefrontal modulation of anxiety through a lens of noradrenergic signaling. *Frontiers in Systems Neuroscience*, *17*, 1173326. https://doi.org/10.3389/fnsys.2023.1173326

Bouret, S., & Richmond, B. J. (2015). Sensitivity of Locus Ceruleus Neurons to Reward Value for Goal-Directed Actions. *The Journal of Neuroscience*, *35*(9), 4005–4014. https://doi.org/10.1523/JNEUROSCI.4553-14.2015

Britt, J. P., Benaliouad, F., McDevitt, R. A., Stuber, G. D., Wise, R. A., & Bonci, A. (2012). Synaptic and Behavioral Profile of Multiple Glutamatergic Inputs to the Nucleus Accumbens. *Neuron*, *76*(4), 790–803. https://doi.org/10.1016/j.neuron.2012.09.040

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Brown, V. J., & Bowman, E. M. (1995). Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. *The European Journal of Neuroscience*, *7*(12), 2479–2485. https://doi.org/10.1111/j.1460-9568.1995.tb01046.x

Calipari, E. S., Bagot, R. C., Purushothaman, I., Davidson, T. J., Yorgason, J. T., Peña, C. J., Walker, D. M., Pirpinias, S. T., Guise, K. G., Ramakrishnan, C., Deisseroth, K., & Nestler, E. J. (2016). In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proceedings of the National Academy of Sciences*, *113*(10), 2726–2731. https://doi.org/10.1073/pnas.1521238113

Carter, A. G., Soler-Llavina, G. J., & Sabatini, B. L. (2007). Timing and Location of Synaptic Inputs Determine Modes of Subthreshold Integration in Striatal Medium Spiny Neurons. *The Journal of Neuroscience*, *27*(33), 8967–8977. https://doi.org/10.1523/JNEUROSCI.2798-07.2007

Carvalho Poyraz, F., Holzner, E., Bailey, M. R., Meszaros, J., Kenney, L., Kheirbek, M. A., Balsam, P. D., & Kellendonk, C. (2016). Decreasing Striatopallidal Pathway Function Enhances Motivation by Energizing the Initiation of Goal-Directed Action. *The Journal of Neuroscience*, *36*(22), 5988–6001. https://doi.org/10.1523/JNEUROSCI.0444-16.2016

Chen, C. S., Knep, E., Han, A., Ebitz, R. B., & Grissom, N. M. (2021). Sex differences in learning from exploration. *eLife*, *10*, e69748. https://doi.org/10.7554/eLife.69748

Chen, C. S., Mueller, D., Knep, E., Ebitz, R. B., & Grissom, N. M. (2024). Dopamine and Norepinephrine Differentially Mediate the Exploration-Exploitation Tradeoff. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *44*(44), e1194232024. https://doi.org/10.1523/JNEUROSCI.1194-23.2024

Chen, G., Lai, S., Bao, G., Ke, J., Meng, X., Lu, S., Wu, X., Xu, H., Wu, F., Xu, Y., Xu, F., Bi, G.-Q., Peng, G., Zhou, K., & Zhu, Y. (2023). Distinct reward processing by subregions of the nucleus accumbens. *Cell Reports*, *42*(2). https://doi.org/10.1016/j.celrep.2023.112069

Christoffel, D. J., Walsh, J. J., Hoerbelt, P., Heifets, B. D., Llorach, P., Lopez, R. C., Ramakrishnan, C., Deisseroth, K., & Malenka, R. C. (2021). Selective filtering of excitatory inputs to nucleus accumbens by dopamine and serotonin. *Proceedings of the National Academy of Sciences*, *118*(24), e2106648118. https://doi.org/10.1073/pnas.2106648118

Cid-Pellitero, E. del, & Garzón, M. (2011). Medial prefrontal cortex receives input from dorsal raphe nucleus neurons targeted by Hypocretin1/OrexinA-containing axons. *Neuroscience*, *172*, 30–43. https://doi.org/10.1016/j.neuroscience.2010.10.058

Cohen, J. Y., Amoroso, M. W., & Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. *eLife*, *4*, e06346. https://doi.org/10.7554/eLife.06346

Collins, D. P., Anastasiades, P. G., Marlin, J. J., & Carter, A. G. (2018). Reciprocal circuits linking the prefrontal cortex with dorsal and ventral thalamic nuclei. *Neuron*, *98*(2), 366-379.e4. https://doi.org/10.1016/j.neuron.2018.03.024

Cover, T. M., & Thomas, J. A. (1991). Entropy, relative entropy and mutual information. *Elements of Information Theory*, *2*(1), 12–13.

Cox, J., Minerva, A. R., Fleming, W. T., Zimmerman, C. A., Hayes, C., Zorowitz, S., Bandi, A., Ornelas, S., McMannon, B., Parker, N. F., & Witten, I. B. (2023). A neural substrate of sex-dependent modulation of motivation. *Nature Neuroscience*, *26*(2), Article 2. https://doi.org/10.1038/s41593-022-01229-9

Dalton, G. L., Phillips, A. G., & Floresco, S. B. (2014). Preferential Involvement by Nucleus Accumbens Shell in Mediating Probabilistic Learning and Reversal Shifts. *Journal of Neuroscience*, *34*(13), 4618–4626. https://doi.org/10.1523/JNEUROSCI.5058-13.2014

Dana, H., Sun, Y., Mohar, B., Hulse, B. K., Kerlin, A. M., Hasseman, J. P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., Macklin, J. J., Chen, Y., Konnerth, A., Jayaraman, V., Looger, L. L., Schreiter, E. R., Svoboda, K., & Kim, D. S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods*, *16*(7), 649–657. https://doi.org/10.1038/s41592-019-0435-6

Day, J. J., Jones, J. L., & Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs. *The European Journal of Neuroscience*, *33*(2), 308–321. https://doi.org/10.1111/j.1460-9568.2010.07531.x

De Groote, A., & de Kerchove d'Exaerde, A. (2021). Thalamo-Nucleus Accumbens Projections in Motivated Behaviors and Addiction. *Frontiers in Systems Neuroscience*, *15*, 711350. https://doi.org/10.3389/fnsys.2021.711350

de Jong, J. W., Afjei, S. A., Dorocic, I. P., Peck, J. R., Liu, C., Kim, C. K., Tian, L., Deisseroth, K., & Lammel, S. (2019). A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron*, *101*(1), 133-151.e7. https://doi.org/10.1016/j.neuron.2018.11.005

Dezfouli, A., Balleine, B. W., & Nock, R. (2019). Optimal response vigor and choice under non-stationary outcome values. *Psychonomic Bulletin & Review*, *26*(1), 182–204. https://doi.org/10.3758/s13423-018-1500-3

Dieterich, A., Floeder, J., Stech, K., Lee, J., Srivastava, P., Barker, D. J., & Samuels, B. A. (2021). Activation of Basolateral Amygdala to Nucleus Accumbens Projection Neurons Attenuates Chronic Corticosterone-Induced Behavioral Deficits in Male Mice. *Frontiers in Behavioral Neuroscience*, *15*. https://doi.org/10.3389/fnbeh.2021.643272

Do-Monte, F. H., Minier-Toribio, A., Quiñones-Laracuente, K., Medina-Colón, E. M., & Quirk, G. J. (2017). Thalamic regulation of sucrose-seeking during unexpected reward omission. *Neuron*, *94*(2), 388-400.e4. https://doi.org/10.1016/j.neuron.2017.03.036

Dong, X., Li, S., & Kirouac, G. J. (2017). Collateralization of projections from the paraventricular nucleus of the thalamus to the nucleus accumbens, bed nucleus of the stria terminalis, and central nucleus of the amygdala. *Brain Structure & Function*, *222*(9), 3927–3943. https://doi.org/10.1007/s00429-017-1445-8

Ebitz, R. B., Albarran, E., & Moore, T. (2018). Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex. *Neuron*, *97*(2), 450-461.e9. https://doi.org/10.1016/j.neuron.2017.12.007

Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). Reinforcement learning and Bayesian inference provide complementary models for the unique advantage of adolescents in stochastic reversal. *Developmental Cognitive Neuroscience*, *55*, 101106. https://doi.org/10.1016/j.dcn.2022.101106

Engelke, D. S., Zhang, X. O., O'Malley, J. J., Fernandez-Leon, J. A., Li, S., Kirouac, G. J., Beierlein, M., & Do-Monte, F. H. (2021). A hypothalamic-thalamostriatal circuit that controls approach-avoidance conflict in rats. *Nature Communications*, *12*, 2517. https://doi.org/10.1038/s41467-021-22730-y

Fetcho, R. N., Hall, B. S., Estrin, D. J., Walsh, A. P., Schuette, P. J., Kaminsky, J., Singh, A., Roshgodal, J., Bavley, C. C., Nadkarni, V., Antigua, S., Huynh, T. N., Grosenick, L., Carthy, C., Komer, L., Adhikari, A., Lee, F. S., Rajadhyaksha, A. M., & Liston, C. (2023). Regulation of social interaction in mice by a frontostriatal circuit modulated by established hierarchical relationships. *Nature Communications*, *14*, 2487. https://doi.org/10.1038/s41467-023-37460-6

Fisone, G., Bonito-Oliva, A., & Feyder, M. (2011). Deciphering the Actions of Antiparkinsonian and Antipsychotic Drugs on cAMP/DARPP-32 Signaling. *Frontiers in Neuroanatomy*, *5*. https://doi.org/10.3389/fnana.2011.00038

Floresco, S. B. (2007). Dopaminergic regulation of limbic-striatal interplay. *Journal of Psychiatry & Neuroscience : JPN*, *32*(6), 400–411.

Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of Psychology*, *66*(1), 25–52. https://doi.org/10.1146/annurev-psych-010213-115159

Floresco, S. B., Montes, D. R., Tse, M. M. T., & van Holstein, M. (2018). Differential Contributions of Nucleus Accumbens Subregions to Cue-Guided Risk/Reward Decision Making and Implementation of Conditional Rules. *The Journal of Neuroscience*, *38*(8), 1901–1914. https://doi.org/10.1523/JNEUROSCI.3191-17.2018

Francis, T. C., & Lobo, M. K. (2017). Emerging role for nucleus accumbens medium spiny neuron subtypes in depression. *Biological Psychiatry*, *81*(8), 645–653. https://doi.org/10.1016/j.biopsych.2016.09.007

French, S. J., & Totterdell, S. (2002). Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *Journal of Comparative Neurology*, *446*(2), 151–165.

Gallo, E. F., Meszaros, J., Sherman, J. D., Chohan, M. O., Teboul, E., Choi, C. S., Moore, H., Javitch, J. A., & Kellendonk, C. (2018). Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nature Communications*, *9*, 1086. https://doi.org/10.1038/s41467-018-03272-2

Garcia-Garcia, A., Tancredi, A. N.-, & Leonardo, E. D. (2014). 5-HT1A receptors in mood and anxiety: Recent insights into autoreceptor versus heteroreceptor function. *Psychopharmacology*, *231*(4), 623–636. https://doi.org/10.1007/s00213-013-3389-x

Gerfen, C. R. (2023). Segregation of D1 and D2 dopamine receptors in the striatal direct and indirect pathways: An historical perspective. *Frontiers in Synaptic Neuroscience*, *14*, 1002960. https://doi.org/10.3389/fnsyn.2022.1002960

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D1 and D2 Dopamine Receptor-regulated Gene Expression of Striatonigral and Striatopallidal Neurons. *Science*, *250*(4986), 1429–1432. https://doi.org/10.1126/science.2147780

Goto, Y., & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, *8*(6), Article 6. https://doi.org/10.1038/nn1471

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences*, *30*(5), 220–227. https://doi.org/10.1016/j.tins.2007.03.003

Guillaumin, M. C. C., Viskaitis, P., Bracey, E., Burdakov, D., & Peleg-Raibstein, D. (2023). Disentangling the role of NAc D1 and D2 cells in hedonic eating. *Molecular Psychiatry*, *28*(8), 3531–3547. https://doi.org/10.1038/s41380-023-02131-x

Guitart-Masip, M., Beierholm, U. R., Dolan, R., Duzel, E., & Dayan, P. (2011). Vigor in the Face of Fluctuating Rates of Reward: An Experimental Examination. *Journal of Cognitive Neuroscience*, *23*(12), 3933–3938. https://doi.org/10.1162/jocn_a_00090

Gustafson, N. J., & Daw, N. D. (2011). Grid Cells, Place Cells, and Geodesic Generalization for Spatial Reinforcement Learning. *PLOS Computational Biology*, *7*(10), e1002235. https://doi.org/10.1371/journal.pcbi.1002235

Hamel, L., Cavdaroglu, B., Yeates, D., Nguyen, D., Riaz, S., Patterson, D., Khan, N., Kirolos, N., Roper, K., Ha, Q. A., & Ito, R. (2022). Cortico-Striatal Control over Adaptive Goal-Directed Responding Elicited by Cues Signaling Sucrose Reward or Punishment. *Journal of Neuroscience*, *42*(18), 3811–3822. https://doi.org/10.1523/JNEUROSCI.2175-21.2022

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2016). Mesolimbic Dopamine Signals the Value of Work. *Nature Neuroscience*, *19*(1), 117–126. https://doi.org/10.1038/nn.4173

Han, X., Jing, M., Zhao, T., Wu, N., Song, R., & Li, J. (2017). Role of dopamine projections from ventral tegmental area to nucleus accumbens and medial prefrontal cortex in reinforcement behaviors assessed using optogenetic manipulation. *Metabolic Brain Disease*, *32*(5), 1491–1502. https://doi.org/10.1007/s11011-017-0023-3

Hanson, E., Brandel-Ankrapp, K. L., & Arenkiel, B. R. (2021). Dynamic Cholinergic Tone in the Basal Forebrain Reflects Reward-Seeking and Reinforcement During Olfactory Behavior. *Frontiers in Cellular Neuroscience*, *15*, 635837. https://doi.org/10.3389/fncel.2021.635837

He, Y., Huang, Y. H., Schlüter, O. M., & Dong, Y. (2023). Cue- versus reward-encoding basolateral amygdala projections to nucleus accumbens. *eLife*, *12*, e89766. https://doi.org/10.7554/eLife.89766

Hernandes, P. M., Batistela, M. F., Vilela-Costa, H. H., Sant'Ana, A. B., Kumpel, V. D., Tirapelle, M. C., Bom, A. de O. P., de Andrade, T. G. C. S., & Zangrossi, H. (2021). Role

of 5-HT1A receptors in the ventral hippocampus in the regulation of anxiety- and panic-related defensive behaviors in rats. *Behavioural Brain Research*, *408*, 113296. https://doi.org/10.1016/j.bbr.2021.113296

Herrnstein, R. J. (1997). *The matching law: Papers in psychology and economics*. Harvard University Press.

Hofmeister, J., & Sterpenich, V. (2015). A Role for the Locus Ceruleus in Reward Processing: Encoding Behavioral Energy Required for Goal-Directed Actions. *Journal of Neuroscience*, *35*(29), 10387–10389. https://doi.org/10.1523/JNEUROSCI.1734-15.2015

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, *500*(7464), Article 7464. https://doi.org/10.1038/nature12475

Huang, S., Borgland, S. L., & Zamponi, G. W. (2019). Dopaminergic modulation of pain signals in the medial prefrontal cortex: Challenges and perspectives. *Neuroscience Letters*, *702*, 71–76. https://doi.org/10.1016/j.neulet.2018.11.043

Ikemoto, S., Glazier, B. S., Murphy, J. M., & McBride, W. J. (1997). Role of Dopamine D1 and D2 Receptors in the Nucleus Accumbens in Mediating Reward. *The Journal of Neuroscience*, *17*(21), 8580–8587. https://doi.org/10.1523/JNEUROSCI.17-21-08580.1997

Iyer, E. S., Vitaro, P., Wu, S., Muir, J., Tse, Y. C., Cvetkovska, V., & Bagot, R. C. (2024). *Reward integration in prefrontal-cortical and ventral-hippocampal nucleus accumbens*

inputs cooperatively modulates engagement (p. 2024.06.27.601063). bioRxiv. https://doi.org/10.1101/2024.06.27.601063

Jeong, H., Kim, D., Song, M., Paik, S.-B., & Jung, M. W. (2020). Distinct roles of parvalbumin- and somatostatin-expressing neurons in flexible representation of task variables in the prefrontal cortex. *Progress in Neurobiology*, *187*, 101773. https://doi.org/10.1016/j.pneurobio.2020.101773

Jeong, H., Taylor, A., Floeder, J. R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D. A., & Namboodiri, V. M. K. (2022). Mesolimbic dopamine release conveys causal associations. *Science*, *378*(6626), eabq6740. https://doi.org/10.1126/science.abq6740

Jing, M., Zhang, P., Wang, G., Feng, J., Mesik, L., Zeng, J., Jiang, H., Wang, S., Looby, J. C., Guagliardo, N. A., Langma, L. W., Lu, J., Zuo, Y., Talmage, D. A., Role, L. W., Barrett, P. Q., Zhang, L. I., Luo, M., Song, Y., … Li, Y. (2018). A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nature Biotechnology*, *36*(8), 726–737. https://doi.org/10.1038/nbt.4184

Kagiampaki, Z., Rohner, V., Kiss, C., Curreli, S., Dieter, A., Wilhelm, M., Harada, M., Duss, S. N., Dernic, J., Bhat, M. A., Zhou, X., Ravotto, L., Ziebarth, T., Wasielewski, L. M., Sönmez, L., Benke, D., Weber, B., Bohacek, J., Reiner, A., … Patriarchi, T. (2023). Sensitive multicolor indicators for monitoring norepinephrine in vivo. *Nature Methods*, *20*(9), 1426–1436. https://doi.org/10.1038/s41592-023-01959-z

Kato, T., Mitsukura, Y., Yoshida, K., Mimura, M., Takata, N., & Tanaka, K. F. (2022). Oscillatory Population-Level Activity of Dorsal Raphe Serotonergic Neurons Is

Inscribed in Sleep Structure. *The Journal of Neuroscience*, *42*(38), 7244–7255. https://doi.org/10.1523/JNEUROSCI.2288-21.2022

Kebabian, J. W., & Calne, D. B. (1979). Multiple receptors for dopamine. *Nature*, *277*(5692), 93–96. https://doi.org/10.1038/277093a0

Keiflin, R., & Janak, P. H. (2015). Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron*, *88*(2), 247–263. https://doi.org/10.1016/j.neuron.2015.08.037

Keistler, C., Barker, J. M., & Taylor, J. R. (2015). Infralimbic prefrontal cortex interacts with nucleus accumbens shell to unmask expression of outcome-selective Pavlovian-to-instrumental transfer. *Learning & Memory*, *22*(10), 509–513. https://doi.org/10.1101/lm.038810.115

Klausberger, T., & Somogyi, P. (2008). Neuronal Diversity and Temporal Dynamics: The Unity of Hippocampal Circuit Operations. *Science (New York, N.Y.)*, *321*(5885), 53–57. https://doi.org/10.1126/science.1149381

Koo, J. W., Lobo, M. K., Chaudhury, D., Labonté, B., Friedman, A., Heller, E., Peña, C. J., Han, M.-H., & Nestler, E. J. (2014). Loss of BDNF Signaling in D1R-Expressing NAc Neurons Enhances Morphine Reward by Reducing GABA Inhibition. *Neuropsychopharmacology*, *39*(11), 2646–2653. https://doi.org/10.1038/npp.2014.118

Köther, A. K., Alpers, G. W., Büdenbender, B., Lenhart, M., Michel, M. S., & Kriegmair, M. C. (2021). Predicting decisional conflict: Anxiety and depression in shared decision

making. *Patient Education and Counseling*, *104*(5), 1229–1236. https://doi.org/10.1016/j.pec.2020.10.037

Kravitz, A. V., & Kreitzer, A. C. (2012). Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. *Physiology (Bethesda, Md.)*, *27*(3), 10.1152/physiol.00004.2012. https://doi.org/10.1152/physiol.00004.2012

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818. https://doi.org/10.1038/nn.3100

Kuchibhotla, K. V., Hindmarsh Sten, T., Papadoyannis, E. S., Elnozahy, S., Fogelson, K. A., Kumar, R., Boubenec, Y., Holland, P. C., Ostojic, S., & Froemke, R. C. (2019). Dissociating task acquisition from expression during learning reveals latent knowledge. *Nature Communications*, *10*(1), 2151. https://doi.org/10.1038/s41467-019-10089-0

Kupchik, Y. M., Brown, R. M., Heinsbroek, J. A., Lobo, M. K., Schwartz, D. J., & Kalivas, P. W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nature Neuroscience*, *18*(9), 1230–1232. https://doi.org/10.1038/nn.4068

Kutlu, M. G., Zachry, J. E., Melugin, P. R., Cajigas, S. A., Chevee, M. F., Kelly, S. J., Kutlu, B., Tian, L., Siciliano, C. A., & Calipari, E. S. (2021). Dopamine release in the nucleus accumbens core signals perceived saliency. *Current Biology*, *31*(21), 4748-4761.e8. https://doi.org/10.1016/j.cub.2021.08.052

Lafferty, C. K., Yang, A. K., Mendoza, J. A., & Britt, J. P. (2020). Nucleus Accumbens Cell Type- and Input-Specific Suppression of Unproductive Reward Seeking. *Cell Reports*, *30*(11), 3729-3742.e3. https://doi.org/10.1016/j.celrep.2020.02.095

Langdon, A. J., Hathaway, B. A., Zorowitz, S., Harris, C. B. W., & Winstanley, C. A. (2019). Relative insensitivity to time-out punishments induced by win-paired cues in a rat gambling task. *Psychopharmacology*, *236*(8), 2543–2556. https://doi.org/10.1007/s00213-019-05308-x

Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., & Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological Medicine*, *50*(10), 1613–1622. https://doi.org/10.1017/S0033291719001570

LeGates, T. A., Kvarta, M. D., Tooley, J. R., Francis, T. C., Lobo, M. K., Creed, M. C., & Thompson, S. M. (2018). Reward behavior is regulated by the strength of hippocampus-nucleus accumbens synapses. *Nature*, *564*(7735), 258–262. https://doi.org/10.1038/s41586-018-0740-8

Lenth, R. V., Buerkner, P., Herve, M., Love, J., Riebl, H., & Singmann, H. (2021). *Emmeans: Estimated marginal means, aka least-squares means*. *CRAN*.

Lind, E. B., Sweis, B. M., Asp, A. J., Esguerra, M., Silvis, K. A., David Redish, A., & Thomas, M. J. (2023). A quadruple dissociation of reward-related behaviour in mice across excitatory inputs to the nucleus accumbens shell. *Communications Biology*, *6*(1), Article 1. https://doi.org/10.1038/s42003-023-04429-6

Lindenbach, D., Vacca, G., Ahn, S., Seamans, J. K., & Phillips, A. G. (2022). Optogenetic modulation of glutamatergic afferents from the ventral subiculum to the nucleus

accumbens: Effects on dopamine function, response vigor and locomotor activity. *Behavioural Brain Research*, *434*, 114028. https://doi.org/10.1016/j.bbr.2022.114028

Lipski, W. J., & Grace, A. A. (2013). Activation and inhibition of neurons in the hippocampal ventral subiculum by norepinephrine and locus coeruleus stimulation. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *38*(2), 285–292. https://doi.org/10.1038/npp.2012.157

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703–713. https://doi.org/10.1016/j.neuron.2005.05.002

Liu, X., & Carter, A. G. (2018). Ventral Hippocampal Inputs Preferentially Drive Corticocortical Neurons in the Infralimbic Prefrontal Cortex. *The Journal of Neuroscience*, *38*(33), 7351–7363. https://doi.org/10.1523/JNEUROSCI.0378-18.2018

Lladó-Pelfort, L., Santana, N., Ghisi, V., Artigas, F., & Celada, P. (2012). 5-HT1A receptor agonists enhance pyramidal cell firing in prefrontal cortex through a preferential action on GABA interneurons. *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(7), 1487–1497. https://doi.org/10.1093/cercor/bhr220

Lloyd, A., Roiser, J. P., Skeen, S., Freeman, Z., Badalova, A., Agunbiade, A., Busakhwe, C., DeFlorio, C., Marcu, A., Pirie, H., Saleh, R., Snyder, T., Fearon, P., & Viding, E. (2024). Reviewing explore/exploit decision-making as a transdiagnostic target for psychosis,

depression, and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *24*(5), 793–815. https://doi.org/10.3758/s13415-024-01186-9

Lobo, M. K., Covington, H. E., Chaudhury, D., Friedman, A. K., Sun, H., Damez-Werno, D., Dietz, D., Zaman, S., Koo, J. W., Kennedy, P. J., Mouzon, E., Mogri, M., Neve, R. L., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2010). Cell Type Specific Loss of BDNF Signaling Mimics Optogenetic Control of Cocaine Reward. *Science (New York, N.Y.)*, *330*(6002), 385–390. https://doi.org/10.1126/science.1188472

Loriaux, A. L., Roitman, J. D., & Roitman, M. F. (2011). Nucleus accumbens shell, but not core, tracks motivational value of salt. *Journal of Neurophysiology*, *106*(3), 1537–1544. https://doi.org/10.1152/jn.00153.2011

McGarry, L. M., & Carter, A. G. (2017). Prefrontal cortex drives distinct projection neurons in the basolateral amygdala. *Cell Reports*, *21*(6), 1426–1433. https://doi.org/10.1016/j.celrep.2017.10.046

Mingote, S., Amsellem, A., Kempf, A., Rayport, S., & Chuhma, N. (2019). Dopamine-Glutamate Neuron Projections to the Nucleus Accumbens Medial Shell and Behavioral Switching. *Neurochemistry International*, *129*, 104482. https://doi.org/10.1016/j.neuint.2019.104482

Mishchanchuk, K., Gregoriou, G., Qü, A., Kastler, A., Huys, Q. J. M., Wilbrecht, L., & MacAskill, A. F. (2024). Hidden state inference requires abstract contextual representations in the ventral hippocampus. *Science*, *386*(6724), 926–932. https://doi.org/10.1126/science.adq5874

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), Article 7759. https://doi.org/10.1038/s41586-019-1235-y

Moore, D. G., Valentini, G., Walker, S. I., & Levin, M. (2018). Inform: Efficient information-theoretic analysis of collective behaviors. *Frontiers in Robotics and AI*, *5*, 60.

Muir, J., Iyer, E. S., Tse, Y.-C., Sorensen, J., Wu, S., Eid, R. S., Cvetkovska, V., Wassef, K., Gostlin, S., Vitaro, P., Spencer, N. J., & Bagot, R. C. (2024). Sex-biased neural encoding of threat discrimination in nucleus accumbens afferents drives suppression of reward behavior. *Nature Neuroscience*, *27*(10), 1966–1976. https://doi.org/10.1038/s41593-024-01748-7

Muir, J., Tse, Y. C., Iyer, E. S., Biris, J., Cvetkovska, V., Lopez, J., & Bagot, R. C. (2020). Ventral Hippocampal Afferents to Nucleus Accumbens Encode Both Latent Vulnerability and Stress-Induced Susceptibility. *Biological Psychiatry*, *88*(11), 843–854. https://doi.org/10.1016/j.biopsych.2020.05.021

Nestler, E. J., & Carlezon, W. A. (2006). The Mesolimbic Dopamine Reward Circuit in Depression. *Biological Psychiatry*, *59*(12), 1151–1159. https://doi.org/10.1016/j.biopsych.2005.09.018

Nevin, J. A., & Mandell, C. (2017). Comparing positive and negative reinforcement: A fantasy experiment. *Journal of the Experimental Analysis of Behavior*, *107*(1), 34–38. https://doi.org/10.1002/jeab.237

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Nowend, K. L., Arizzi, M., Carlson, B. B., & Salamone, J. D. (2001). D1 or D2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. *Pharmacology Biochemistry and Behavior*, *69*(3), 373–382. https://doi.org/10.1016/S0091-3057(01)00524-X

Noworyta-Sokolowska, K., Kozub, A., Jablonska, J., Rodriguez Parkitna, J., Drozd, R., & Rygula, R. (2019). Sensitivity to negative and positive feedback as a stable and enduring behavioural trait in rats. *Psychopharmacology*, *236*(8), 2389–2403. https://doi.org/10.1007/s00213-019-05333-w

O'Donnell, P., & Grace, A. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: Hippocampal gating of prefrontal cortical input. *The Journal of Neuroscience*, *15*(5), 3622–3639. https://doi.org/10.1523/JNEUROSCI.15-05-03622.1995

O'Hearn, E., & Molliver, M. E. (1984). Organization of raphe-cortical projections in rat: A quantitative retrograde study. *Brain Research Bulletin*, *13*(6), 709–726. https://doi.org/10.1016/0361-9230(84)90232-6

Onge, J. R. S., Abhari, H., & Floresco, S. B. (2011). Dissociable Contributions by Prefrontal D1 and D2 Receptors to Risk-Based Decision Making. *Journal of Neuroscience*, *31*(23), 8625–8633. https://doi.org/10.1523/JNEUROSCI.1020-11.2011

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., & Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), Article 7643. https://doi.org/10.1038/nature21376

Otis, J. M., Zhu, M., Namboodiri, V. M. K., Cook, C. A., Kosyk, O., Matan, A. M., Ying, R., Hashikawa, Y., Hashikawa, K., Trujillo-Pisanty, I., Guo, J., Ung, R. L., Rodriguez-Romaguera, J., Anton, E. S., & Stuber, G. D. (2019). Paraventricular Thalamus Projection Neurons Integrate Cortical and Hypothalamic Signals for Cue-Reward Processing. *Neuron*, *103*(3), 423-431.e4. https://doi.org/10.1016/j.neuron.2019.05.018

Pan, W.-X., Brown, J., & Dudman, J. T. (2013). Neural signals of extinction in the inhibitory microcircuit of the ventral midbrain. *Nature Neuroscience*, *16*(1), 71–78. https://doi.org/10.1038/nn.3283

Pardo-Garcia, T. R., Garcia-Keller, C., Penaloza, T., Richie, C. T., Pickel, J., Hope, B. T., Harvey, B. K., Kalivas, P. W., & Heinsbroek, J. A. (2019). Ventral Pallidum Is the Primary Target for Accumbens D1 Projections Driving Cocaine Seeking. *The Journal of Neuroscience*, *39*(11), 2041–2051. https://doi.org/10.1523/JNEUROSCI.2822-18.2018

Parker, N. F., Baidya, A., Cox, J., Haetzel, L. M., Zhukovskaya, A., Murugan, M., Engelhard, B., Goldman, M. S., & Witten, I. B. (2022). Choice-selective sequences dominate in cortical relative to thalamic inputs to NAc to support reinforcement learning. *Cell Reports*, *39*(7). https://doi.org/10.1016/j.celrep.2022.110756

Parsons, M. P., Li, S., & Kirouac, G. J. (2007). Functional and anatomical connection between the paraventricular nucleus of the thalamus and dopamine fibers of the nucleus accumbens. *The Journal of Comparative Neurology*, *500*(6), 1050–1063. https://doi.org/10.1002/cne.21224

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: Affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476–483. https://doi.org/10.1016/j.tics.2012.07.009

Piantadosi, P. T., Yeates, D. C. M., & Floresco, S. B. (2020). Prefrontal cortical and nucleus accumbens contributions to discriminative conditioned suppression of reward-seeking. *Learning & Memory*, *27*(10), 429–440. https://doi.org/10.1101/lm.051912.120

Qi, J., Zhang, S., Wang, H.-L., Barker, D. J., Miranda-Barrientos, J., & Morales, M. (2016). VTA glutamatergic inputs to nucleus accumbens drive aversion by acting on GABAergic interneurons. *Nature Neuroscience*, *19*(5), 725–733. https://doi.org/10.1038/nn.4281

Rapanelli, M., Wang, W., Hurley, E., Feltri, M. L., Pittenger, C., Frick, L. R., & Yan, Z. (2023). Cholinergic neurons in the basal forebrain are involved in behavioral abnormalities associated with Cul3 deficiency: Role of prefrontal cortex projections in cognitive deficits. *Translational Psychiatry*, *13*(1), 1–11. https://doi.org/10.1038/s41398-023-02306-8

Reed, S. J., Lafferty, C. K., Mendoza, J. A., Yang, A. K., Davidson, T. J., Grosenick, L., Deisseroth, K., & Britt, J. P. (2018). Coordinated Reductions in Excitatory Input to the

Nucleus Accumbens Underlie Food Consumption. *Neuron*, *99*(6), 1260-1273.e4. https://doi.org/10.1016/j.neuron.2018.07.051

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory: Vol. Vol. 2*.

Richard, J. M., & Berridge, Kent. C. (2013). Prefrontal cortex modulates desire and dread generated by nucleus accumbens glutamate disruption. *Biological Psychiatry*, *73*(4), 360–370. https://doi.org/10.1016/j.biopsych.2012.08.009

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *The Journal of Neuroscience*, *29*(42), 13365–13376. https://doi.org/10.1523/JNEUROSCI.2572-09.2009

Sackett, D. A., Saddoris, M. P., & Carelli, R. M. (2017). Nucleus Accumbens Shell Dopamine Preferentially Tracks Information Related to Outcome Value of Reward. *eNeuro*, *4*(3). https://doi.org/10.1523/ENEURO.0058-17.2017

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015a). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. *The Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015b). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal

Complementary Signals for Error Prediction and Incentive Motivation. *Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Salamone, J. D., Correa, M., Yang, J.-H., Rotolo, R., & Presby, R. (2018). Dopamine, Effort-Based Choice, and Behavioral Economics: Basic and Translational Research. *Frontiers in Behavioral Neuroscience*, *12*, 52. https://doi.org/10.3389/fnbeh.2018.00052

Santana, N., & Artigas, F. (2017). Laminar and Cellular Distribution of Monoamine Receptors in Rat Medial Prefrontal Cortex. *Frontiers in Neuroanatomy*, *11*, 87. https://doi.org/10.3389/fnana.2017.00087

Sayegh, F. J. P., Mouledous, L., Macri, C., Pi Macedo, J., Lejards, C., Rampon, C., Verret, L., & Dahan, L. (2024). Ventral tegmental area dopamine projections to the hippocampus trigger long-term potentiation and contextual learning. *Nature Communications*, *15*(1), 4100. https://doi.org/10.1038/s41467-024-47481-4

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Scott, A., Palmer, D., Newell, B., Lin, I., Cayton, C. A., Paulson, A., Remde, P., & Richard, J. M. (2023). Ventral Pallidal GABAergic Neuron Calcium Activity Encodes Cue-Driven

Reward Seeking and Persists in the Absence of Reward Delivery. *Journal of Neuroscience*, *43*(28), 5191–5203. https://doi.org/10.1523/JNEUROSCI.0013-23.2023

Segal, M. (1990). Serotonin attenuates a slow inhibitory postsynaptic potential in rat hippocampal neurons. *Neuroscience*, *36*(3), 631–641. https://doi.org/10.1016/0306-4522(90)90006-P

Smith, R. J., Lobo, M. K., Spencer, S., & Kalivas, P. W. (2013). Cocaine-induced adaptations in D1 and D2 accumbens projection neurons (a dichotomy not necessarily synonymous with direct and indirect pathways). *Current Opinion in Neurobiology*, *23*(4), 546–552. https://doi.org/10.1016/j.conb.2013.01.026

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., Sousa, N., & Rodrigues, A. J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. *Nature Communications*, *7*(1), 11829. https://doi.org/10.1038/ncomms11829

Soares-Cunha, C., Coimbra, B., Sousa, N., & Rodrigues, A. J. (2016). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neuroscience & Biobehavioral Reviews*, *68*, 370–386. https://doi.org/10.1016/j.neubiorev.2016.05.021

Soares-Cunha, C., de Vasconcelos, N. A. P., Coimbra, B., Domingues, A. V., Silva, J. M., Loureiro-Campos, E., Gaspar, R., Sotiropoulos, I., Sousa, N., & Rodrigues, A. J. (2020). Nucleus accumbens medium spiny neurons subtypes signal both reward and aversion. *Molecular Psychiatry*, *25*(12), 3241–3255. https://doi.org/10.1038/s41380-019-0484-3

Soares-Cunha, C., & Heinsbroek, J. A. (2023). Ventral pallidal regulation of motivated behaviors and reinforcement. *Frontiers in Neural Circuits*, *17*, 1086053. https://doi.org/10.3389/fncir.2023.1086053

Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., & Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell*, *184*(10), 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

St-Amand, D., Sheldon, S., & Otto, A. R. (2018). Modulating Episodic Memory Alters Risk Preference during Decision-making. *Journal of Cognitive Neuroscience*, *30*(10), 1433–1441. https://doi.org/10.1162/jocn_a_01253

Stern, E. A., Kincaid, A. E., & Wilson, C. J. (1997). Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons in vivo. *Journal of Neurophysiology*, *77*(4), 1697–1715. https://doi.org/10.1152/jn.1997.77.4.1697

Stopper, C. M., & Floresco, S. B. (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(1), 97–112. https://doi.org/10.3758/s13415-010-0015-9

Stuber, G. D., Sparta, D. R., Stamatakis, A. M., van Leeuwen, W. A., Hardjoprajitno, J. E., Cho, S., Tye, K. M., Kempadoo, K. A., Zhang, F., Deisseroth, K., & Bonci, A. (2011). Amygdala to nucleus accumbens excitatory transmission facilitates reward seeking. *Nature*, *475*(7356), 377–380. https://doi.org/10.1038/nature10194

Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making. *Neuron*, *66*(3), 449–460. https://doi.org/10.1016/j.neuron.2010.03.033

Syed, E. C. J., Grima, L. L., Magill, P. J., Bogacz, R., Brown, P., & Walton, M. E. (2016). Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nature Neuroscience*, *19*(1), Article 1. https://doi.org/10.1038/nn.4187

Szczypka, M. S., Kwok, K., Brot, M. D., Marck, B. T., Matsumoto, A. M., Donahue, B. A., & Palmiter, R. D. (2001). Dopamine Production in the Caudate Putamen Restores Feeding in Dopamine-Deficient Mice. *Neuron*, *30*(3), 819–828. https://doi.org/10.1016/S0896-6273(01)00319-1

Tecuapetla, F., Patel, J. C., Xenias, H., English, D., Tadros, I., Shah, F., Berlin, J., Deisseroth, K., Rice, M. E., Tepper, J. M., & Koos, T. (2010). Glutamatergic Signaling by Mesolimbic Dopamine Neurons in the Nucleus Accumbens. *The Journal of Neuroscience*, *30*(20), 7105–7110. https://doi.org/10.1523/JNEUROSCI.0265-10.2010

Thibeault, K. C., Gunes Kutlu, M., Sanders, C., & Calipari, E. S. (2019). Cell-type and projection-specific dopaminergic encoding of aversive stimuli in addiction. *Brain Research*, *1713*, 1–15. https://doi.org/10.1016/j.brainres.2018.12.024

Thierry, A. M., Gioanni, Y., Dégénétais, E., & Glowinski, J. (2000). Hippocampo-prefrontal cortex pathway: Anatomical and electrophysiological characteristics. *Hippocampus*, *10*(4), 411–419. https://doi.org/10.1002/1098-1063(2000)10:4<411::AID-HIPO7>3.0.CO;2-A

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*, 212–222. https://doi.org/10.2307/1415413

Titulaer, J., Björkholm, C., Feltmann, K., Malmlöf, T., Mishra, D., Bengtsson Gonzales, C., Schilström, B., & Konradsson-Geuken, Å. (2021). The Importance of Ventral Hippocampal Dopamine and Norepinephrine in Recognition Memory. *Frontiers in Behavioral Neuroscience*, *15*, 667244. https://doi.org/10.3389/fnbeh.2021.667244

Tremblay, R., Lee, S., & Rudy, B. (2016). GABAergic interneurons in the neocortex: From cellular properties to circuits. *Neuron*, *91*(2), 260–292. https://doi.org/10.1016/j.neuron.2016.06.033

Tzschentke, T. M., & Schmidt, W. J. (2000). Functional relationship among medial prefrontal cortex, nucleus accumbens, and ventral tegmental area in locomotion and reward. *Critical Reviews in Neurobiology*, *14*(2), 131–142.

van Holstein, M., MacLeod, P. E., & Floresco, S. B. (2020). Basolateral amygdala—Nucleus accumbens circuitry regulates optimal cue-guided risk/reward decision making. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *98*, 109830. https://doi.org/10.1016/j.pnpbp.2019.109830

Verharen, J. P. H., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2019). Differential contributions of striatal dopamine D1 and D2 receptors to component processes of

value-based decision making. *Neuropsychopharmacology*, *44*(13), 2195–2204. https://doi.org/10.1038/s41386-019-0454-0

Vierock, J., Rodriguez-Rozada, S., Dieter, A., Pieper, F., Sims, R., Tenedini, F., Bergs, A. C. F., Bendifallah, I., Zhou, F., Zeitzschel, N., Ahlbeck, J., Augustin, S., Sauter, K., Papagiakoumou, E., Gottschalk, A., Soba, P., Emiliani, V., Engel, A. K., Hegemann, P., & Wiegert, J. S. (2021). BiPOLES is an optogenetic tool developed for bidirectional dual-color control of neurons. *Nature Communications*, *12*(1), 4527. https://doi.org/10.1038/s41467-021-24759-5

Voorn, P., Vanderschuren, L. J. M. J., Groenewegen, H. J., Robbins, T. W., & Pennartz, C. M. A. (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends in Neurosciences*, *27*(8), 468–474. https://doi.org/10.1016/j.tins.2004.06.006

Wan, J., Peng, W., Li, X., Qian, T., Song, K., Zeng, J., Deng, F., Hao, S., Feng, J., Zhang, P., Zhang, Y., Zou, J., Pan, S., Shin, M., Venton, B. J., Zhu, J. J., Jing, M., Xu, M., & Li, Y. (2021). A genetically encoded sensor for measuring serotonin dynamics. *Nature Neuroscience*, *24*(5), 746–752. https://doi.org/10.1038/s41593-021-00823-7

Wang, A. Y., Miura, K., & Uchida, N. (2013). The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nature Neuroscience*, *16*(5), 639–647. https://doi.org/10.1038/nn.3377

Warlow, S. M., Singhal, S. M., Hollon, N. G., Faget, L., Dowlat, D. S., Zell, V., Hunker, A. C., Zweifel, L. S., & Hnasko, T. S. (2024). Mesoaccumbal glutamate neurons drive reward via glutamate release but aversion via dopamine co-release. *Neuron*, *112*(3), 488-499.e5. https://doi.org/10.1016/j.neuron.2023.11.002

Wassum, K. M., Ostlund, S. B., & Maidment, N. T. (2012). Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task. *Biological Psychiatry*, *71*(10), 846–854. https://doi.org/10.1016/j.biopsych.2011.12.019

Wenzel, J. M., Zlebnik, N. E., Patton, M. H., Smethells, J. R., Ayvazian, V. M., Dantrassy, H. M., Zhang, L.-Y., Mathur, B. N., & Cheer, J. F. (2023). Selective chemogenetic inactivation of corticoaccumbal projections disrupts trait choice impulsivity. *Neuropsychopharmacology*, *48*(12), Article 12. https://doi.org/10.1038/s41386-023-01604-5

West, E. A., & Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after Reinforcer Devaluation. *The Journal of Neuroscience*, *36*(4), 1128–1139. https://doi.org/10.1523/JNEUROSCI.2976-15.2016

White, N. M. (1989). Reward or reinforcement: What's the difference? *Neuroscience & Biobehavioral Reviews*, *13*(2), 181–186. https://doi.org/10.1016/S0149-7634(89)80028-4

Willmore, L., Cameron, C., Yang, J., Witten, I. B., & Falkner, A. L. (2022). Behavioural and dopaminergic signatures of resilience. *Nature*, *611*(7934), 124–132. https://doi.org/10.1038/s41586-022-05328-2

Wilson, C. (2008). Up and down states. *Scholarpedia Journal*, *3*(6), 1410.

Wilson, C. J., & Groves, P. M. (1981). Spontaneous firing patterns of identified spiny neurons in the rat neostriatum. *Brain Research*, *220*(1), 67–80. https://doi.org/10.1016/0006-8993(81)90211-0

Wilson, C., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *The Journal of Neuroscience*, *16*(7), 2397–2410. https://doi.org/10.1523/JNEUROSCI.16-07-02397.1996

Wu, J.-L., Li, Z.-M., Chen, H., Chen, W.-J., Hu, N.-Y., Jin, S.-Y., Li, X.-W., Chen, Y.-H., Yang, J.-M., & Gao, T.-M. (2024). Distinct septo-hippocampal cholinergic projections separately mediate stress-induced emotional and cognitive deficits. *Science Advances*, *10*(45), eado1508. https://doi.org/10.1126/sciadv.ado1508

Yamaguchi, T., Wang, H.-L., Li, X., Ng, T. H., & Morales, M. (2011). Mesocorticolimbic glutamatergic pathway. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(23), 8476–8490. https://doi.org/10.1523/JNEUROSCI.1598-11.2011

Yang, H., Jong, J. W. de, Tak, Y., Peck, J., Bateup, H. S., & Lammel, S. (2018). Nucleus Accumbens Subnuclei Regulate Motivated Behavior via Direct Inhibition and Disinhibition of VTA Dopamine Subpopulations. *Neuron*, *97*(2), 434-449.e4. https://doi.org/10.1016/j.neuron.2017.12.022

Yoo, J. H., Zell, V., Gutierrez-Reed, N., Wu, J., Ressler, R., Shenasa, M. A., Johnson, A. B., Fife, K. H., Faget, L., & Hnasko, T. S. (2016). Ventral tegmental area glutamate neurons co-release GABA and promote positive reinforcement. *Nature Communications*, *7*(1), 13697. https://doi.org/10.1038/ncomms13697

Yoshida, K., Drew, M. R., Kono, A., Mimura, M., Takata, N., & Tanaka, K. F. (2021). Chronic social defeat stress impairs goal-directed behavior through dysregulation of ventral

hippocampal activity in male mice. *Neuropsychopharmacology*, *46*(9), 1606–1616. https://doi.org/10.1038/s41386-021-00990-y

Yoshida, K., Drew, M. R., Mimura, M., & Tanaka, K. F. (2019). Serotonin-mediated inhibition of ventral hippocampus is required for sustained goal-directed behavior. *Nature Neuroscience*, *22*(5), Article 5. https://doi.org/10.1038/s41593-019-0376-5

Yoshida, K., Tsutsui-Kimura, I., Kono, A., Yamanaka, A., Kobayashi, K., Watanabe, M., Mimura, M., & Tanaka, K. F. (2020). Opposing Ventral Striatal Medium Spiny Neuron Activities Shaped by Striatal Parvalbumin-Expressing Interneurons during Goal-Directed Behaviors. *Cell Reports*, *31*(13), 107829. https://doi.org/10.1016/j.celrep.2020.107829

Yu, Z., Guindani, M., Grieco, S. F., Chen, L., Holmes, T. C., & Xu, X. (2022). Beyond t test and ANOVA: Applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron*, *110*(1), 21–35. https://doi.org/10.1016/j.neuron.2021.10.030

Zhu, Y., Wienecke, C. F. R., Nachtrab, G., & Chen, X. (2016). A thalamic input to the nucleus accumbens mediates opiate dependence. *Nature*, *530*(7589), 219–222. https://doi.org/10.1038/nature16954

Zhuo, Y., Luo, B., Yi, X., Dong, H., Miao, X., Wan, J., Williams, J. T., Campbell, M. G., Cai, R., Qian, T., Li, F., Weber, S. J., Wang, L., Li, B., Wei, Y., Li, G., Wang, H., Zheng, Y., Zhao, Y., … Li, Y. (2024). Improved green and red GRAB sensors for monitoring dopaminergic activity in vivo. *Nature Methods*, *21*(4), 680–691. https://doi.org/10.1038/s41592-023-02100-w

Chapter 4

Discussion

Integrating information about actions and their outcomes is essential for goal-directed behavior. For behavior to remain flexible and adaptive in dynamic environments, outcome information must be continuously integrated across multiple timescales. This thesis sought to elucidate the role of glutamatergic inputs to the NAc in supporting reward processing and motivated behavior, with the expectation that different inputs and regions would play distinct roles in supporting these processes. Using dual-site fiber photometry in two important glutamatergic NAc-projecting regions, the mPFC and vHip, each with distinct attributed behavioral and computational functions, we surprisingly demonstrate that both projections similarly encode information about outcomes, with reward driving suppression of neural activity and unrewarded outcomes gradually restoring activity. This pattern of encoding continues into subsequent trials, tracking prior reward history as a graded function of unrewarded outcomes. By degrading task requirements, performing conditional entropy analyses, and through carefully examination of how outcome encoding interacts with reward history, we find a degree a specialization across inputs such that, while mPFC projections to NAc consistently encode outcome, encoding in vHip projections to NAc is anchored to unrewarded outcomes. Examining the behavioral relevance of activity in these pathways revealed that higher activity in either pathway associates with decreased engagement and optogenetic stimulation of either pathway suppressed engagement, with an additive effect of simultaneous dual-pathway stimulation. Together this

demonstrates that mPFC and vHip projections to NAc integrate outcomes to modulate engagement. While each pathway executes this function under somewhat differing demands, once engaged, there is redundancy in their capacity to modulate behavior pointing to complementary roles in control of reward seeking.

Thinking that inhibitory interneurons in the mPFC and vHip could be responsible for the patterns of outcome encoding observed in NAc-projecting neurons, we used dual-site fiber photometry to record the activity of GABAergic neurons in these regions. Expecting to see a pattern of activity inverse to that described in the first study, we instead found a similar pattern of reward integration in the GABAergic neurons of these regions. Once again, reward drives suppression of neural activity while unrewarded outcomes gradually restore activity. Using a hidden Markov model to decompose behavior into bouts of exploration and exploitation, we demonstrate that exploration gates outcome integration in both mPFC and vHip. As in the first study, examining the behavioral relevance of activity in the GABAergic neurons of these regions revealed that higher activity in either region associates with decreased engagement, regardless of state. However, we show a state-dependent role in modulating choice behavior in each region. This demonstrates that GABAergic neurons also play an important role in outcome integration and state-sensitive modulation of reward-motivated behavior. The similarities in encoding motifs across both studies suggest perhaps a common input, implicating neuromodulators such as serotonin and acetylcholine in mediating outcome integration in mPFC and vHip.

Potential mechanisms of reward-mediated suppression of neural activity

A principal finding of this thesis is the existence of a reward-encoding motif present both in glutamatergic projection from mPFC and vHip to NAc as well as in GABAergic neurons in these regions. Here reward drives a suppression of neural activity while unrewarded outcomes gradually restore activity, defining a pattern of outcome integration where neural activity is inversely correlated with the amount of reward experienced. Initially, following the first study wherein we found reward-mediated suppression in glutamatergic projections to NAc, we hypothesized that these processes could be mediated by local inhibitory neurons in mPFC and vHip. Accordingly, we hypothesized that neural activity of these inhibitory neurons would be the inverse of the activity in the glutamatergic projections. Instead, when we recorded neural activity from GABAergic populations in mPFC and vHip, we found a very similar pattern of neural encoding to that in the glutamatergic projections. Together, this suggests that both NAc projecting cells as well as GABAergic cells in mPFC and vHip are subject to a common source of modulation. Further, the similarity in the modulation across regions, as well as the lack of evidence that one region leads the other in encoding suggests that likely both mPFC and vHip are modulated by a common source, that is most likely a neuromodulatory input. Both mPFC and vHip receive a range of neuromodulatory inputs, including dopamine, serotonin, norepinephrine and acetylcholine.

Dopamine

Dopaminergic modulation originating from the ventral tegmental area (VTA) is one potential source of this modulation. VTA dopamine is known to convey information about reward in addition to information about associated cues and actions (Berridge, 2007; Jeong et al., 2022; Nestler & Carlezon, 2006; Schultz et al., 1997). Additionally, both mPFC and vHip receive dopaminergic input from this source (Huang et al., 2019; Lisman & Grace, 2005; Titulaer et al., 2021; Tzschentke & Schmidt, 2000). The delivery of a reward is usually associated with an increase in dopaminergic signaling whereas the omission of a reward is associated with a decrease (Schultz et al., 1997). This should result in a pattern of modulation where the concentration of dopamine in mPFC and vHip correlates with the amount of reward received, effectively the inverse of the pattern of activity we observe in mPFC and vHip. Dopaminergic signaling at D2 receptors could mediate that changing sign of the signaling, with increasing dopamine concentration decreasing the excitability of neurons (Gerfen, 2023; Kebabian & Calne, 1979; Thibeault et al., 2019). One major caveat is that D1 and D2 receptors are not completely segregated in mPFC and NAc, implying a level of D1 receptor expression and modulation in both GABAergic and glutamatergic NAc projecting neurons, thus likely resulting in a more complex pattern of neural activity in response to dopamine. More work is needed to fully understand the expression patterns of dopamine receptors of GABAergic and glutamatergic NAc projecting neurons and the consequences of rising dopamine on neuronal excitability.

Serotonin

Serotonergic modulation, originating from the dorsal raphe nucleus, is another potential source of the reward-mediated suppression we observe in mPFC and vHip (Cid-Pellitero & Garzón, 2011; O'Hearn & Molliver, 1984; Segal, 1990; Yoshida et al., 2019). Projecting to both mPFC and vHip, serotonin neurons in the dorsal raphe nucleus have been shown to modulate their firing rate to the level of reward in the environment, a pattern of encoding that parallels patterns of outcome encoding we observe in mPFC and vHip (Cohen et al., 2015). Similar to dopamine, this could define a pattern of modulation where the concentration of serotonin in mPFC and vHip correlates with the amount of reward received. Serotonergic signaling at 5-HT1A receptors, a common serotonin receptor that decreases excitability and is highly expressed in both glutamatergic and GABAergic cells of mPFC and vHip, could then change the sign of the signaling, decreasing the excitability of neurons as the concentration of serotonin increases (Garcia-Garcia et al., 2014; Hernandes et al., 2021; Lladó-Pelfort et al., 2012, 2012; Santana & Artigas, 2017; Segal, 1990).

Norepinephrine

Both the mPFC and the vHip also receive noradrenergic modulation originating from the locus coeruleus (Bouras et al., 2023; Lipski & Grace, 2013). Though much less is known about the role of noradrenergic modulation in these regions, there is increasing evidence that this might play an important role in supporting motivation and goal directed-behavior (Hofmeister & Sterpenich, 2015). For example, when monkeys are taught to perform a task where they press a bar for a reward following a cue that

predicts reward size, locus coeruleus activity before action initiation increases as the cued reward size decreases (Bouret & Richmond, 2015). Though this study did not directly look at the relationship between latency and locus coeruleus activity, the findings of this study follow a pattern of activity that is reminiscent of the findings of this thesis whereas activity increases as a function of the amount of unrewarded outcomes encountered, response latency decreases. This could define a pattern of modulation whereby as the amount of reward decreases in the environment, noradrenergic input to mPFC and vHip increases, increasing the excitability of neurons in these regions and giving rise to the neural encoding of reward history we observe.

Acetylcholine

The basal forebrain sends cholinergic projections to both the mPFC and vHip and is thought to play an important role in regulating attention (Bloem et al., 2014; Rapanelli et al., 2023; Wu et al., 2024). However, much less is known about the role of cholinergic modulation in supporting outcome integration and reward-motivated behavior. Recent work has shown that cholinergic tone in the basal forebrain is modulated by reward such that when performing reward-seeking behavior, the delivery of reward itself strongly suppresses acetylcholine release (Hanson et al., 2021). This is similar to the general pattern of activity we describe in mPFC and vHip populations, suggesting a potential role of cholinergic modulation in mediating outcome integration.

While there is a case to be made for the involvement of each of these neuromodulators in supporting integration of reward history across mPFC and vHip, the observed modulation of activity that emerges across the ITI is likely a product of the interactions between a variety of neuromodulators signaling reward-associated

information at different timescales. This invites further study to better characterize the role of each of these neuromodulators in reward integration, both individually and in tandem. Recent advances now allow for the use of genetically encoded fluorescent indicators to record neurotransmitter release *in vivo*. This includes constructs such as GrabDA for dopamine, GRAB_5-HT for serotonin, GACh for acetocholine, and nLightG for norepinephrine (Jing et al., 2018; Kagiampaki et al., 2023; Wan et al., 2021; Zhuo et al., 2024). Systematically recording the release of these neurotransmitters in mPFC and vHip while animals perform a two-armed bandit task and subsequent task degredations, as described in chapters 2 and 3, would reveal the extent of their contribution towards outcome encoding and reward integration in mPFC and vHip.

Reward integration across mPFC, vHip, and NAc

Chapter 2 of this thesis revealed that mPFC and vHip projections to NAc integrate reward history by suppressing activity to reward. In chapter 3 once again we find this pattern of encoding in the GABAergic neurons of these regions. The similarity in reward-encoding motifs across these two studies suggests common sources of modulation in mPFC and vHip, which we have speculated on in the previous section. However, one notable difference in the patterns of reward integration across the two studies is the granularity of reward encoding. While GABAergic neurons in mPFC and vHip seem to encode a complete outcome history, at least when exploiting, glutamatergic NAc-projecting neurons from these regions seem to be encoding more of a history of unrewarded outcomes with a singular reward maximally suppressing activity with each consecutive unrewarded outcome gradually increasing activity. This suggests that while mPFC and vHip may be encoding a full outcome history at a region-level,

there is some sort of processing or filtering of information that is occurring, transforming a fully gradated outcome history to a simplified history of unrewarded outcomes such that neural activity is inversely correlated with the amount of reward experienced. We also note that this transformed encoding pattern is also observed in mPFC and vHip GABAergic neurons when animals are in an exploit state. However, because we did not analyze how state modulates outcome encoding in mPFC and vHip projection activity, we cannot say for certain if encoding is further transformed in NAc-projecting neurons when animals are exploiting. However, this body of works suggests that even after further transformation, the information would be more oriented towards a history of unrewarded outcomes.

This history of unrewarded outcomes is then transmitted downstream such that many unrewarded outcomes should increase input from mPFC and vHip to the NAc to suppress engagement. Optogenetic manipulations from the first set of studies confirm this, demonstrating that strong activation of either mPFC of vHip can suppress engagement. However, weaker activation requires synchronous recruitment to suppress engagement. This suggests that while more balanced activity, resulting from similar integration of outcome history across mPFC and vHip inputs to NAc modulates engagement, strong activation of either input may allow for a given circuit to exert more direct behavioral control. For example, the vHip input to NAc appears to be more sensitive to unrewarded outcomes, ceasing to encode outcome history after consecutive rewards. This potentially defines a mechanism whereby in environments where reward is less consistent, the vHip input to NAc exerts more behavioral control.

mPFC and vHip projections to NAc both integrate reward-related information, regulating the activity of MSNs in the NAc shell. These MSNs go on to modulate activity in downstream regions that exert more direct control over motor output. The ventral pallidum (VP) is one such region, representing the primary output of MSNs in the NAc shell and receiving input from both D1 and D2 MSNs, with D1 MSNs also collateralizing to VTA (Kupchik et al., 2015; Pardo-Garcia et al., 2019). Activity in this region has been implicated in motivated behavior, encoding information about relative value of reward, reward-associated cues, and modulating motivation and projecting to motor regions of the brain stem (Soares-Cunha & Heinsbroek, 2023). Specifically, VP activity appears to modulate engagement, with optogenetic stimulation in a lever pressing task, decreasing latency to lever press, consistent with our observations of mPFC-NAc and vHip-NAc. Altogether this work suggests a cascade of activity whereby reward-relevant information is conveyed to and integrated by mPFC and vHip, wherein reward suppresses and unrewarded outcomes increase neural activity. This information is then transformed to encode a history of unrewarded outcomes where reward now maximally suppresses NAc-projecting cells and activity scales with the number of unrewarded outcomes encountered. This should then modulate the activity of MSNs in the NAc such that they provide increasing inhibition of the VP as the number of unrewarded outcomes increase, thus decreasing engagement.

Neural encoding and behavioral demands

Much of the extant work on reward is predicated on the concept that reward-encoding is static and invariant. That is, that these neural processes and correlates hold true regardless of task demands or behavioral state. A large part of this is due to the

statistical approaches Yet existing studies have not generally modulated task requirements to test this assumption. By explicitly varying task demands, this thesis, provides strong evidence that behavioral demands shape neural encoding. In chapter 2 where we examine neural encoding of reward history in mPFC and vHip projections to NAc, we demonstrate that when animals no longer have to perform a specific action for a reward, encoding of reward history simplifies to only encode the previous outcome in mPFC-NAc and to encode consecutively unrewarded outcomes against other outcome histories in vHip-NAc. This demonstrates that the degree of complexity in reward encoding can be modulated directly by the complexity of the task. These changes in behavioral demands likely affect the behavioral state the animal is in and thus the neural mechanisms utilized to integrate outcomes and support reward-motivated behavior. We provide further evidence of this in chapter 3 where we examine neural encoding of reward history in mPFC and vHip GABAergic neurons. Here we find variations in behavioral state map on to differences in encoding. When exploring, mPFC and vHip GABAergic neurons encode a full reward history, however, once exploiting, the neural representation becomes anchored to unrewarded outcomes, likely driven by the decreased utility of reward-oriented information during exploitation. This demonstrates that as behavioral demands change as animals shift between exploration and exploitation states, neural encoding of the same reward information changes. Together this provides strong evidence that neural encoding is highly intertwined with behavioral demands, and thus behavioral state.

Limitations and future directions

This thesis presents a set of studies that identify a novel motif of reward integration implemented across glutamatergic and GABAergic neuronal populations in mPFC and vHip and demonstrate how these patterns of neural encoding influence reward-motivated behavior. This thesis also presents clear evidence that neural encoding of reward is shaped by behavioral state. However, within this body of work there are technical and methodological limitations that leave some questions open and invite further investigation. A clear limitation of this thesis is the exclusive use of males in chapter 3. Though in chapter 2 we show that patterns of reward integration in mPFC and vHip projections to NAc are largely similar across sexes, this does not preclude the possibility that there exist sex difference in how GABAergic neurons integrate outcomes. As one example of a surprising sex difference, in other research, I found that while males use the mPFC projection to NAc to distinguish between threatening and non-threatening cues, females use the vHip projection to NAc underlining the possibility of latent sex differences in neural encoding (Muir et al., 2024).

In chapter 2, we utilized excitatory opsins to increase activity in mPFC and vHip projections to NAc during the ITI. This allowed us to establish that increases in neural activity in mPFC and vHip projections to NAc lead to decreases in engagement. However, I did not perform an inverse experiment using inhibitory opsins to decrease activity in these projections. Without exploring the effects of inhibition, we cannot say for certain if decreasing neural activity in these projections would lead to an increase in engagement. However, for a number of reasons, it is likely that such an experiment with inhibitory opsin would not provide definitive results. The signal that we observe in NAc

projecting cells is maximally suppressed by reward, with unrewarded outcomes gradually restoring activity. This means that the richest modulation occurs in the excitatory direction and suggests a lower bound for the effect of bulk inhibition of NAc projections on behavior. Additionally, a key finding of this thesis is the redundancy across the brain in control over engagement and reward-motivated behavior. Thus, it is equally likely that inhibitory manipulations would be without effect given compensatory input from other NAc projecting regions.

In chapter 2, stimulation was applied either across the entirety of the ITI or during the lever extension period. This allowed us to link neural activity specifically during the ITI and engagement. However, activity across the ITI is dynamic in both mPFC and vHip, with suppression emerging across the ITI. Additionally, the signal appears to have several components. In mPFC-NAc we see a peak to lever press, a slight increase in activity to reward delivery, and then suppression to reward. In vHip-NAc we see a peak to lever press, a suppression of activity, followed by an increase in activity to unrewarded outcomes. Further studies clarifying the role of these features in supporting reward-motivated behavior provide an exciting avenue of future research. For example, we note that each of these features appear to emerge at specific time-points across the ITI. Using BiPOLES for bidirectional optogenetic manipulation of activity in NAc-projecting neurons we can systematically ablate these features, for example suppressing mPFC-NAc activity when we expect a peak to lever press or stimulating vHip-NAc activity when we expect a suppression of activity (Vierock et al., 2021). This would reveal the specific contributions of each of these signal components towards supporting reward-motivated behavior and outcome integration. We could also use this

approach to manipulate GABAergic neuron activity in mPFC and vHip following our findings in chapter 3 to better understand the causal relationship between these neural populations and reward motivated behavior.

In this thesis, I describe how population-level activity is modulated as a consequence of interactions with rewarded and unrewarded outcomes. However, the specific neural basis of this modulation remains unclear. This limitation is due to the use of fiber photometry as the technique for recording neural activity. Fiber photometry, though useful for its high throughput and cell-type specificity, lacks single cell resolution. That is, while we are able to describe at a population level how various projections and cell-types process reward, we are unable to determine how this manifests at a single cell level. This leads to several interesting questions that remain unanswered. For example, it is unclear if outcome integration is being performed by an entire population or just a subset of cells. Likewise, we observe an increase in activity following unrewarded outcomes compared to rewarded outcomes. This could be encoded by two separate populations, one signaling reward and another signaling the omission of a rewarded outcome with the latter population comprising of a larger number of cells and representing the primary site of modulation by outcome history. The other possibility is that both reward and the omission of reward are signaled by the same population of cells with overall activity modulated by outcome history. Each possibility would have a different implication for the neural mechanisms governing outcome integrations.

Another unanswered question is the neural substrate of the change observed in mPFC-NAc and vHip-NAc outcome integration following degradations in task structure. As elements such as choice and action are removed from the environment, the neural

signal in these pathways become more and more noisy. It could be the case that this increase in noise is due to the fact that encoding becomes less specific as the complexity of task structure decreases. The other possibility is that the as task complexity decreases so does the size of the neural population encoding outcome. Future studies utilizing neural recording techniques at a cellular resolution such as microendoscopy or *in vivo* electrophysiology are needed to answer these questions.

Chapter 5

Conclusions

Prior to this thesis, the role of glutamatergic afferents in supporting reward processing remained relatively unexplored. The primary goal of this thesis was to address this gap, by determining redundancy and specialization in how two distinct glutamatergic inputs to the NAc, mPFC-NAc and vHip-NAc, process reward and contribute towards reward-motivated behavior.

In the first set of studies, I examined how population-level activity in mPFC and vHip projections to NAc encode information about outcomes, finding that reward drives suppression of neural activity and unrewarded outcomes gradually restore activity. I then described how this motif can be implemented across many trials, tracking prior reward history as a graded function of unrewarded outcomes. Using task degradations and conditional entropy analyses I find that while the mPFC projection to NAc consistently encodes outcome, encoding in the projection from vHip to NAc is anchored to unrewarded outcomes. Examining the behavioral relevance of reward integration in mPFC-NAc and vHip-NAc revealed a co-operative and somewhat redundant role of these pathways in dynamically modulating engagement in rewarding environments.

The second study expanded upon this work and was based on the hypothesis that inhibitory populations in mPFC and vHip might contribute to the reward-mediated suppression identified in the first study. Surprisingly, I again found a similar pattern of reward integration as described in the glutamatergic projection neurons. Inhibitory

neurons in mPFC and vHip suppressed activity to reward, tracking prior reward history. However, in inhibitory neurons, this seemed to be gated by behavioral state. Examining the behavioral relevance of state-gated reward integration in mPFC and vHip again revealed a role in dynamically modulating task engagement in both regions, but a state-dependent role in modulating choice behavior in each region.

Together, this work identifies a novel motif of outcome integration implemented across various important reward-processing neural populations in mPFC and vHip. In this motif, the level of activity inversely correlates with the amount and recency of reward received. Critically, we show that at a neural level, implementation of this outcome-integrating motif is dependent on task demands and behavioral state. While we demonstrate some state-dependent associations between neural activity and choice behavior in GABAergic neurons, we find that the primary role of these neurons as well as of NAc-projecting neurons is to mediate engagement, a relationship that has historically been less well characterized than that of choice. Finally, this work highlights the cooperative roles of mPFC and vHip in supporting outcome integration and reward-motivated behavior.

These results challenge the traditional view that reward processing can be neatly divided among different brain regions and cell types, and instead propose that neural encoding is more likely a distributed function, simultaneously occurring across multiple brain regions and cell types. This is evidenced in the by the repeated discovery of a reward integration motif with similar behavioral correlates across GABAergic neurons and glutamatergic projections neurons in both the mPFC and vHip. These findings underline not only the high level of redundancy in the brain but also the importance of

putting neural circuits in context with each other. At first glance, each cell type examined in this thesis appears to be performing the same function: suppressing activity to reward, however, we find that the specialization of each cell type in reward encoding only becomes apparent when compared directly to other cells in the broader circuit. This suggests that differences in cell-type and region-specific encoding are highly nuanced and that coarse-grained interrogations of encoding are likely inadequate to explain the relevance of encoding. This is further complicated by evidence provided in this thesis that encoding can also be differentially modulated by environmental, behavioral, and other latent factors, highlighting the importance in exploring and defining the limits associated with encoding. Overall, the significant redundancy within the brain likely enables the maintenance of high-level behavior, utilizing gradual shifts in the balance between brain regions and cell types to further fine-tune behavior, resulting in continuous and highly adaptive behavior.

The brain is a complex and highly redundant structure. Taken together, this work underlines the importance of simultaneously considering multiple cell-types and circuits in context as we work towards developing an integrative understanding of how the brain continuously processes information to produce ever-adapting behaviors.

References

Abboud, S., & Sadeh, D. (1984). The use of cross-correlation function for the alignment of ECG waveforms and rejection of extrasystoles. *Computers and Biomedical Research*, *17*(3), 258–266. https://doi.org/10.1016/S0010-4809(84)80017-8

Abela, A. R., Duan, Y., & Chudasama, Y. (2015). Hippocampal interplay with the nucleus accumbens is critical for decisions about time. *The European Journal of Neuroscience*, *42*(5), 2224–2233. https://doi.org/10.1111/ejn.13009

Aberman, J. E., & Salamone, J. D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience*, *92*(2), 545–552. https://doi.org/10.1016/S0306-4522(99)00004-4

Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron*, *59*(4), 648–661. https://doi.org/10.1016/j.neuron.2008.07.004

Amir, N., Prouvost, C., & Kuckertz, J. M. (2012). Lack of a Benign Interpretation Bias in Social Anxiety Disorder. *Cognitive Behaviour Therapy*, *41*(2), 119–129. https://doi.org/10.1080/16506073.2012.662655

Bagot, R. C., Parise, E. M., Peña, C. J., Zhang, H.-X., Maze, I., Chaudhury, D., Persaud, B., Cachope, R., Bolaños-Guzmán, C. A., Cheer, J. F., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2015). Ventral hippocampal afferents to the nucleus accumbens regulate susceptibility to depression. *Nature Communications*, *6*, 7062. https://doi.org/10.1038/ncomms8062

Baldo, B. A., Sadeghian, K., Basso, A. M., & Kelley, A. E. (2002). Effects of selective dopamine D1 or D2 receptor blockade within nucleus accumbens subregions on ingestive behavior and associated motor activity. *Behavioural Brain Research*, *137*(1), 165–177. https://doi.org/10.1016/S0166-4328(02)00293-0

Barbano, M. F., Qi, J., Chen, E., Mohammad, U., Espinoza, O., Candido, M., Wang, H., Liu, B., Hahn, S., Vautier, F., & Morales, M. (2024). VTA glutamatergic projections to the nucleus accumbens suppress psychostimulant-seeking behavior. *Neuropsychopharmacology*, 1–11. https://doi.org/10.1038/s41386-024-01905-3

Bari, B. A., Grossman, C. D., Lubin, E. E., Rajagopalan, A. E., Cressy, J. I., & Cohen, J. Y. (2019). Stable representations of decision variables for flexible behavior. *Neuron*, *103*(5), 922-933.e7. https://doi.org/10.1016/j.neuron.2019.06.001

Barker, J. M., Bryant, K. G., & Chandler, L. J. (2019). Inactivation of ventral hippocampus projections promotes sensitivity to changes in contingency. *Learning & Memory*, *26*(1), 1–8.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv Preprint arXiv:1406.5823*.

Bathellier, B., Tee, S. P., Hrovat, C., & Rumpel, S. (2013). A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. *Proceedings of the National Academy of Sciences*, *110*(49), 19950–19955. https://doi.org/10.1073/pnas.1312125110

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., & Dayan, P. (2013). Dopamine Modulates Reward-Related Vigor.

Neuropsychopharmacology, *38*(8), 1495–1503. https://doi.org/10.1038/npp.2013.48

Belujon, P., & Grace, A. A. (2008). Critical role of the prefrontal cortex in the regulation of hippocampus-accumbens information flow. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(39), 9797–9805. https://doi.org/10.1523/JNEUROSCI.2200-08.2008

Bercovici, D. A., Princz-Lebel, O., Tse, M. T., Moorman, D. E., & Floresco, S. B. (2018). Optogenetic Dissection of Temporal Dynamics of Amygdala-Striatal Interplay during Risk/Reward Decision Making. *eNeuro*, *5*(6), ENEURO.0422-18.2018. https://doi.org/10.1523/ENEURO.0422-18.2018

Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology*, *191*(3), 391–431. https://doi.org/10.1007/s00213-006-0578-x

Beyeler, A., Namburi, P., Glober, G. F., Simonnet, C., Calhoon, G. G., Conyers, G. F., Luck, R., Wildes, C. P., & Tye, K. M. (2016). Divergent routing of positive and negative information from the amygdala during memory retrieval. *Neuron*, *90*(2), 348–361. https://doi.org/10.1016/j.neuron.2016.03.004

Bishop, S. J., & Gagne, C. (2018). Anxiety, Depression, and Decision Making: A Computational Perspective. *Annual Review of Neuroscience*, *41*, 371–388. https://doi.org/10.1146/annurev-neuro-080317-062007

Bloem, B., Poorthuis, R. B., & Mansvelder, H. D. (2014). Cholinergic modulation of the medial prefrontal cortex: The role of nicotinic receptors in attention and regulation

of neuronal activity. *Frontiers in Neural Circuits*, *8*, 17. https://doi.org/10.3389/fncir.2014.00017

Borrego, M. B., Grigsby, K. B., Townsley, K. G., Chan, A., Firsick, E. F., Tran, A., & Ozburn, A. R. (2022). Central nucleus of the amygdala projections onto the nucleus accumbens core regulate binge-like alcohol drinking in a CRF-dependent manner. *Neuropharmacology*, *203*, 108874. https://doi.org/10.1016/j.neuropharm.2021.108874

Bouarab, C., Thompson, B., & Polter, A. M. (2019). VTA GABA Neurons at the Interface of Stress and Reward. *Frontiers in Neural Circuits*, *13*. https://doi.org/10.3389/fncir.2019.00078

Bouras, N. N., Mack, N. R., & Gao, W.-J. (2023). Prefrontal modulation of anxiety through a lens of noradrenergic signaling. *Frontiers in Systems Neuroscience*, *17*, 1173326. https://doi.org/10.3389/fnsys.2023.1173326

Bouret, S., & Richmond, B. J. (2015). Sensitivity of Locus Ceruleus Neurons to Reward Value for Goal-Directed Actions. *The Journal of Neuroscience*, *35*(9), 4005–4014. https://doi.org/10.1523/JNEUROSCI.4553-14.2015

Britt, J. P., Benaliouad, F., McDevitt, R. A., Stuber, G. D., Wise, R. A., & Bonci, A. (2012). Synaptic and Behavioral Profile of Multiple Glutamatergic Inputs to the Nucleus Accumbens. *Neuron*, *76*(4), 790–803. https://doi.org/10.1016/j.neuron.2012.09.040

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Brown, V. J., & Bowman, E. M. (1995). Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. *The European Journal of Neuroscience*, *7*(12), 2479–2485. https://doi.org/10.1111/j.1460-9568.1995.tb01046.x

Calipari, E. S., Bagot, R. C., Purushothaman, I., Davidson, T. J., Yorgason, J. T., Peña, C. J., Walker, D. M., Pirpinias, S. T., Guise, K. G., Ramakrishnan, C., Deisseroth, K., & Nestler, E. J. (2016). In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proceedings of the National Academy of Sciences*, *113*(10), 2726–2731. https://doi.org/10.1073/pnas.1521238113

Carter, A. G., Soler-Llavina, G. J., & Sabatini, B. L. (2007). Timing and Location of Synaptic Inputs Determine Modes of Subthreshold Integration in Striatal Medium Spiny Neurons. *The Journal of Neuroscience*, *27*(33), 8967–8977. https://doi.org/10.1523/JNEUROSCI.2798-07.2007

Carvalho Poyraz, F., Holzner, E., Bailey, M. R., Meszaros, J., Kenney, L., Kheirbek, M. A., Balsam, P. D., & Kellendonk, C. (2016). Decreasing Striatopallidal Pathway Function Enhances Motivation by Energizing the Initiation of Goal-Directed Action. *The Journal of Neuroscience*, *36*(22), 5988–6001. https://doi.org/10.1523/JNEUROSCI.0444-16.2016

Chen, C. S., Knep, E., Han, A., Ebitz, R. B., & Grissom, N. M. (2021). Sex differences in learning from exploration. *eLife*, *10*, e69748. https://doi.org/10.7554/eLife.69748

Chen, C. S., Mueller, D., Knep, E., Ebitz, R. B., & Grissom, N. M. (2024). Dopamine and Norepinephrine Differentially Mediate the Exploration-Exploitation Tradeoff. *The*

Journal of Neuroscience: The Official Journal of the Society for Neuroscience, *44*(44), e1194232024. https://doi.org/10.1523/JNEUROSCI.1194-23.2024

Chen, G., Lai, S., Bao, G., Ke, J., Meng, X., Lu, S., Wu, X., Xu, H., Wu, F., Xu, Y., Xu, F., Bi, G.-Q., Peng, G., Zhou, K., & Zhu, Y. (2023). Distinct reward processing by subregions of the nucleus accumbens. *Cell Reports*, *42*(2). https://doi.org/10.1016/j.celrep.2023.112069

Christoffel, D. J., Walsh, J. J., Hoerbelt, P., Heifets, B. D., Llorach, P., Lopez, R. C., Ramakrishnan, C., Deisseroth, K., & Malenka, R. C. (2021). Selective filtering of excitatory inputs to nucleus accumbens by dopamine and serotonin. *Proceedings of the National Academy of Sciences*, *118*(24), e2106648118. https://doi.org/10.1073/pnas.2106648118

Cid-Pellitero, E. del, & Garzón, M. (2011). Medial prefrontal cortex receives input from dorsal raphe nucleus neurons targeted by Hypocretin1/OrexinA-containing axons. *Neuroscience*, *172*, 30–43. https://doi.org/10.1016/j.neuroscience.2010.10.058

Cohen, J. Y., Amoroso, M. W., & Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. *eLife*, *4*, e06346. https://doi.org/10.7554/eLife.06346

Collins, D. P., Anastasiades, P. G., Marlin, J. J., & Carter, A. G. (2018). Reciprocal circuits linking the prefrontal cortex with dorsal and ventral thalamic nuclei. *Neuron*, *98*(2), 366-379.e4. https://doi.org/10.1016/j.neuron.2018.03.024

Cover, T. M., & Thomas, J. A. (1991). Entropy, relative entropy and mutual information. *Elements of Information Theory*, *2*(1), 12–13.

Cox, J., Minerva, A. R., Fleming, W. T., Zimmerman, C. A., Hayes, C., Zorowitz, S., Bandi, A., Ornelas, S., McMannon, B., Parker, N. F., & Witten, I. B. (2023). A neural substrate of sex-dependent modulation of motivation. *Nature Neuroscience*, *26*(2), Article 2. https://doi.org/10.1038/s41593-022-01229-9

Dalton, G. L., Phillips, A. G., & Floresco, S. B. (2014). Preferential Involvement by Nucleus Accumbens Shell in Mediating Probabilistic Learning and Reversal Shifts. *Journal of Neuroscience*, *34*(13), 4618–4626. https://doi.org/10.1523/JNEUROSCI.5058-13.2014

Dana, H., Sun, Y., Mohar, B., Hulse, B. K., Kerlin, A. M., Hasseman, J. P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., Macklin, J. J., Chen, Y., Konnerth, A., Jayaraman, V., Looger, L. L., Schreiter, E. R., Svoboda, K., & Kim, D. S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods*, *16*(7), 649–657. https://doi.org/10.1038/s41592-019-0435-6

Day, J. J., Jones, J. L., & Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs. *The European Journal of Neuroscience*, *33*(2), 308–321. https://doi.org/10.1111/j.1460-9568.2010.07531.x

De Groote, A., & de Kerchove d'Exaerde, A. (2021). Thalamo-Nucleus Accumbens Projections in Motivated Behaviors and Addiction. *Frontiers in Systems Neuroscience*, *15*, 711350. https://doi.org/10.3389/fnsys.2021.711350

de Jong, J. W., Afjei, S. A., Dorocic, I. P., Peck, J. R., Liu, C., Kim, C. K., Tian, L., Deisseroth, K., & Lammel, S. (2019). A neural circuit mechanism for encoding

aversive stimuli in the mesolimbic dopamine system. *Neuron*, *101*(1), 133-151.e7. https://doi.org/10.1016/j.neuron.2018.11.005

Dezfouli, A., Balleine, B. W., & Nock, R. (2019). Optimal response vigor and choice under non-stationary outcome values. *Psychonomic Bulletin & Review*, *26*(1), 182–204. https://doi.org/10.3758/s13423-018-1500-3

Dieterich, A., Floeder, J., Stech, K., Lee, J., Srivastava, P., Barker, D. J., & Samuels, B. A. (2021). Activation of Basolateral Amygdala to Nucleus Accumbens Projection Neurons Attenuates Chronic Corticosterone-Induced Behavioral Deficits in Male Mice. *Frontiers in Behavioral Neuroscience*, *15*. https://doi.org/10.3389/fnbeh.2021.643272

Do-Monte, F. H., Minier-Toribio, A., Quiñones-Laracuente, K., Medina-Colón, E. M., & Quirk, G. J. (2017). Thalamic regulation of sucrose-seeking during unexpected reward omission. *Neuron*, *94*(2), 388-400.e4. https://doi.org/10.1016/j.neuron.2017.03.036

Dong, X., Li, S., & Kirouac, G. J. (2017). Collateralization of projections from the paraventricular nucleus of the thalamus to the nucleus accumbens, bed nucleus of the stria terminalis, and central nucleus of the amygdala. *Brain Structure & Function*, *222*(9), 3927–3943. https://doi.org/10.1007/s00429-017-1445-8

Ebitz, R. B., Albarran, E., & Moore, T. (2018). Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex. *Neuron*, *97*(2), 450-461.e9. https://doi.org/10.1016/j.neuron.2017.12.007

Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). Reinforcement learning and Bayesian inference provide complementary models

for the unique advantage of adolescents in stochastic reversal. *Developmental Cognitive Neuroscience*, *55*, 101106. https://doi.org/10.1016/j.dcn.2022.101106

Engelke, D. S., Zhang, X. O., O'Malley, J. J., Fernandez-Leon, J. A., Li, S., Kirouac, G. J., Beierlein, M., & Do-Monte, F. H. (2021). A hypothalamic-thalamostriatal circuit that controls approach-avoidance conflict in rats. *Nature Communications*, *12*, 2517. https://doi.org/10.1038/s41467-021-22730-y

Fetcho, R. N., Hall, B. S., Estrin, D. J., Walsh, A. P., Schuette, P. J., Kaminsky, J., Singh, A., Roshgodal, J., Bavley, C. C., Nadkarni, V., Antigua, S., Huynh, T. N., Grosenick, L., Carthy, C., Komer, L., Adhikari, A., Lee, F. S., Rajadhyaksha, A. M., & Liston, C. (2023). Regulation of social interaction in mice by a frontostriatal circuit modulated by established hierarchical relationships. *Nature Communications*, *14*, 2487. https://doi.org/10.1038/s41467-023-37460-6

Fisone, G., Bonito-Oliva, A., & Feyder, M. (2011). Deciphering the Actions of Antiparkinsonian and Antipsychotic Drugs on cAMP/DARPP-32 Signaling. *Frontiers in Neuroanatomy*, *5*. https://doi.org/10.3389/fnana.2011.00038

Floresco, S. B. (2007). Dopaminergic regulation of limbic-striatal interplay. *Journal of Psychiatry & Neuroscience : JPN*, *32*(6), 400–411.

Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of Psychology*, *66*(1), 25–52. https://doi.org/10.1146/annurev-psych-010213-115159

Floresco, S. B., Montes, D. R., Tse, M. M. T., & van Holstein, M. (2018). Differential Contributions of Nucleus Accumbens Subregions to Cue-Guided Risk/Reward Decision Making and Implementation of Conditional Rules. *The Journal of*

Neuroscience, *38*(8), 1901–1914. https://doi.org/10.1523/JNEUROSCI.3191-17.2018

Francis, T. C., & Lobo, M. K. (2017). Emerging role for nucleus accumbens medium spiny neuron subtypes in depression. *Biological Psychiatry*, *81*(8), 645–653. https://doi.org/10.1016/j.biopsych.2016.09.007

French, S. J., & Totterdell, S. (2002). Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *Journal of Comparative Neurology*, *446*(2), 151–165.

Gallo, E. F., Meszaros, J., Sherman, J. D., Chohan, M. O., Teboul, E., Choi, C. S., Moore, H., Javitch, J. A., & Kellendonk, C. (2018). Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nature Communications*, *9*, 1086. https://doi.org/10.1038/s41467-018-03272-2

Garcia-Garcia, A., Tancredi, A. N.-, & Leonardo, E. D. (2014). 5-HT1A receptors in mood and anxiety: Recent insights into autoreceptor versus heteroreceptor function. *Psychopharmacology*, *231*(4), 623–636. https://doi.org/10.1007/s00213-013-3389-x

Gerfen, C. R. (2023). Segregation of D1 and D2 dopamine receptors in the striatal direct and indirect pathways: An historical perspective. *Frontiers in Synaptic Neuroscience*, *14*, 1002960. https://doi.org/10.3389/fnsyn.2022.1002960

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D1 and D2 Dopamine Receptor-regulated Gene Expression

of Striatonigral and Striatopallidal Neurons. *Science*, *250*(4986), 1429–1432. https://doi.org/10.1126/science.2147780

Goto, Y., & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, *8*(6), Article 6. https://doi.org/10.1038/nn1471

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences*, *30*(5), 220–227. https://doi.org/10.1016/j.tins.2007.03.003

Guillaumin, M. C. C., Viskaitis, P., Bracey, E., Burdakov, D., & Peleg-Raibstein, D. (2023). Disentangling the role of NAc D1 and D2 cells in hedonic eating. *Molecular Psychiatry*, *28*(8), 3531–3547. https://doi.org/10.1038/s41380-023-02131-x

Guitart-Masip, M., Beierholm, U. R., Dolan, R., Duzel, E., & Dayan, P. (2011). Vigor in the Face of Fluctuating Rates of Reward: An Experimental Examination. *Journal of Cognitive Neuroscience*, *23*(12), 3933–3938. https://doi.org/10.1162/jocn_a_00090

Gustafson, N. J., & Daw, N. D. (2011). Grid Cells, Place Cells, and Geodesic Generalization for Spatial Reinforcement Learning. *PLOS Computational Biology*, *7*(10), e1002235. https://doi.org/10.1371/journal.pcbi.1002235

Hamel, L., Cavdaroglu, B., Yeates, D., Nguyen, D., Riaz, S., Patterson, D., Khan, N., Kirolos, N., Roper, K., Ha, Q. A., & Ito, R. (2022). Cortico-Striatal Control over Adaptive Goal-Directed Responding Elicited by Cues Signaling Sucrose Reward

or Punishment. *Journal of Neuroscience*, *42*(18), 3811–3822. https://doi.org/10.1523/JNEUROSCI.2175-21.2022

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2016). Mesolimbic Dopamine Signals the Value of Work. *Nature Neuroscience*, *19*(1), 117–126. https://doi.org/10.1038/nn.4173

Han, X., Jing, M., Zhao, T., Wu, N., Song, R., & Li, J. (2017). Role of dopamine projections from ventral tegmental area to nucleus accumbens and medial prefrontal cortex in reinforcement behaviors assessed using optogenetic manipulation. *Metabolic Brain Disease*, *32*(5), 1491–1502. https://doi.org/10.1007/s11011-017-0023-3

Hanson, E., Brandel-Ankrapp, K. L., & Arenkiel, B. R. (2021). Dynamic Cholinergic Tone in the Basal Forebrain Reflects Reward-Seeking and Reinforcement During Olfactory Behavior. *Frontiers in Cellular Neuroscience*, *15*, 635837. https://doi.org/10.3389/fncel.2021.635837

He, Y., Huang, Y. H., Schlüter, O. M., & Dong, Y. (2023). Cue- versus reward-encoding basolateral amygdala projections to nucleus accumbens. *eLife*, *12*, e89766. https://doi.org/10.7554/eLife.89766

Hernandes, P. M., Batistela, M. F., Vilela-Costa, H. H., Sant'Ana, A. B., Kumpel, V. D., Tirapelle, M. C., Bom, A. de O. P., de Andrade, T. G. C. S., & Zangrossi, H. (2021). Role of 5-HT1A receptors in the ventral hippocampus in the regulation of anxiety- and panic-related defensive behaviors in rats. *Behavioural Brain Research*, *408*, 113296. https://doi.org/10.1016/j.bbr.2021.113296

Herrnstein, R. J. (1997). *The matching law: Papers in psychology and economics*. Harvard University Press.

Hofmeister, J., & Sterpenich, V. (2015). A Role for the Locus Ceruleus in Reward Processing: Encoding Behavioral Energy Required for Goal-Directed Actions. *Journal of Neuroscience*, *35*(29), 10387–10389. https://doi.org/10.1523/JNEUROSCI.1734-15.2015

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, *500*(7464), Article 7464. https://doi.org/10.1038/nature12475

Huang, S., Borgland, S. L., & Zamponi, G. W. (2019). Dopaminergic modulation of pain signals in the medial prefrontal cortex: Challenges and perspectives. *Neuroscience Letters*, *702*, 71–76. https://doi.org/10.1016/j.neulet.2018.11.043

Ikemoto, S., Glazier, B. S., Murphy, J. M., & McBride, W. J. (1997). Role of Dopamine D1 and D2 Receptors in the Nucleus Accumbens in Mediating Reward. *The Journal of Neuroscience*, *17*(21), 8580–8587. https://doi.org/10.1523/JNEUROSCI.17-21-08580.1997

Iyer, E. S., Vitaro, P., Wu, S., Muir, J., Tse, Y. C., Cvetkovska, V., & Bagot, R. C. (2024). *Reward integration in prefrontal-cortical and ventral-hippocampal nucleus accumbens inputs cooperatively modulates engagement* (p. 2024.06.27.601063). bioRxiv. https://doi.org/10.1101/2024.06.27.601063

Jeong, H., Kim, D., Song, M., Paik, S.-B., & Jung, M. W. (2020). Distinct roles of parvalbumin- and somatostatin-expressing neurons in flexible representation of

task variables in the prefrontal cortex. *Progress in Neurobiology*, *187*, 101773. https://doi.org/10.1016/j.pneurobio.2020.101773

Jeong, H., Taylor, A., Floeder, J. R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D. A., & Namboodiri, V. M. K. (2022). Mesolimbic dopamine release conveys causal associations. *Science*, *378*(6626), eabq6740. https://doi.org/10.1126/science.abq6740

Jing, M., Zhang, P., Wang, G., Feng, J., Mesik, L., Zeng, J., Jiang, H., Wang, S., Looby, J. C., Guagliardo, N. A., Langma, L. W., Lu, J., Zuo, Y., Talmage, D. A., Role, L. W., Barrett, P. Q., Zhang, L. I., Luo, M., Song, Y., … Li, Y. (2018). A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nature Biotechnology*, *36*(8), 726–737. https://doi.org/10.1038/nbt.4184

Kagiampaki, Z., Rohner, V., Kiss, C., Curreli, S., Dieter, A., Wilhelm, M., Harada, M., Duss, S. N., Dernic, J., Bhat, M. A., Zhou, X., Ravotto, L., Ziebarth, T., Wasielewski, L. M., Sönmez, L., Benke, D., Weber, B., Bohacek, J., Reiner, A., … Patriarchi, T. (2023). Sensitive multicolor indicators for monitoring norepinephrine in vivo. *Nature Methods*, *20*(9), 1426–1436. https://doi.org/10.1038/s41592-023-01959-z

Kato, T., Mitsukura, Y., Yoshida, K., Mimura, M., Takata, N., & Tanaka, K. F. (2022). Oscillatory Population-Level Activity of Dorsal Raphe Serotonergic Neurons Is Inscribed in Sleep Structure. *The Journal of Neuroscience*, *42*(38), 7244–7255. https://doi.org/10.1523/JNEUROSCI.2288-21.2022

Kebabian, J. W., & Calne, D. B. (1979). Multiple receptors for dopamine. *Nature*, *277*(5692), 93–96. https://doi.org/10.1038/277093a0

Keiflin, R., & Janak, P. H. (2015). Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron*, *88*(2), 247–263. https://doi.org/10.1016/j.neuron.2015.08.037

Keistler, C., Barker, J. M., & Taylor, J. R. (2015). Infralimbic prefrontal cortex interacts with nucleus accumbens shell to unmask expression of outcome-selective Pavlovian-to-instrumental transfer. *Learning & Memory*, *22*(10), 509–513. https://doi.org/10.1101/lm.038810.115

Klausberger, T., & Somogyi, P. (2008). Neuronal Diversity and Temporal Dynamics: The Unity of Hippocampal Circuit Operations. *Science (New York, N.Y.)*, *321*(5885), 53–57. https://doi.org/10.1126/science.1149381

Koo, J. W., Lobo, M. K., Chaudhury, D., Labonté, B., Friedman, A., Heller, E., Peña, C. J., Han, M.-H., & Nestler, E. J. (2014). Loss of BDNF Signaling in D1R-Expressing NAc Neurons Enhances Morphine Reward by Reducing GABA Inhibition. *Neuropsychopharmacology*, *39*(11), 2646–2653. https://doi.org/10.1038/npp.2014.118

Köther, A. K., Alpers, G. W., Büdenbender, B., Lenhart, M., Michel, M. S., & Kriegmair, M. C. (2021). Predicting decisional conflict: Anxiety and depression in shared decision making. *Patient Education and Counseling*, *104*(5), 1229–1236. https://doi.org/10.1016/j.pec.2020.10.037

Kravitz, A. V., & Kreitzer, A. C. (2012). Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. *Physiology (Bethesda, Md.)*, *27*(3), 10.1152/physiol.00004.2012. https://doi.org/10.1152/physiol.00004.2012

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818. https://doi.org/10.1038/nn.3100

Kuchibhotla, K. V., Hindmarsh Sten, T., Papadoyannis, E. S., Elnozahy, S., Fogelson, K. A., Kumar, R., Boubenec, Y., Holland, P. C., Ostojic, S., & Froemke, R. C. (2019). Dissociating task acquisition from expression during learning reveals latent knowledge. *Nature Communications*, *10*(1), 2151. https://doi.org/10.1038/s41467-019-10089-0

Kupchik, Y. M., Brown, R. M., Heinsbroek, J. A., Lobo, M. K., Schwartz, D. J., & Kalivas, P. W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nature Neuroscience*, *18*(9), 1230–1232. https://doi.org/10.1038/nn.4068

Kutlu, M. G., Zachry, J. E., Melugin, P. R., Cajigas, S. A., Chevee, M. F., Kelly, S. J., Kutlu, B., Tian, L., Siciliano, C. A., & Calipari, E. S. (2021). Dopamine release in the nucleus accumbens core signals perceived saliency. *Current Biology*, *31*(21), 4748-4761.e8. https://doi.org/10.1016/j.cub.2021.08.052

Lafferty, C. K., Yang, A. K., Mendoza, J. A., & Britt, J. P. (2020). Nucleus Accumbens Cell Type- and Input-Specific Suppression of Unproductive Reward Seeking. *Cell Reports*, *30*(11), 3729-3742.e3. https://doi.org/10.1016/j.celrep.2020.02.095

Langdon, A. J., Hathaway, B. A., Zorowitz, S., Harris, C. B. W., & Winstanley, C. A. (2019). Relative insensitivity to time-out punishments induced by win-paired cues in a rat gambling task. *Psychopharmacology*, *236*(8), 2543–2556. https://doi.org/10.1007/s00213-019-05308-x

Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., & Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological Medicine*, *50*(10), 1613–1622. https://doi.org/10.1017/S0033291719001570

LeGates, T. A., Kvarta, M. D., Tooley, J. R., Francis, T. C., Lobo, M. K., Creed, M. C., & Thompson, S. M. (2018). Reward behavior is regulated by the strength of hippocampus-nucleus accumbens synapses. *Nature*, *564*(7735), 258–262. https://doi.org/10.1038/s41586-018-0740-8

Lenth, R. V., Buerkner, P., Herve, M., Love, J., Riebl, H., & Singmann, H. (2021). *Emmeans: Estimated marginal means, aka least-squares means. CRAN*.

Lind, E. B., Sweis, B. M., Asp, A. J., Esguerra, M., Silvis, K. A., David Redish, A., & Thomas, M. J. (2023). A quadruple dissociation of reward-related behaviour in mice across excitatory inputs to the nucleus accumbens shell. *Communications Biology*, *6*(1), Article 1. https://doi.org/10.1038/s42003-023-04429-6

Lindenbach, D., Vacca, G., Ahn, S., Seamans, J. K., & Phillips, A. G. (2022). Optogenetic modulation of glutamatergic afferents from the ventral subiculum to the nucleus accumbens: Effects on dopamine function, response vigor and locomotor activity. *Behavioural Brain Research*, *434*, 114028. https://doi.org/10.1016/j.bbr.2022.114028

Lipski, W. J., & Grace, A. A. (2013). Activation and inhibition of neurons in the hippocampal ventral subiculum by norepinephrine and locus coeruleus stimulation. *Neuropsychopharmacology: Official Publication of the American*

College of Neuropsychopharmacology, *38*(2), 285–292. https://doi.org/10.1038/npp.2012.157

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703–713. https://doi.org/10.1016/j.neuron.2005.05.002

Liu, X., & Carter, A. G. (2018). Ventral Hippocampal Inputs Preferentially Drive Corticocortical Neurons in the Infralimbic Prefrontal Cortex. *The Journal of Neuroscience*, *38*(33), 7351–7363. https://doi.org/10.1523/JNEUROSCI.0378-18.2018

Lladó-Pelfort, L., Santana, N., Ghisi, V., Artigas, F., & Celada, P. (2012). 5-HT1A receptor agonists enhance pyramidal cell firing in prefrontal cortex through a preferential action on GABA interneurons. *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(7), 1487–1497. https://doi.org/10.1093/cercor/bhr220

Lloyd, A., Roiser, J. P., Skeen, S., Freeman, Z., Badalova, A., Agunbiade, A., Busakhwe, C., DeFlorio, C., Marcu, A., Pirie, H., Saleh, R., Snyder, T., Fearon, P., & Viding, E. (2024). Reviewing explore/exploit decision-making as a transdiagnostic target for psychosis, depression, and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *24*(5), 793–815. https://doi.org/10.3758/s13415-024-01186-9

Lobo, M. K., Covington, H. E., Chaudhury, D., Friedman, A. K., Sun, H., Damez-Werno, D., Dietz, D., Zaman, S., Koo, J. W., Kennedy, P. J., Mouzon, E., Mogri, M., Neve, R. L., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2010). Cell Type Specific Loss of BDNF Signaling Mimics Optogenetic Control of Cocaine

Reward. *Science (New York, N.Y.)*, *330*(6002), 385–390. https://doi.org/10.1126/science.1188472

Loriaux, A. L., Roitman, J. D., & Roitman, M. F. (2011). Nucleus accumbens shell, but not core, tracks motivational value of salt. *Journal of Neurophysiology*, *106*(3), 1537–1544. https://doi.org/10.1152/jn.00153.2011

McGarry, L. M., & Carter, A. G. (2017). Prefrontal cortex drives distinct projection neurons in the basolateral amygdala. *Cell Reports*, *21*(6), 1426–1433. https://doi.org/10.1016/j.celrep.2017.10.046

Mingote, S., Amsellem, A., Kempf, A., Rayport, S., & Chuhma, N. (2019). Dopamine-Glutamate Neuron Projections to the Nucleus Accumbens Medial Shell and Behavioral Switching. *Neurochemistry International*, *129*, 104482. https://doi.org/10.1016/j.neuint.2019.104482

Mishchanchuk, K., Gregoriou, G., Qü, A., Kastler, A., Huys, Q. J. M., Wilbrecht, L., & MacAskill, A. F. (2024). Hidden state inference requires abstract contextual representations in the ventral hippocampus. *Science*, *386*(6724), 926–932. https://doi.org/10.1126/science.adq5874

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), Article 7759. https://doi.org/10.1038/s41586-019-1235-y

Moore, D. G., Valentini, G., Walker, S. I., & Levin, M. (2018). Inform: Efficient information-theoretic analysis of collective behaviors. *Frontiers in Robotics and AI*, *5*, 60.

Muir, J., Iyer, E. S., Tse, Y.-C., Sorensen, J., Wu, S., Eid, R. S., Cvetkovska, V., Wassef, K., Gostlin, S., Vitaro, P., Spencer, N. J., & Bagot, R. C. (2024). Sex-biased neural encoding of threat discrimination in nucleus accumbens afferents drives suppression of reward behavior. *Nature Neuroscience*, *27*(10), 1966–1976. https://doi.org/10.1038/s41593-024-01748-7

Muir, J., Tse, Y. C., Iyer, E. S., Biris, J., Cvetkovska, V., Lopez, J., & Bagot, R. C. (2020). Ventral Hippocampal Afferents to Nucleus Accumbens Encode Both Latent Vulnerability and Stress-Induced Susceptibility. *Biological Psychiatry*, *88*(11), 843–854. https://doi.org/10.1016/j.biopsych.2020.05.021

Nestler, E. J., & Carlezon, W. A. (2006). The Mesolimbic Dopamine Reward Circuit in Depression. *Biological Psychiatry*, *59*(12), 1151–1159. https://doi.org/10.1016/j.biopsych.2005.09.018

Nevin, J. A., & Mandell, C. (2017). Comparing positive and negative reinforcement: A fantasy experiment. *Journal of the Experimental Analysis of Behavior*, *107*(1), 34–38. https://doi.org/10.1002/jeab.237

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Nowend, K. L., Arizzi, M., Carlson, B. B., & Salamone, J. D. (2001). D1 or D2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. *Pharmacology Biochemistry and Behavior*, *69*(3), 373–382. https://doi.org/10.1016/S0091-3057(01)00524-X

Noworyta-Sokolowska, K., Kozub, A., Jablonska, J., Rodriguez Parkitna, J., Drozd, R., & Rygula, R. (2019). Sensitivity to negative and positive feedback as a stable and enduring behavioural trait in rats. *Psychopharmacology*, *236*(8), 2389–2403. https://doi.org/10.1007/s00213-019-05333-w

O'Donnell, P., & Grace, A. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: Hippocampal gating of prefrontal cortical input. *The Journal of Neuroscience*, *15*(5), 3622–3639. https://doi.org/10.1523/JNEUROSCI.15-05-03622.1995

O'Hearn, E., & Molliver, M. E. (1984). Organization of raphe-cortical projections in rat: A quantitative retrograde study. *Brain Research Bulletin*, *13*(6), 709–726. https://doi.org/10.1016/0361-9230(84)90232-6

Onge, J. R. S., Abhari, H., & Floresco, S. B. (2011). Dissociable Contributions by Prefrontal D1 and D2 Receptors to Risk-Based Decision Making. *Journal of Neuroscience*, *31*(23), 8625–8633. https://doi.org/10.1523/JNEUROSCI.1020-11.2011

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., & Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), Article 7643. https://doi.org/10.1038/nature21376

Otis, J. M., Zhu, M., Namboodiri, V. M. K., Cook, C. A., Kosyk, O., Matan, A. M., Ying, R., Hashikawa, Y., Hashikawa, K., Trujillo-Pisanty, I., Guo, J., Ung, R. L., Rodriguez-Romaguera, J., Anton, E. S., & Stuber, G. D. (2019). Paraventricular

Thalamus Projection Neurons Integrate Cortical and Hypothalamic Signals for Cue-Reward Processing. *Neuron*, *103*(3), 423-431.e4. https://doi.org/10.1016/j.neuron.2019.05.018

Pan, W.-X., Brown, J., & Dudman, J. T. (2013). Neural signals of extinction in the inhibitory microcircuit of the ventral midbrain. *Nature Neuroscience*, *16*(1), 71–78. https://doi.org/10.1038/nn.3283

Pardo-Garcia, T. R., Garcia-Keller, C., Penaloza, T., Richie, C. T., Pickel, J., Hope, B. T., Harvey, B. K., Kalivas, P. W., & Heinsbroek, J. A. (2019). Ventral Pallidum Is the Primary Target for Accumbens D1 Projections Driving Cocaine Seeking. *The Journal of Neuroscience*, *39*(11), 2041–2051. https://doi.org/10.1523/JNEUROSCI.2822-18.2018

Parker, N. F., Baidya, A., Cox, J., Haetzel, L. M., Zhukovskaya, A., Murugan, M., Engelhard, B., Goldman, M. S., & Witten, I. B. (2022). Choice-selective sequences dominate in cortical relative to thalamic inputs to NAc to support reinforcement learning. *Cell Reports*, *39*(7). https://doi.org/10.1016/j.celrep.2022.110756

Parsons, M. P., Li, S., & Kirouac, G. J. (2007). Functional and anatomical connection between the paraventricular nucleus of the thalamus and dopamine fibers of the nucleus accumbens. *The Journal of Comparative Neurology*, *500*(6), 1050–1063. https://doi.org/10.1002/cne.21224

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: Affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476–483. https://doi.org/10.1016/j.tics.2012.07.009

Piantadosi, P. T., Yeates, D. C. M., & Floresco, S. B. (2020). Prefrontal cortical and nucleus accumbens contributions to discriminative conditioned suppression of reward-seeking. *Learning & Memory*, *27*(10), 429–440. https://doi.org/10.1101/lm.051912.120

Qi, J., Zhang, S., Wang, H.-L., Barker, D. J., Miranda-Barrientos, J., & Morales, M. (2016). VTA glutamatergic inputs to nucleus accumbens drive aversion by acting on GABAergic interneurons. *Nature Neuroscience*, *19*(5), 725–733. https://doi.org/10.1038/nn.4281

Rapanelli, M., Wang, W., Hurley, E., Feltri, M. L., Pittenger, C., Frick, L. R., & Yan, Z. (2023). Cholinergic neurons in the basal forebrain are involved in behavioral abnormalities associated with Cul3 deficiency: Role of prefrontal cortex projections in cognitive deficits. *Translational Psychiatry*, *13*(1), 1–11. https://doi.org/10.1038/s41398-023-02306-8

Reed, S. J., Lafferty, C. K., Mendoza, J. A., Yang, A. K., Davidson, T. J., Grosenick, L., Deisseroth, K., & Britt, J. P. (2018). Coordinated Reductions in Excitatory Input to the Nucleus Accumbens Underlie Food Consumption. *Neuron*, *99*(6), 1260-1273.e4. https://doi.org/10.1016/j.neuron.2018.07.051

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory: Vol. Vol. 2*.

Richard, J. M., & Berridge, Kent. C. (2013). Prefrontal cortex modulates desire and dread generated by nucleus accumbens glutamate disruption. *Biological Psychiatry*, *73*(4), 360–370. https://doi.org/10.1016/j.biopsych.2012.08.009

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *The Journal of Neuroscience*, *29*(42), 13365–13376. https://doi.org/10.1523/JNEUROSCI.2572-09.2009

Sackett, D. A., Saddoris, M. P., & Carelli, R. M. (2017). Nucleus Accumbens Shell Dopamine Preferentially Tracks Information Related to Outcome Value of Reward. *eNeuro*, *4*(3). https://doi.org/10.1523/ENEURO.0058-17.2017

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015a). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. *The Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015b). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. *Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Salamone, J. D., Correa, M., Yang, J.-H., Rotolo, R., & Presby, R. (2018). Dopamine, Effort-Based Choice, and Behavioral Economics: Basic and Translational Research. *Frontiers in Behavioral Neuroscience*, *12*, 52. https://doi.org/10.3389/fnbeh.2018.00052

Santana, N., & Artigas, F. (2017). Laminar and Cellular Distribution of Monoamine Receptors in Rat Medial Prefrontal Cortex. *Frontiers in Neuroanatomy*, *11*, 87. https://doi.org/10.3389/fnana.2017.00087

Sayegh, F. J. P., Mouledous, L., Macri, C., Pi Macedo, J., Lejards, C., Rampon, C., Verret, L., & Dahan, L. (2024). Ventral tegmental area dopamine projections to the hippocampus trigger long-term potentiation and contextual learning. *Nature Communications*, *15*(1), 4100. https://doi.org/10.1038/s41467-024-47481-4

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Scott, A., Palmer, D., Newell, B., Lin, I., Cayton, C. A., Paulson, A., Remde, P., & Richard, J. M. (2023). Ventral Pallidal GABAergic Neuron Calcium Activity Encodes Cue-Driven Reward Seeking and Persists in the Absence of Reward Delivery. *Journal of Neuroscience*, *43*(28), 5191–5203. https://doi.org/10.1523/JNEUROSCI.0013-23.2023

Segal, M. (1990). Serotonin attenuates a slow inhibitory postsynaptic potential in rat hippocampal neurons. *Neuroscience*, *36*(3), 631–641. https://doi.org/10.1016/0306-4522(90)90006-P

Smith, R. J., Lobo, M. K., Spencer, S., & Kalivas, P. W. (2013). Cocaine-induced adaptations in D1 and D2 accumbens projection neurons (a dichotomy not

necessarily synonymous with direct and indirect pathways). *Current Opinion in Neurobiology*, *23*(4), 546–552. https://doi.org/10.1016/j.conb.2013.01.026

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., Sousa, N., & Rodrigues, A. J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. *Nature Communications*, *7*(1), 11829. https://doi.org/10.1038/ncomms11829

Soares-Cunha, C., Coimbra, B., Sousa, N., & Rodrigues, A. J. (2016). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neuroscience & Biobehavioral Reviews*, *68*, 370–386. https://doi.org/10.1016/j.neubiorev.2016.05.021

Soares-Cunha, C., de Vasconcelos, N. A. P., Coimbra, B., Domingues, A. V., Silva, J. M., Loureiro-Campos, E., Gaspar, R., Sotiropoulos, I., Sousa, N., & Rodrigues, A. J. (2020). Nucleus accumbens medium spiny neurons subtypes signal both reward and aversion. *Molecular Psychiatry*, *25*(12), 3241–3255. https://doi.org/10.1038/s41380-019-0484-3

Soares-Cunha, C., & Heinsbroek, J. A. (2023). Ventral pallidal regulation of motivated behaviors and reinforcement. *Frontiers in Neural Circuits*, *17*, 1086053. https://doi.org/10.3389/fncir.2023.1086053

Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., & Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell*, *184*(10), 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

St-Amand, D., Sheldon, S., & Otto, A. R. (2018). Modulating Episodic Memory Alters Risk Preference during Decision-making. *Journal of Cognitive Neuroscience*, *30*(10), 1433–1441. https://doi.org/10.1162/jocn_a_01253

Stern, E. A., Kincaid, A. E., & Wilson, C. J. (1997). Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons in vivo. *Journal of Neurophysiology*, *77*(4), 1697–1715. https://doi.org/10.1152/jn.1997.77.4.1697

Stopper, C. M., & Floresco, S. B. (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(1), 97–112. https://doi.org/10.3758/s13415-010-0015-9

Stuber, G. D., Sparta, D. R., Stamatakis, A. M., van Leeuwen, W. A., Hardjoprajitno, J. E., Cho, S., Tye, K. M., Kempadoo, K. A., Zhang, F., Deisseroth, K., & Bonci, A. (2011). Amygdala to nucleus accumbens excitatory transmission facilitates reward seeking. *Nature*, *475*(7356), 377–380. https://doi.org/10.1038/nature10194

Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making. *Neuron*, *66*(3), 449–460. https://doi.org/10.1016/j.neuron.2010.03.033

Syed, E. C. J., Grima, L. L., Magill, P. J., Bogacz, R., Brown, P., & Walton, M. E. (2016). Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nature Neuroscience*, *19*(1), Article 1. https://doi.org/10.1038/nn.4187

Szczypka, M. S., Kwok, K., Brot, M. D., Marck, B. T., Matsumoto, A. M., Donahue, B. A., & Palmiter, R. D. (2001). Dopamine Production in the Caudate Putamen Restores Feeding in Dopamine-Deficient Mice. *Neuron*, *30*(3), 819–828. https://doi.org/10.1016/S0896-6273(01)00319-1

Tecuapetla, F., Patel, J. C., Xenias, H., English, D., Tadros, I., Shah, F., Berlin, J., Deisseroth, K., Rice, M. E., Tepper, J. M., & Koos, T. (2010). Glutamatergic Signaling by Mesolimbic Dopamine Neurons in the Nucleus Accumbens. *The Journal of Neuroscience*, *30*(20), 7105–7110. https://doi.org/10.1523/JNEUROSCI.0265-10.2010

Thibeault, K. C., Gunes Kutlu, M., Sanders, C., & Calipari, E. S. (2019). Cell-type and projection-specific dopaminergic encoding of aversive stimuli in addiction. *Brain Research*, *1713*, 1–15. https://doi.org/10.1016/j.brainres.2018.12.024

Thierry, A. M., Gioanni, Y., Dégénétais, E., & Glowinski, J. (2000). Hippocampo-prefrontal cortex pathway: Anatomical and electrophysiological characteristics. *Hippocampus*, *10*(4), 411–419. https://doi.org/10.1002/1098-1063(2000)10:4<411::AID-HIPO7>3.0.CO;2-A

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*, 212–222. https://doi.org/10.2307/1415413

Titulaer, J., Björkholm, C., Feltmann, K., Malmlöf, T., Mishra, D., Bengtsson Gonzales, C., Schilström, B., & Konradsson-Geuken, Å. (2021). The Importance of Ventral

Hippocampal Dopamine and Norepinephrine in Recognition Memory. *Frontiers in Behavioral Neuroscience*, *15*, 667244. https://doi.org/10.3389/fnbeh.2021.667244

Tremblay, R., Lee, S., & Rudy, B. (2016). GABAergic interneurons in the neocortex: From cellular properties to circuits. *Neuron*, *91*(2), 260–292. https://doi.org/10.1016/j.neuron.2016.06.033

Tzschentke, T. M., & Schmidt, W. J. (2000). Functional relationship among medial prefrontal cortex, nucleus accumbens, and ventral tegmental area in locomotion and reward. *Critical Reviews in Neurobiology*, *14*(2), 131–142.

van Holstein, M., MacLeod, P. E., & Floresco, S. B. (2020). Basolateral amygdala—Nucleus accumbens circuitry regulates optimal cue-guided risk/reward decision making. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *98*, 109830. https://doi.org/10.1016/j.pnpbp.2019.109830

Verharen, J. P. H., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2019). Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology*, *44*(13), 2195–2204. https://doi.org/10.1038/s41386-019-0454-0

Vierock, J., Rodriguez-Rozada, S., Dieter, A., Pieper, F., Sims, R., Tenedini, F., Bergs, A. C. F., Bendifallah, I., Zhou, F., Zeitzschel, N., Ahlbeck, J., Augustin, S., Sauter, K., Papagiakoumou, E., Gottschalk, A., Soba, P., Emiliani, V., Engel, A. K., Hegemann, P., & Wiegert, J. S. (2021). BiPOLES is an optogenetic tool developed for bidirectional dual-color control of neurons. *Nature Communications*, *12*(1), 4527. https://doi.org/10.1038/s41467-021-24759-5

Voorn, P., Vanderschuren, L. J. M. J., Groenewegen, H. J., Robbins, T. W., & Pennartz, C. M. A. (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends in Neurosciences*, *27*(8), 468–474. https://doi.org/10.1016/j.tins.2004.06.006

Wan, J., Peng, W., Li, X., Qian, T., Song, K., Zeng, J., Deng, F., Hao, S., Feng, J., Zhang, P., Zhang, Y., Zou, J., Pan, S., Shin, M., Venton, B. J., Zhu, J. J., Jing, M., Xu, M., & Li, Y. (2021). A genetically encoded sensor for measuring serotonin dynamics. *Nature Neuroscience*, *24*(5), 746–752. https://doi.org/10.1038/s41593-021-00823-7

Wang, A. Y., Miura, K., & Uchida, N. (2013). The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nature Neuroscience*, *16*(5), 639–647. https://doi.org/10.1038/nn.3377

Warlow, S. M., Singhal, S. M., Hollon, N. G., Faget, L., Dowlat, D. S., Zell, V., Hunker, A. C., Zweifel, L. S., & Hnasko, T. S. (2024). Mesoaccumbal glutamate neurons drive reward via glutamate release but aversion via dopamine co-release. *Neuron*, *112*(3), 488-499.e5. https://doi.org/10.1016/j.neuron.2023.11.002

Wassum, K. M., Ostlund, S. B., & Maidment, N. T. (2012). Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task. *Biological Psychiatry*, *71*(10), 846–854. https://doi.org/10.1016/j.biopsych.2011.12.019

Wenzel, J. M., Zlebnik, N. E., Patton, M. H., Smethells, J. R., Ayvazian, V. M., Dantrassy, H. M., Zhang, L.-Y., Mathur, B. N., & Cheer, J. F. (2023). Selective chemogenetic inactivation of corticoaccumbal projections disrupts trait choice

impulsivity. *Neuropsychopharmacology*, *48*(12), Article 12. https://doi.org/10.1038/s41386-023-01604-5

West, E. A., & Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after Reinforcer Devaluation. *The Journal of Neuroscience*, *36*(4), 1128–1139. https://doi.org/10.1523/JNEUROSCI.2976-15.2016

White, N. M. (1989). Reward or reinforcement: What's the difference? *Neuroscience & Biobehavioral Reviews*, *13*(2), 181–186. https://doi.org/10.1016/S0149-7634(89)80028-4

Willmore, L., Cameron, C., Yang, J., Witten, I. B., & Falkner, A. L. (2022). Behavioural and dopaminergic signatures of resilience. *Nature*, *611*(7934), 124–132. https://doi.org/10.1038/s41586-022-05328-2

Wilson, C. (2008). Up and down states. *Scholarpedia Journal*, *3*(6), 1410.

Wilson, C. J., & Groves, P. M. (1981). Spontaneous firing patterns of identified spiny neurons in the rat neostriatum. *Brain Research*, *220*(1), 67–80. https://doi.org/10.1016/0006-8993(81)90211-0

Wilson, C., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *The Journal of Neuroscience*, *16*(7), 2397–2410. https://doi.org/10.1523/JNEUROSCI.16-07-02397.1996

Wu, J.-L., Li, Z.-M., Chen, H., Chen, W.-J., Hu, N.-Y., Jin, S.-Y., Li, X.-W., Chen, Y.-H., Yang, J.-M., & Gao, T.-M. (2024). Distinct septo-hippocampal cholinergic projections separately mediate stress-induced emotional and cognitive deficits. *Science Advances*, *10*(45), eado1508. https://doi.org/10.1126/sciadv.ado1508

Yamaguchi, T., Wang, H.-L., Li, X., Ng, T. H., & Morales, M. (2011). Mesocorticolimbic glutamatergic pathway. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(23), 8476–8490. https://doi.org/10.1523/JNEUROSCI.1598-11.2011

Yang, H., Jong, J. W. de, Tak, Y., Peck, J., Bateup, H. S., & Lammel, S. (2018). Nucleus Accumbens Subnuclei Regulate Motivated Behavior via Direct Inhibition and Disinhibition of VTA Dopamine Subpopulations. *Neuron*, *97*(2), 434-449.e4. https://doi.org/10.1016/j.neuron.2017.12.022

Yoo, J. H., Zell, V., Gutierrez-Reed, N., Wu, J., Ressler, R., Shenasa, M. A., Johnson, A. B., Fife, K. H., Faget, L., & Hnasko, T. S. (2016). Ventral tegmental area glutamate neurons co-release GABA and promote positive reinforcement. *Nature Communications*, *7*(1), 13697. https://doi.org/10.1038/ncomms13697

Yoshida, K., Drew, M. R., Kono, A., Mimura, M., Takata, N., & Tanaka, K. F. (2021). Chronic social defeat stress impairs goal-directed behavior through dysregulation of ventral hippocampal activity in male mice. *Neuropsychopharmacology*, *46*(9), 1606–1616. https://doi.org/10.1038/s41386-021-00990-y

Yoshida, K., Drew, M. R., Mimura, M., & Tanaka, K. F. (2019). Serotonin-mediated inhibition of ventral hippocampus is required for sustained goal-directed behavior. *Nature Neuroscience*, *22*(5), Article 5. https://doi.org/10.1038/s41593-019-0376-5

Yoshida, K., Tsutsui-Kimura, I., Kono, A., Yamanaka, A., Kobayashi, K., Watanabe, M., Mimura, M., & Tanaka, K. F. (2020). Opposing Ventral Striatal Medium Spiny Neuron Activities Shaped by Striatal Parvalbumin-Expressing Interneurons during

Goal-Directed Behaviors. *Cell Reports*, *31*(13), 107829. https://doi.org/10.1016/j.celrep.2020.107829

Yu, Z., Guindani, M., Grieco, S. F., Chen, L., Holmes, T. C., & Xu, X. (2022). Beyond t test and ANOVA: Applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron*, *110*(1), 21–35. https://doi.org/10.1016/j.neuron.2021.10.030

Zhu, Y., Wienecke, C. F. R., Nachtrab, G., & Chen, X. (2016). A thalamic input to the nucleus accumbens mediates opiate dependence. *Nature*, *530*(7589), 219–222. https://doi.org/10.1038/nature16954

Zhuo, Y., Luo, B., Yi, X., Dong, H., Miao, X., Wan, J., Williams, J. T., Campbell, M. G., Cai, R., Qian, T., Li, F., Weber, S. J., Wang, L., Li, B., Wei, Y., Li, G., Wang, H., Zheng, Y., Zhao, Y., … Li, Y. (2024). Improved green and red GRAB sensors for monitoring dopaminergic activity in vivo. *Nature Methods*, *21*(4), 680–691. https://doi.org/10.1038/s41592-023-02100-w