Uncertainty gates redundancy in reward integration in prefrontal-cortical and ventral-hippocampal nucleus accumbens inputs to tune 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.

Abstract

Circuit neuroscience commonly seeks to assign specific functions to specific circuits. Yet, redundancy can be highly adaptive and is therefore a critical motif in circuit organization. The NAc, a highly integrative brain region controlling motivated behavior, is thought to receive distinct information from its various glutamatergic inputs yet strong evidence of functional specialization of inputs is lacking. 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 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 requires goal-directed action and uncertainty. Ultimately, using optogenetic stimulation we demonstrate that both inputs co-operatively modulate task engagement. Taken together, we illustrate how similar encoding, with differential gating by behavioral state, supports outcome encoding to tune engagement to recent history of reward.

INTRODUCTION

Redundancy is a defining property of nervous system organization (Marder, 2011; Marder & Goaillard, 2006; Mizusaki & O'Donnell, 2021) yet there has been limited consideration of the role of redundancy in neural circuit mechanisms of motivated behavior. Redundancy in neural circuits may confer various advantages, including increasing robustness of cognition and behavior to perturbation, enhancing encoding accuracy, and facilitating coherent integration of multiple inputs, suggesting it should be a frequently observed motif (Ghanbari et al., 2023; Hiratani & Fukai, 2018; Nguyen et al., 2019). The literature abounds with examples of apparent circuit-specific cognitive and behavioral functions yet redundancy is rarely examined. To better understand these opposing processes in nucleus accumbens (NAc) circuits we leveraged dual circuit recordings and computational modeling to rigorously test the specificity and redundancy of information processing in a fully controlled, within animal comparison of two NAc glutamatergic inputs.

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., 2022; 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 (Chen et al., 2023; Reed et al., 2018; Parker et al., 2022).

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 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 outcomes under uncertainty. 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 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, an auditory cue 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 lever 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).

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.

Reward strongly suppressed mPFC-NAc and vHip-NAc activity in female and male mice (Fig. 1H, K). We focused analysis on two epochs: ITI start (0-2 sec after lever press) when outcome is signaled and delivered, and ITI end (8-10 sec after lever press) when trial outcome has been integrated prior to next trial start. At ITI start, reward modestly elevated mPFC-NAc activity, significantly in male mice (Fig. 1I). By ITI end, reward robustly suppressed mPFC-NAc activity in female and male mice (Fig. 1J). At ITI start, reward did not significantly modulate vHip-NAc activity (Fig. 1L). By ITI end, however, reward robustly suppressed vHip-NAc activity in female and male mice (Fig. 1M). This indicates that outcome encoding emerges across the ITI and that reward similarly suppresses mPFC-NAc and vHip-NAc activity.

We observed that mPFC-NAc and vHip-NAc are similarly modulated by reward so we examined if one circuit leads the other. To do this, we calculated the time lag for the maximum cross-correlation and found that the regression coefficients of the mPFC-NAc and vHip-NAc cross-

correlation 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. S1). Rather, mPFC-NAc and vHip-NAc encoding emerge on similar timescales, and both are simultaneously conveying similar choice outcome information to NAc.

mPFC-NAc and vHip-NAc similarly integrate reward history

We find that mPFC-NAc and vHip-NAc similarly encode outcomes. Noting that reward-induced suppression lasted throughout the ITI, we 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, rewarded then unrewarded, unrewarded then rewarded, and unrewarded then unrewarded. 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 and no effect of sex (Fig. S2).We therefore report sex-combined analyses. Both previous and current outcome modulate mPFC-NAc activity (Fig. 2A). At ITI start, in previously rewarded mice, current reward increased mPFC-NAc activity, an effect not seen in previously unrewarded mice. This suggests that initial reward processing is facilitated by prior reward (Fig. 2B). At ITI end, reward on the current trial similarly suppressed mPFC-NAc, regardless of prior outcome. However, when mice were unrewarded on the current trial, suppression of mPFC-NAc by prior reward was maintained to 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. At ITI start, vHip-NAc activity reflected the prior trial outcome (Fig 2D). In mice previously rewarded, current reward increased vHip-NAc activity, an effect not seen in previously unrewarded mice. Similar to mPFC-NAc, this again suggests that prior reward facilitates initial encoding of current reward (Fig. 2E). At ITI end current reward suppressed vHip-NAc regardless of prior outcome, similar to mPFC-NAc. However, in mice unrewarded on the current trial, suppression of vHip-NAc activity by prior reward was maintained to 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 subsequent unrewarded outcomes gradually increase activity, such that, by ITI end, the relative degree of suppression represents an integrated reward outcome history.

mPFC-NAc and vHip-NAc are differentially sensitive to uncertainty

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, and if this varies with trial outcome (Fig. 3A). Reward reduced mPFC-NAc conditional entropy, indicating that vHip-NAc and mPFC-NAc signals are more redundant after reward than non-reward. (Fig. 3B). In contrast, following previous unreward, reward increased vHip-NAc conditional entropy (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 surprising reward.

A major function attributed to the vHip is resolving uncertainty (Iyer et al., 2022; Schumacher et al., 2016, 2018). Therefore, we reasoned that the increased information carried by vHip-NAc in encoding reward that follows unreward might indicate that vHip-NAc preferentially encodes outcomes when reward is less certain. If this is true, across consecutive rewards, vHip-NAc should become insensitive to outcome while mPFC-NAc continues to encode outcome. To test this, we examined current outcome encoding at ITI end while considering prior outcomes up to three trials back. Consistent with this prediction, mPFC-NAc encoded current outcome regardless of prior outcome history (Fig. 3D; Supplementary Table 29), however vHip-NAc failed to encode current outcome after two or more consecutive rewards (Fig. 3E; Supplementary Table 29). This supports our hypothesis that vHip-NAc is specifically engaged by uncertainty whereas mPFC-NAc invariantly encodes outcome.

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 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 with subtle modification (Fig. 4B). At ITI start, mPFC-NAc activity reflected only the prior outcome, differing from the two-lever task, where current outcome was encoded if the prior trial was rewarded (Fig. 4C). This suggests that eliminating choice slows integration of the current outcome. At ITI end, current and prior outcomes were encoded, as in the two-lever task. Examining vHip-NAc in the one-lever task revealed largely similar modulation (Fig. 4E). At ITI start, vHip-NAc only encoded the prior outcome (Fig. 4F) and, at ITI end, prior and current outcomes were integrated similar to the two-lever task (Fig. 4G). Taken together, this demonstrate 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 in either circuit. 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. 4H). Eliminating the response requirement markedly and distinctly altered reward integration in both circuits. In mPFC-NAc (Fig. 4I), encoding at trial start continued to represent prior outcome (Fig. 4J), however, at ITI end, prior outcome was erased and only the current outcome encoded (Fig. 4K). 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.

In contrast, vHip-NAc, failed to encode either prior or current outcome when reward was no longer contingent upon lever press (Fig. 4L-N). Critically, this cannot be explained by modulation of task engagement given that mPFC-NAc continued to represent reward in these same animals (4I-K) and licking bouts were similarly maintained across task variants (FigS3). Rather, this reveals that task demands distinctly shape neural encoding of reward and reward-history in mPFC-NAc and vHip-NAc, with vHip-NAc characterized by a marked behavioral state dependence. 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 disengages and fails to encode passively encountered reward.

mPFC-NAc and vHip-NAc modulate task engagement

Examining neural representation of outcomes identified both mechanistic redundancy and situational 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 may modulate task engagement. To test this, we examined if neural activity at ITI end predicted 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 mPFC-NAc, vHip-NAc, and the interaction of mPFC-NAc and vHip-NAc activity (Fig. 5A, Supplementary Table 30). This suggests that increased activity during outcome integration in either circuit increases latency to lever press, indicating reduced behavioral engagement.

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. S4). To test if mPFC-NAc and vHip-NAc uniquely or redundantly 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 (Fig. 5F). 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. This shows that total glutamatergic input modulates engagement, independent of input identity (Fig. 5G). 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 simultaneous 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 is specifically recruited by action-induced uncertainty. 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. Layered upon this basic functionality, the richness of mPFC-NAc encoding scales with environmental complexity. 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, mPFC-NAc encoding window extends to integrate reward history.

We show that, like the mPFC-NAc, vHip-NAc also integrates reward history, revealing a novel role for this circuit as a parallel and distinct stream of outcome integration. However, in contrast to the state-invariance of mPFC-NAc, we find that vHip-NAc is recruited primarily by action-induced uncertainty such that vHip-NAc only encodes reward (and reward history) when the act of lever pressing elicits an uncertain expectation of reward. vHip-NAc disengages when reward is passively encountered or action-elicited reward is consistently encountered, such that action-induced uncertainty is low. This functional recruitment under uncertainty is reminiscent of the role of vHip in resolving motivational uncertainty and conflict in explicit conflict tasks (Iyer et al., 2022; Schumacher et al., 2016, 2018). Our findings suggest that this function may generalize to other kinds of uncertainty such that vHip-NAc is also sensitive to action-elicited uncertainty in gating neural encoding and behavioral engagement in reward seeking.

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, the 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 unexpected 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. vHip-NAc disengagement under low uncertainty may serve as a fail-safe to ensure continued high behavioral engagement in low complexity and/or high reward behavioral contexts. Conversely, failure to disengage vHip-NAc could inappropriately suppress behavioral engagement with reward, manifesting 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 which

associate with aberrant reward processing. Overall, our findings point to the importance of balanced suppression of NAc glutamatergic inputs during outcome integration, and reward-modulated behavioral engagement.

Here we examined redundancy and specificity in a key neural circuit 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, only action-induced uncertainty elicits vHip-NAc encoding. Overall, our findings illustrate an interplay of redundancy and specificity in circuit control of motivated behavior and demonstrate the need to contextualize events within varied behavioral states to fully understand neural encoding.

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 # 1044pGP-AAV-syn-jGCaMP7f-WPRE was a gift from Douglas Kim & GENIE Project (Addgene plasmid # 104488; http://n2t.net/addgene:104488 ; RRID:Addgene_104488).88; http://n2t.net/addgene:104488 ; RRID:Addgene_104488). 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: -0.3, M/L: +/-1.90, 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: -0.3, M/L: +/-1.90, 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.

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). 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 $\Delta F/F$ 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.

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 will be made available at github.com/BagotLab/IyerEtAll2024

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.

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

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

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.

Barrientos, C., Knowland, D., Wu, M. M. J., Lilascharoen, V., Huang, K. W., Malenka, R. C., & Lim, B. K. (2018). Cocaine-Induced Structural Plasticity in Input Regions to Distinct Cell Types in Nucleus Accumbens. *Biological Psychiatry*, *84*(12), 893–904. https://doi.org/10.1016/j.biopsych.2018.04.019

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

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

Cahill, M. E., Bagot, R. C., Gancarz, A. M., Walker, D. M., Sun, H., Wang, Z.-J., Heller, E. A., Feng, J., Kennedy, P. J., Koo, J. W., Cates, H. M., Neve, R. L., Shen, L., Dietz, D. M., & Nestler, E. J. (2016). Bidirectional Synaptic Structural Plasticity after Chronic Cocaine Administration Occurs through Rap1 Small GTPase Signaling. *Neuron*, *89*(3), 566–582. https://doi.org/10.1016/j.neuron.2016.01.031

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

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

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

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

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

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.

Ghanbari, M., Li, G., Hsu, L., & Yap, P. (2023). Accumulation of network redundancy marks the early stage of Alzheimer's disease. *Human Brain Mapping*, *44*(8), 2993–3006. https://doi.org/10.1002/hbm.26257

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

Griffin, W. C., Lopez, M. F., Woodward, J. J., & Becker, H. C. (2023). Alcohol dependence and the ventral hippocampal influence on alcohol drinking in male mice. *Alcohol*, *106*, 44–54. https://doi.org/10.1016/j.alcohol.2022.10.004

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

Hiratani, N., & Fukai, T. (2018). Redundancy in synaptic connections enables neurons to learn optimally. *Proceedings of the National Academy of Sciences*, *115*(29), E6871–E6879. https://doi.org/10.1073/pnas.1803274115

Iyer, E. S., Weinberg, A., & Bagot, R. C. (2022). Ambiguity and conflict: Dissecting uncertainty in decision-making. *Behavioral Neuroscience*, *136*(1), 1–12. https://doi.org/10.1037/bne0000489

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

Kircher, D. M., Aziz, H., Mangieri, R. A., & Morrisett, R. A. (2019). Ethanol Experience Enhances Glutamatergic Ventral Hippocampal Inputs To D1 Receptor-Expressing Medium Spiny Neurons In The Nucleus Accumbens Shell. *The Journal of Neuroscience*, 3051–18. https://doi.org/10.1523/JNEUROSCI.3051-18.2019

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

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

Marder, E. (2011). Variability, compensation, and modulation in neurons and circuits. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(Suppl 3), 15542–15548. https://doi.org/10.1073/pnas.1010674108

Marder, E., & Goaillard, J.-M. (2006). Variability, compensation and homeostasis in neuron and network function. *Nature Reviews Neuroscience*, *7*(7), 563–574. https://doi.org/10.1038/nrn1949

Millan, E. Z., Kim, H. A., & Janak, P. H. (2017). Optogenetic activation of amygdala projections to nucleus accumbens can arrest conditioned and unconditioned alcohol consummatory behavior. *Neuroscience*, *360*, 106–117. https://doi.org/10.1016/j.neuroscience.2017.07.044

Mizusaki, B. E. P., & O'Donnell, C. (2021). Neural circuit function redundancy in brain disorders. *Current Opinion in Neurobiology*, *70*, 74–80. https://doi.org/10.1016/j.conb.2021.07.008

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., 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

Muir, J., Tse, Y. C., Iyer, E., Sorensen, J., Eid, R., Cvetkovska, V., Wassef, K., Gostlin, S., Spencer, N. J., & Bagot, R. C. (2022). *Sexually dimorphic neural encoding of threat discrimination in nucleus accumbens afferents drives suppression of reward behavior* (p. 2022.10.04.510865). bioRxiv. https://doi.org/10.1101/2022.10.04.510865

Nguyen, A. T., Xu, J., Luu, D. K., Zhao, Q., & Yang, Z. (2019). Advancing System Performance with Redundancy: From Biological to Artificial Designs. *Neural Computation*, *31*(3), 555–573. https://doi.org/10.1162/neco_a_01166

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

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

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

Pascoli, V., Terrier, J., Espallergues, J., Valjent, E., O'Connor, E. C., & Lüscher, C. (2014). Contrasting forms of cocaine-evoked plasticity control components of relapse. *Nature*, *509*(7501), Article 7501. https://doi.org/10.1038/nature13257

Pignatelli, M., Tejeda, H. A., Barker, D. J., Bontempi, L., Wu, J., Lopez, A., Palma Ribeiro, S., Lucantonio, F., Parise, E. M., Torres-Berrio, A., Alvarez-Bagnarol, Y., Marino, R. A. M., Cai, Z.-L., Xue, M., Morales, M., Tamminga, C. A., Nestler, E. J., & Bonci, A. (2021). Cooperative synaptic and intrinsic plasticity in a disynaptic limbic circuit drive stress-induced anhedonia and

passive coping in mice. *Molecular Psychiatry*, *26*(6), Article 6. https://doi.org/10.1038/s41380-020-0686-8

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

Schumacher, A., Villaruel, F. R., Ussling, A., Riaz, S., Lee, A. C. H., & Ito, R. (2018). Ventral Hippocampal CA1 and CA3 Differentially Mediate Learned Approach-Avoidance Conflict Processing. *Current Biology: CB*, *28*(8), 1318-1324.e4. https://doi.org/10.1016/j.cub.2018.03.012

Schumacher, A., Vlassov, E., & Ito, R. (2016). The ventral hippocampus, but not the dorsal hippocampus is critical for learned approach-avoidance decision making. *Hippocampus*, *26*(4), 530–542. https://doi.org/10.1002/hipo.22542

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

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

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

Williams, E. S., Manning, C. E., Eagle, A. L., Swift-Gallant, A., Duque-Wilckens, N., Chinnusamy, S., Moeser, A., Jordan, C., Leinninger, G., & Robison, A. J. (2020). Androgen-Dependent Excitability of Mouse Ventral Hippocampal Afferents to Nucleus Accumbens Underlies Sex-Specific Susceptibility to Stress. *Biological Psychiatry*, *87*(6), 492–501. https://doi.org/10.1016/j.biopsych.2019.08.006

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

Zinsmaier, A. K., Dong, Y., & Huang, Y. H. (2022). Cocaine-induced projection-specific and cell type-specific adaptations in the nucleus accumbens. *Molecular Psychiatry*, *27*(1), Article 1. https://doi.org/10.1038/s41380-021-01112-2

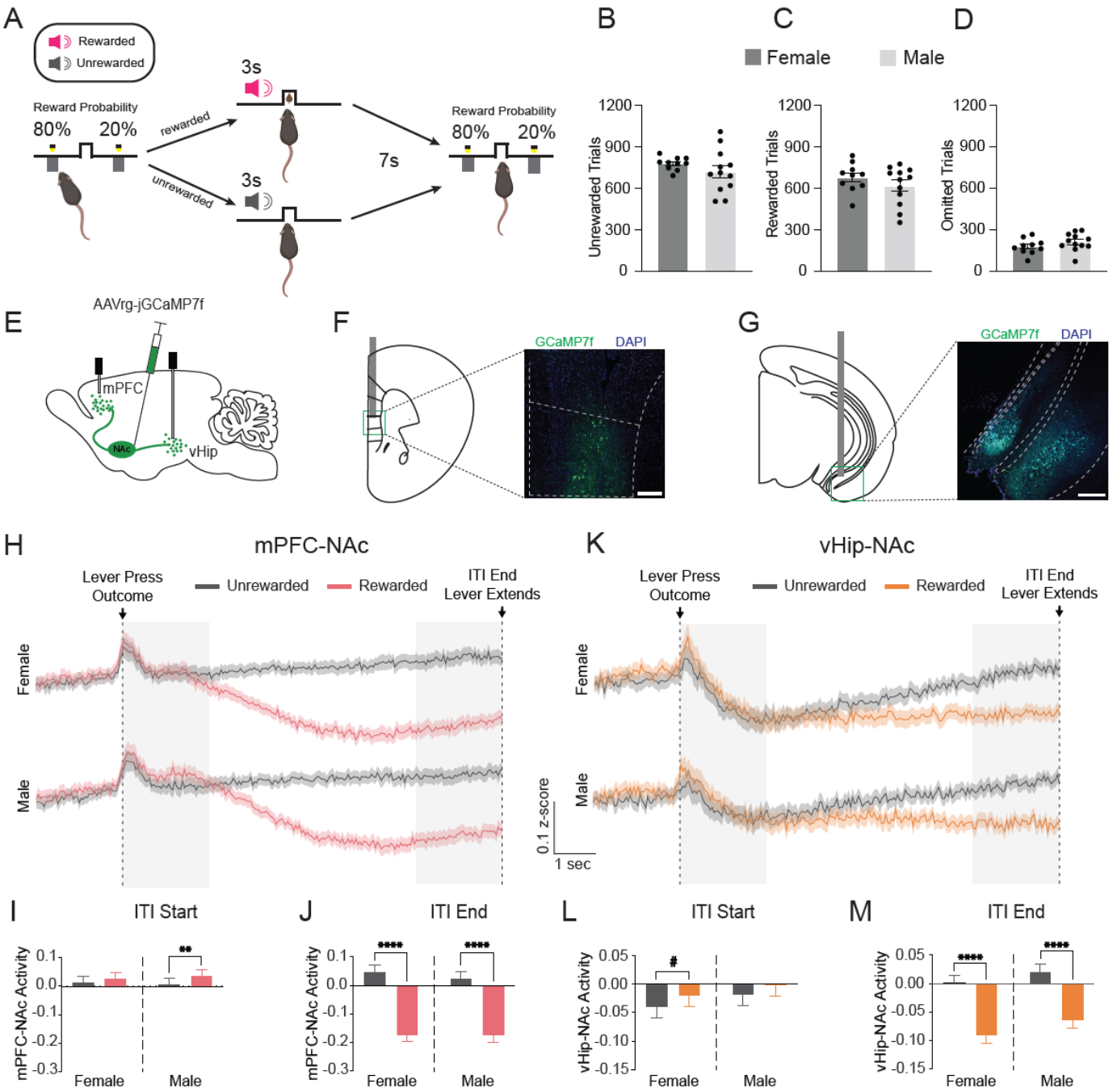

Figure 1. mPFC-NAc and vHip-NAc similarly encode reward in a probabilistically rewarded environment. (A) Schematic of two-armed bandit task. Mice lever press in a two lever task in which one lever is rewarded with chocolate milk on 80% of trials, and the other on 20%. Following a lever press, levers retract, and an auditory cue signals delivery of 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) and optic fibers implanted in medial prefrontal cortex (mPFC) and ventral hippocampus (vHip) to simultaneously probe neural activity indicated by Ca2+-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 male (n=12) mice. Analysis focused on 0-2 sec after lever press (ITI start) or 8-10 sec after lever press (ITI end). (I) At ITI start, mPFC-NAc activity is significantly reduced to unrewarded outcomes in male mice (Z=-2.965, p=0.0060). (J) At ITI end, mPFC-NAc activity is significantly reduced to rewarded outcomes in female (Z=21.348, p<0.0001) and male (Z=19.625, p<0.0001) mice. Estimated mean vHip-NAc activity across all rewarded and unrewarded trials in (K) female and male mice. (L) At ITI start, vHip-NAc activity trends reduction to unrewarded outcomes in female mice (Z=-2.0308, p=0.0815). (M) At ITI end, vHip-NAc activity is significantly reduced to rewarded outcomes in female (Z=8.161; p<0.0001) and male (Z=8.924; p<0.0001) mice. #p<0.1, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

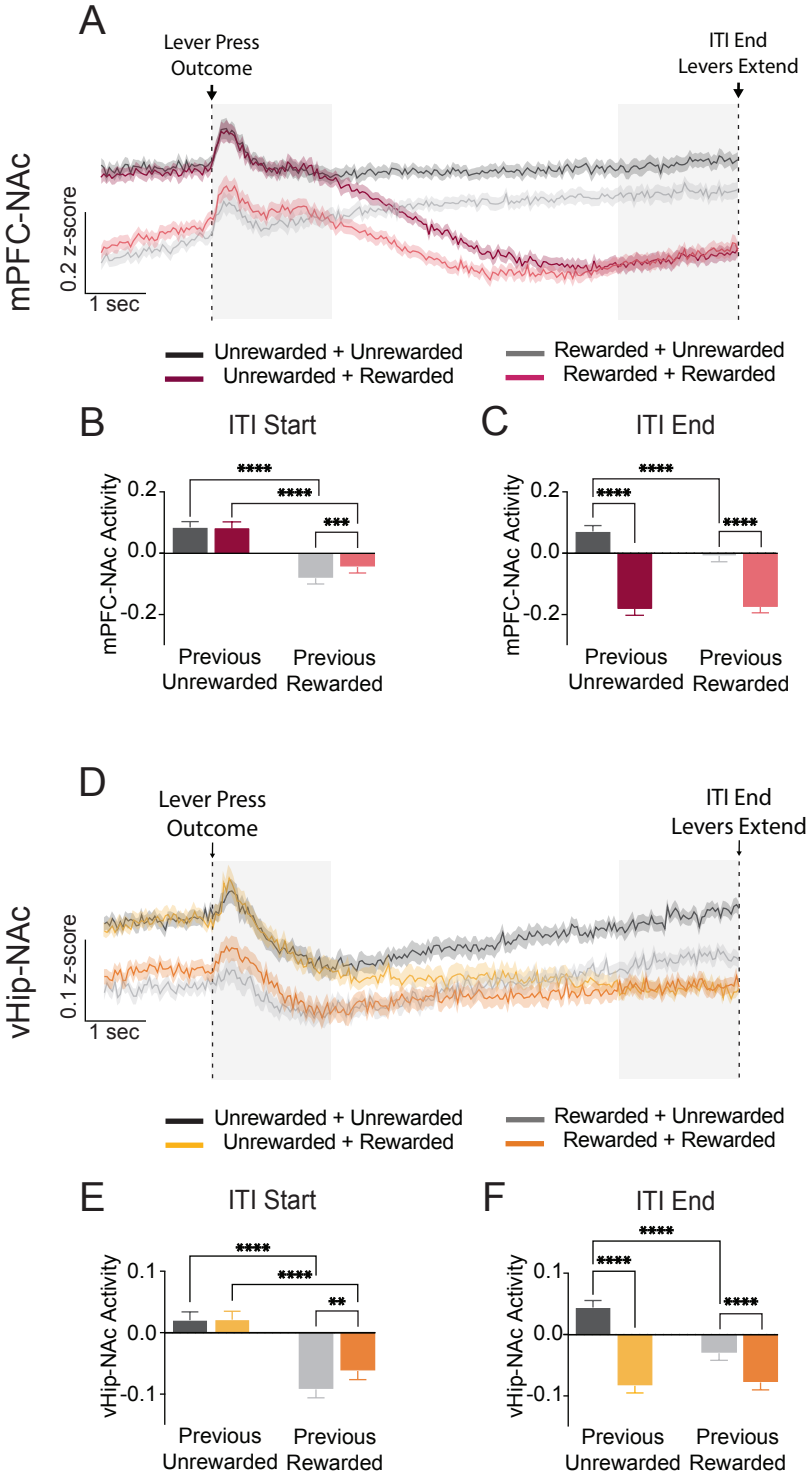

Figure 2. mPFC-NAc and vHip-NAc similarly integrate reward history. (A) Estimated mean mPFC-NAc activity across rewarded+rewarded, rewarded+unrewarded, unrewarded+rewarded, and unrewarded+unrewarded trials in female (n=10) and male (n=12) mice. Analysis focused on 0-2 sec after lever press (ITI start) or 8-10 sec after lever press (ITI end). (B) At ITI start, mPFC-NAc activity is significantly reduced to previous reward (unrewarded +unrewarded vs rewarded+unrewarded: Z=16.9726, p<0.0001; unrewarded+rewarded vs rewarded+rewarded: Z=12.4144, p<0.0001) and significantly more reduced to rewarded+unrewarded than rewarded+rewarded trials (Z=-3.2454, p=0.0023). (C) At ITI end, mPFC-NAc activity is significantly reduced to current reward (unrewarded+ unrewarded vs unrewarded+rewarded: Z=16.0672, p<0.0001; rewarded+ unrewarded vs rewarded+rewarded: Z=25.5127, p<0.0001) and is significantly more reduced on rewarded+unrewarded trials as compared to unrewarded +unrewarded trials (Z= 7.8148, p<0.0001). (D) Estimated mean vHip-NAc activity across rewarded+rewarded, rewarded+unrewarded, unrewarded+rewarded, and unrewarded+unrewarded trials. (E) At ITI start, vHip-NAc activity is significantly reduced to previous reward (unrewarded+unrewarded vs rewarded+unrewarded: Z=11.3306, p<0.0001; unrewarded+rewarded vs rewarded+rewarded: Z=8.1993, p<0.0001) and significantly more reduced to rewarded+unrewarded than rewarded+rewarded trials (Z=-2.6938, p=0.0141). (F) At ITI end, vHip-NAc activity is significantly reduced to current reward (unrewarded+ unrewarded vs unrewarded+rewarded: Z=12.6554, p<0.0001; rewarded+ unrewarded vs rewarded+rewarded: Z=4.6236, p<0.0001) and is significantly more reduced on rewarded +unrewarded trials as compared to unrewarded+unrewarded trials (Z= 7.4329, p<0.0001). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

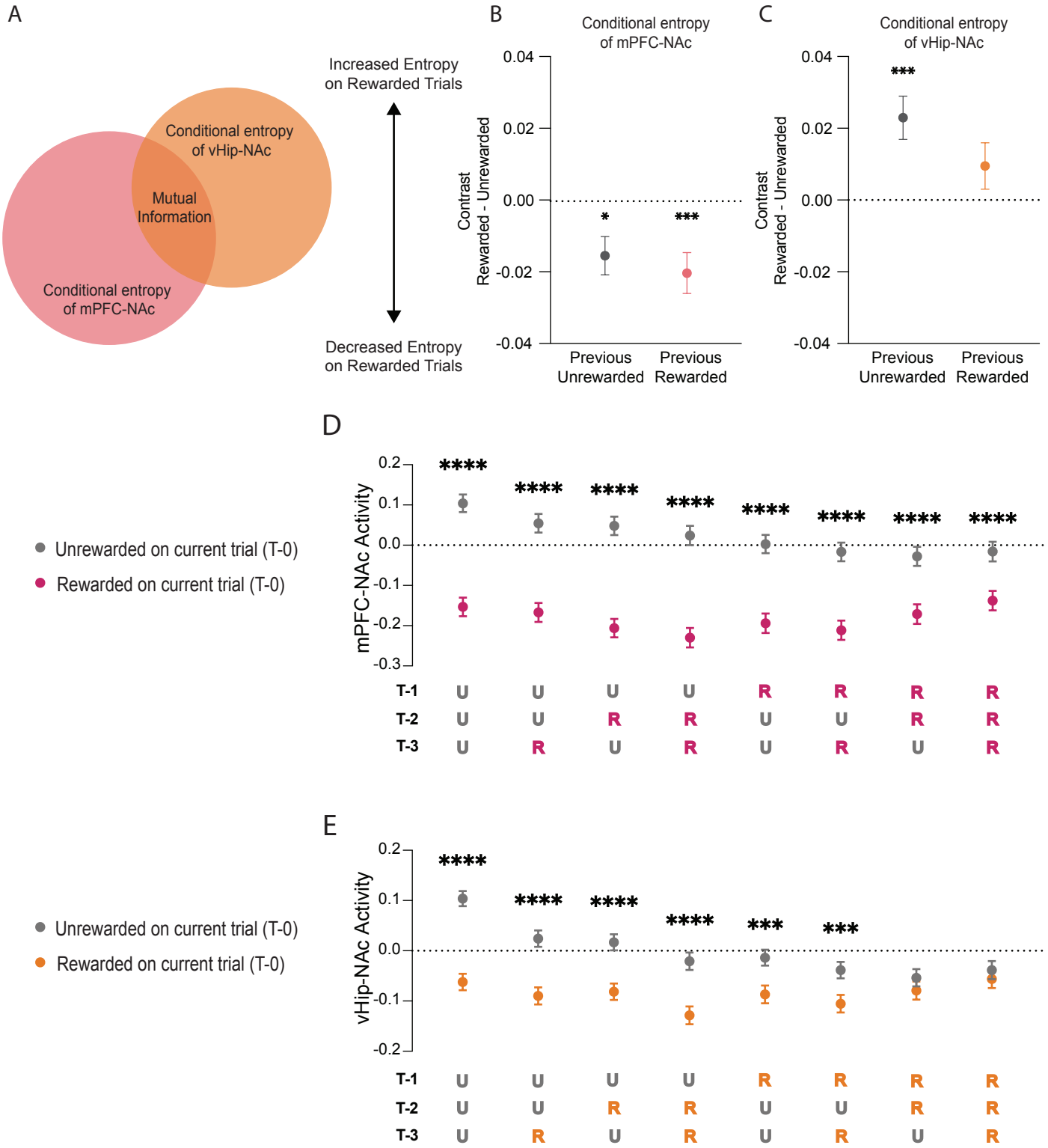

Figure 3. vHip-NAc encodes outcome under uncertainty. (A) Venn diagram representing the relationship between the mutual information and conditional entropy that exists between observed mPFC-NAc and vHip-NAc signals. Conditional entropy is a measure of the additional unique information contributed by 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 (unrewarded+unrewarded vs unrewarded+rewarded: Z=2.8644, p<0.0001; rewarded +unrewarded vs rewarded+rewarded: 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 (unrewarded+ unrewarded vs unrewarded+rewarded: Z=-3.7566, p=0.0003) indicating that more unique information is carried in vHip-NAc specifically 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 significantly reduced on every currently rewarded trial indicating that mPFC-NAc consistently encodes current outcome via relative suppression. In contrast, (E) vHip-NAc activity is significantly reduced on currently rewarded trials but not 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 29 for all comparisons. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

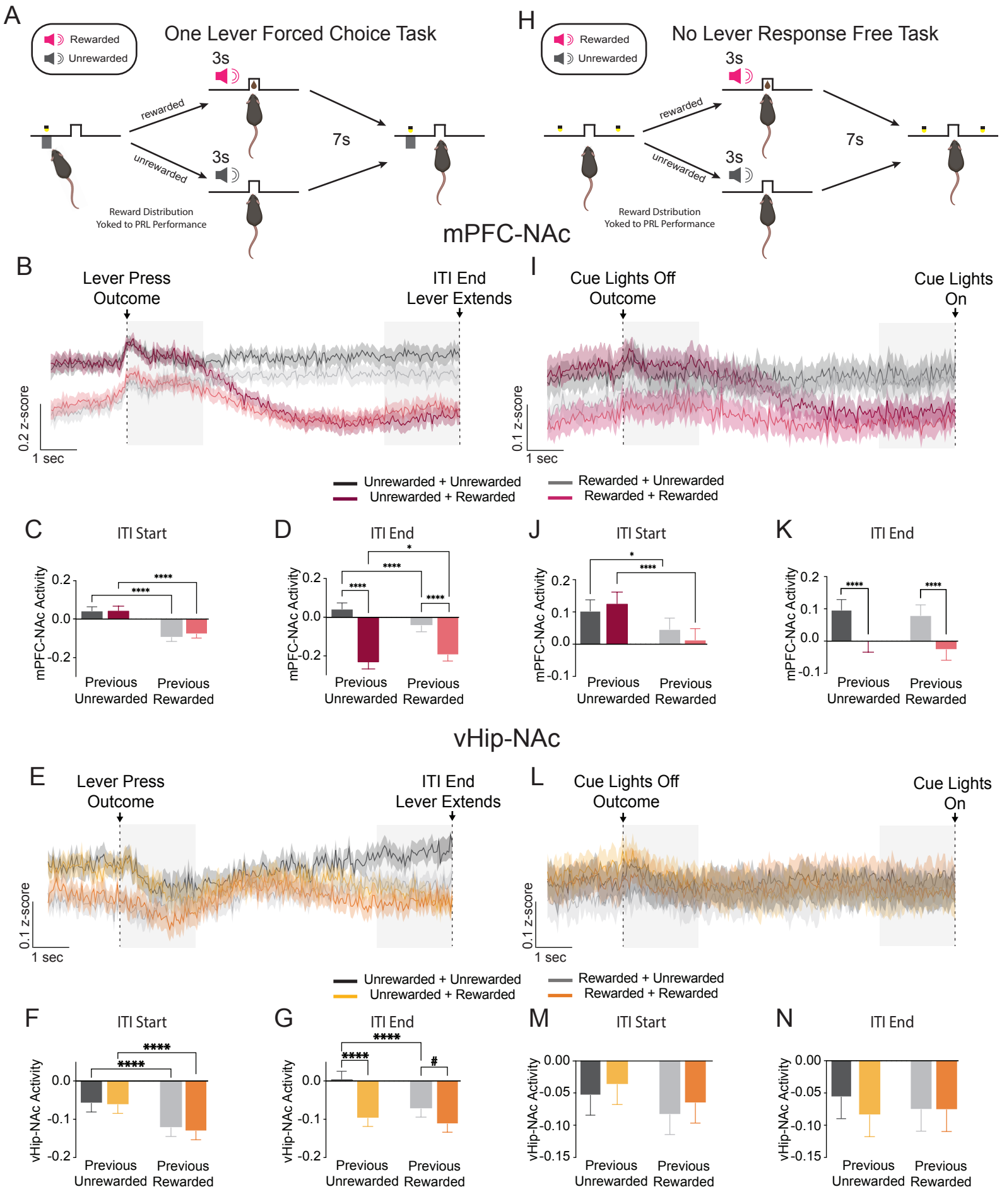

Figure 4. Degrading task requirements reveals circuit specialization in integrating reward history. (A) Schematic of one-lever forced choice task in which lever presses are rewarded on a schedule yoked to each animal's individual performance in the final three days of the two-armed bandit task. Following a lever press, levers retract, and an auditory cue signals delivery of outcome and start of a 10 sec inter-trial interval (ITI). (B) Estimated mean mPFC-NAc activity across rewarded+rewarded, rewarded+unrewarded, unrewarded-+rewarded, and unrewarded+unrewarded trials (n= 14, female n=6). Analysis focused on 0-2 sec after lever press (ITI start) or 8-10 sec after lever press (ITI end). (C) At ITI start, mPFC-NAc activity is significantly reduced to previous reward (unrewarded+unrewarded vs rewarded +unrewarded: Z=7.5901, p<0.0001; unrewarded+rewarded vs rewarded+rewarded: Z=6.9841, p<0.0001). (D) At ITI end, mPFC-NAc activity is significantly

reduced to reward (unrewarded+unrewarded vs unrewarded+rewarded: Z=15.3967, p<0.0001; rewarded+unrewarded vs rewarded+rewarded: Z=8.5014, p<0.0001) and is also modulated by previous reward (unrewarded+unrewarded vs rewarded +unrewarded: Z=4.8267, p<0.0001; unrewarded+rewarded vs rewarded+rewarded: Z=-2.3274, p=0.0395). (E) Estimated mean vHip-NAc activity across rewarded+rewarded, rewarded+unrewarded, unrewarded-+rewarded, and unrewarded +unrewarded trials. (F) At ITI start, vHip-NAc activity is significantly reduced to previous reward (unrewarded+unrewarded vs rewarded +unrewarded: Z=3.5249, p=0.0008; unrewarded+rewarded vs rewarded+rewarded: Z=3.8322, p=0.0003). (G) At ITI end, vHip-NAc activity is significantly reduced to reward (main effect of previous reward: Z=3.2816, P=0.0010; unrewarded +unrewarded vs unrewarded+rewarded: Z=5.0415, p<0.0001; rewarded+unrewarded vs rewarded+rewarded: Z=2.1387, p=0.0623) and to a greater extent for rewarded+unrewarded than unrewarded+unrewarded trials (Z=3.7679, p=0.0003). (H) 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) in the final three days of the two-armed bandit task. Trial structure is signaled by cue-light illumination and after a predetermined delay an auditory cue signals delivery of outcome and start of a 10 sec ITI. (I) Estimated mean mPFC-NAc activity across rewarded+rewarded, rewarded+unrewarded, unrewarded+rewarded, and unrewarded+unrewarded trials (n= 14, female n=6). (J) At ITI start, mPFC-NAc activity is significantly reduced to previous reward (unrewarded+unrewarded vs rewarded +unrewarded: Z=1.1935, p=0.0411; unrewarded+rewarded vs rewarded +rewarded: Z=4.7502, p<0.0001). (K) At ITI end, mPFC-NAc activity is significantly reduced to current reward (unrewarded +unrewarded vs unrewarded+rewarded: Z=3.7213, p=0.0004; rewarded+unrewarded vs rewarded+rewarded: Z=5.0067, p<0.0001). (L) Estimated mean vHip-NAc activity across rewarded+rewarded, rewarded+unrewarded, unrewarded+rewarded, and unrewarded+unrewarded trials. vHip-NAc activity is not modulated by current or previous outcome at (M) ITI start and at (N) ITI end. #p<0.1, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

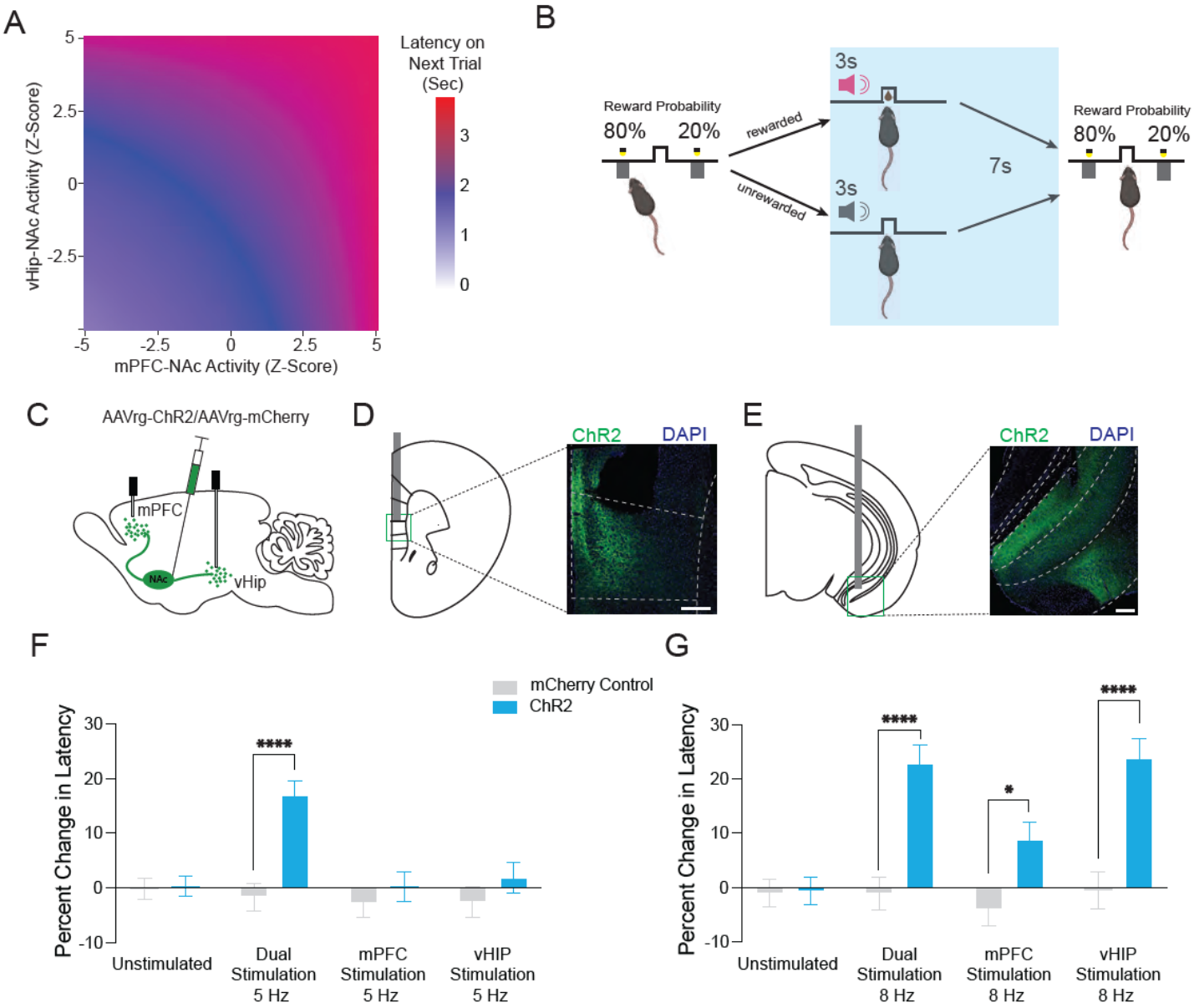

Figure 5. mPFC-NAc and vHip-NAc modulate task engagement. (A) Heatmap of estimated latency to respond on the subsequent trial given mPFC-NAc and vHip-NAc activity at ITI end shows that increased activity associates with longer latency. (B) Optogenetic stimulation protocol in the two-armed bandit task is delivered for the duration of the ITI to either mPFC-NAc, vHip-NAc, or simultaneously to both circuits. (C) Retrograding ChR2 or 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 not each circuit individually, increased latency to respond in ChR2 animals (n=13, n=7 females) compared to mCherry controls (n=12, n=6 females; 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 (n= 11, n=6 females) compared to mCherry controls (n= 12, n=6 females). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001