A corticostriatal learning mechanism linking excess striatal dopamine and auditory hallucinations

Kaushik Lakshminarasimhan[1,†], Justin Buck[1,2,†], Christoph Kellendonk[2,3], and Guillermo Horga[1,2]

[1]Zuckerman Mind Brain Behavior Institute, Department of Neuroscience, Columbia University, New York, NY, USA, [2]Department of Psychiatry, Columbia University, New York, NY, USA, [3]Department of Molecular Pharmacology and Therapeutics, Columbia University, New York, NY, USA, [†]equal contribution

1 Abstract

Auditory hallucinations are linked to elevated striatal dopamine, but their underlying computational mechanisms have been obscured by regional heterogeneity in striatal dopamine signaling. To address this, we developed a normative circuit model in which corticostriatal plasticity in the ventral striatum is modulated by reward prediction errors to drive reinforcement learning while that in the sensory-dorsal striatum is modulated by sensory prediction errors derived from internal belief to drive self-supervised learning. We then validate the key predictions of this model using dopamine recordings across striatal regions in mice, as well as human behavior in a hybrid learning task. Finally, we find that changes in learning resulting from optogenetic stimulation of the sensory striatum in mice and individual variability in hallucination proneness in humans are best explained by selectively enhancing dopamine levels in the model sensory striatum. These findings identify plasticity mechanisms underlying biased learning of sensory expectations as a biologically plausible link between excess dopamine and hallucinations.

2 Introduction

Hallucinations are false percepts frequently experienced by people with psychotic disorders. Our limited understanding of hallucination mechanisms at a biological circuit level precludes improvements to treatments which are often ineffective or poorly tolerated (Sommer et al., 2012). A promising venue to address this urgent need is formulating an end-to-end theory explaining how alterations in neural circuits drive changes in cognitive computations leading to hallucinatory experience—that is, a theory bridging the biology-experience gap (Fletcher & Frith, 2009).

A growing body of work has linked hallucinations to altered learning of sensory statistics (Friston, 2005; Adams et al., 2013; Corlett et al., 2019). Conceptualizing perception as an inferential process that combines learned sensory expectations with sensory evidence to form a percept, this work implicates altered learning or overweighting of sensory expectations in hallucinatory perception. In particular, in noisy sensory environments (e.g., signal detection tasks) psychotic individuals with hallucinations consistently show exaggerated biases toward prior expectations regardless of the true sensory evidence (Moseley et al., 2016; Powers et al., 2017; Cassidy et al., 2018; Corlett et al., 2019; Alderson-Day et al., 2022). While these insights represent substantial progress in our understanding of the cognitive-computational basis of hallucinations, existing theories still lack a detailed, biologically plausible implementation at the neural circuit level—a key requirement for a mechanistic explanation to be complete and actionable through biological interventions (X.-J. Wang & Krystal, 2014).

An established aspect of the neurobiology of hallucinations and other psychotic symptoms is elevated striatal dopamine (Laruelle et al., 1996; Abi-Dargham et al., 1998; Cassidy et al., 2018, 2019; McKetin et al., 2013), with pro-dopaminergic agents worsening and anti-dopaminergic drugs improving this symptom via striatal dopamine blockade (Kesby et al., 2018; Yun et al., 2023). A complete model of hallucinations must therefore incorporate dopamine and striatal circuits, including their increasingly well-documented functional heterogeneity. This is particularly important given that dopamine elevation in psychosis predominates in dorsal striatum (Howes et al., 2009; Kegeles et al., 2010; Mizrahi et al., 2012; McCutcheon et al., 2018). Some previous theories indeed centered around the functional role of dopamine excess in psychosis (Stevens, 1973; Kapur, 2003; Fletcher & Frith, 2009; Heinz & Schlagenhauf, 2010; Maia & Frank, 2017), although knowledge about dopamine heterogeneity has since soared. Specifically, in contrast with the well-established role of ventral striatal dopamine in reward learning, dopamine in dorsal parts of the striatum (including the so-called auditory striatum) responds to sensory features irrespective of reward (Menegas et al., 2017, 2018; Schmack et al., 2021), consistent with distinct anatomical projections from sensory cortex to this "sensory dorsal striatum" (Chen et al., 2021) and its causal contribution to perception (Xiong et al., 2015; Guo et al., 2018; L. Wang et al., 2018; Chen et al., 2022). A plausible circuit model of hallucinations should thus account for such anatomo-functional heterogeneity, specifying the involvement of distinct striatal modules in hallucination generation.

Here, we take a theory-driven approach to constructing a circuit-level computational model of hallucinations. Critically, we first derive a normative circuit model applying constraints from cross-species anatomical, physiological, and behavioral data. Our central hypothesis is that sensory-dorsal-striatal

dopamine signals violations in sensory expectations (i.e., sensory prediction errors). This inspires the corollary hypothesis that excess dorsal-striatal dopamine in hallucinating individuals biases learning of sensory expectation which in turn drives false (hallucinated) percepts. We first strengthen the theoretical validity of the hypotheses by demonstrating that an optimized model learns to represent the nigrostriatal dopamine signal as a sensory prediction error. We then validate key predictions of the intact circuit model using dopamine recordings and optogenetic stimulation of the sensory striatum in behaving rodents, as well as human behavior. Finally, we simulate selective alterations in dopamine function and find that behavioral data in humans with varying degrees of hallucination propensity is consistent with model predictions of excess sensory dorsal-striatal dopamine function. Together, the results suggest that a learning- and corticostriatal-plasticity-based perspective of heterogeneous striatal dopamine signaling can account simultaneously for hallucination-related biological, cognitive, and phenomenological phenotypes.

3 Results

Clinical hallucinations manifest as an increased tendency to exhibit false alarms with high confidence in perceptual signal-detection tasks. These hallucination-related perceptual biases correlate with increased dorsal-striatal dopamine. To investigate the mechanisms by which excess striatal dopamine may lead to high-confidence false alarms, we first take a computational view of the broader class of perceptual decision-making tasks. In such tasks, the state of the world is typically encoded in a series of noisy observations that unfold over time, and choosing an action yields a stochastic performance feedback or reward. One must therefore perform inference to estimate the underlying state from observations and then select an action based on the inferred state. The two processes impose unique constraints on learning. Optimal inference involves learning the statistics of the incoming stimulus (that is, *sensory expectation*) to minimize surprise, while selecting optimal actions in the inferred state requires learning the outcome statistics (that is, *value*) to maximize reward (**Figure 1B**). Since neither stimulus nor reward statistics are stationary in the real world, a normative approach to flexible perceptual decision-making should support learning of both types of statistics. Given that inference and action selection are cascaded processes, one could, in principle, bypass explicit learning of both the sensory expectation and value by relying on implicit trial-and-error learning based on performance feedback. However, previous studies have revealed engagement of both learning processes in tasks involving visual, temporal and auditory judgments, suggesting that perceptual decision-making is a two-stage process in which the first (sensory learning) stage can operate in the absence of external feedback while the second (reinforcement learning) stage requires feedback (Petrov et al., 2005; Dosher & Lu, 2017; Sohn & Jazayeri, 2021; Loewenstein et al., 2021). What biological mechanisms subserve the two learning objectives?

From a neurobiological standpoint, dopamine is known to support learning by modulating plasticity in corticostriatal synapses (**Figure 1C** – left). However, existing models informed by reinforcement-learning theories focus primarily on the role of dopamine in value learning. In such models, learning is mediated by reward prediction errors (rPE) signalled by the mesolimbic pathway, i.e., midbrain dopamine neurons projecting from the ventral tegmental area to ventral striatum (VS). The role of dopamine signalling in

the nigrostriatal pathway, i.e., from the substantia nigra to dorsal striatum is less clear. People with schizophrenia show elevated dopamine levels specifically in dorsal-striatal subregions targeted by this pathway, where dopamine synthesis and release capacity correlate with hallucination (and overall psychosis) severity (Laruelle et al., 1996; Howes et al., 2009; Kegeles et al., 2010; Cassidy et al., 2018, 2019). Furthermore, studies in mice indicate that neurons and dopamine signals in the nonmotor dorsal-striatal subregions play a causal role in perception (Guo et al., 2018; L. Wang et al., 2018; Chen et al., 2022), and that dopamine stimulation in these subregions induces biased (hallucination-like) percepts (Schmack et al., 2021). For simplicity, we refer to this subregion as the sensory striatum (SS) throughout the paper and use more specific terminology when referring to specific results from previous work. Based on the normative learning constraints outlined earlier, we hypothesized that rPE signals in VS dopamine facilitate learning of value while SS dopamine signals sensory prediction errors (sPE) – a specific type of prediction error appropriate for learning expectations about the stimulus.

Temporal-difference reinforcement learning specifies the precise algebraic form of time-varying rPE signals but the composition of sPE signals remains ambiguous. One potential solution is provided by a class of models that rely on the discrepancy between actual and predicted state (Gläscher et al., 2010; Gardner et al., 2018), but constructing such prediction errors requires perfect knowledge of the stimulus identity which may not always be available to the animal. While rewards in instrumental tasks can be used to infer stimulus identity, this does not explain learning of stimulus statistics in the absence of feedback as noted earlier. Therefore, we asked whether sPE could instead be computed without external supervision. To address this, we derived an error signal for minimizing the cumulative change in the internal belief induced by sensory input, on average. Unlike feedback-dependent rPE signals which are useful for maximizing cumulative reward, the resulting sPE signals depended on the momentary changes in internal belief (Methods). To test whether this signal can mediate biologically plausible learning and to generate predictions at the neural level, we constructed a corticostriatal circuit model informed by neuroanatomical constraints (**Figure 1D** – right).

3.1 Modular corticostriatal model of perceptual decision-making

In the model, distinct midbrain nuclei—substantia nigra and ventral tegmental area—project to distinct striatal regions—SS and VS, respectively—to mediate plasticity in corticostriatal synapses according to a biologically plausible three-factor learning rule (Łukasz Kuśmierz et al., 2017; Gerstner et al., 2018). According to this rule, the update to a particular synaptic weight W_{ij}^k from cortical neuron j to striatal neuron i in region $k \in \{\mathrm{SS}, \mathrm{VS}\}$ depends only on the pre- and post-synaptic activities, r_j and r_i, and a global prediction error signal δ_k conveyed by the midbrain dopamine inputs to that striatal region: $\Delta W_{ij}^k \propto r_i \phi'(r_j)\delta_k$, where $\phi(\cdot)$ denotes neuronal nonlinearity (Methods). SS and VS neurons receive inputs from different cortical regions—sensory and frontal, respectively—and summed activities from striatal neurons in each region project back to the same cortical regions, forming parallel loops (Foster et al., 2021). While the precise anatomical pathway from the striatum to the cortex is much more complex, we use summed activity as a proxy since it has previously been shown to be a good indicator of the

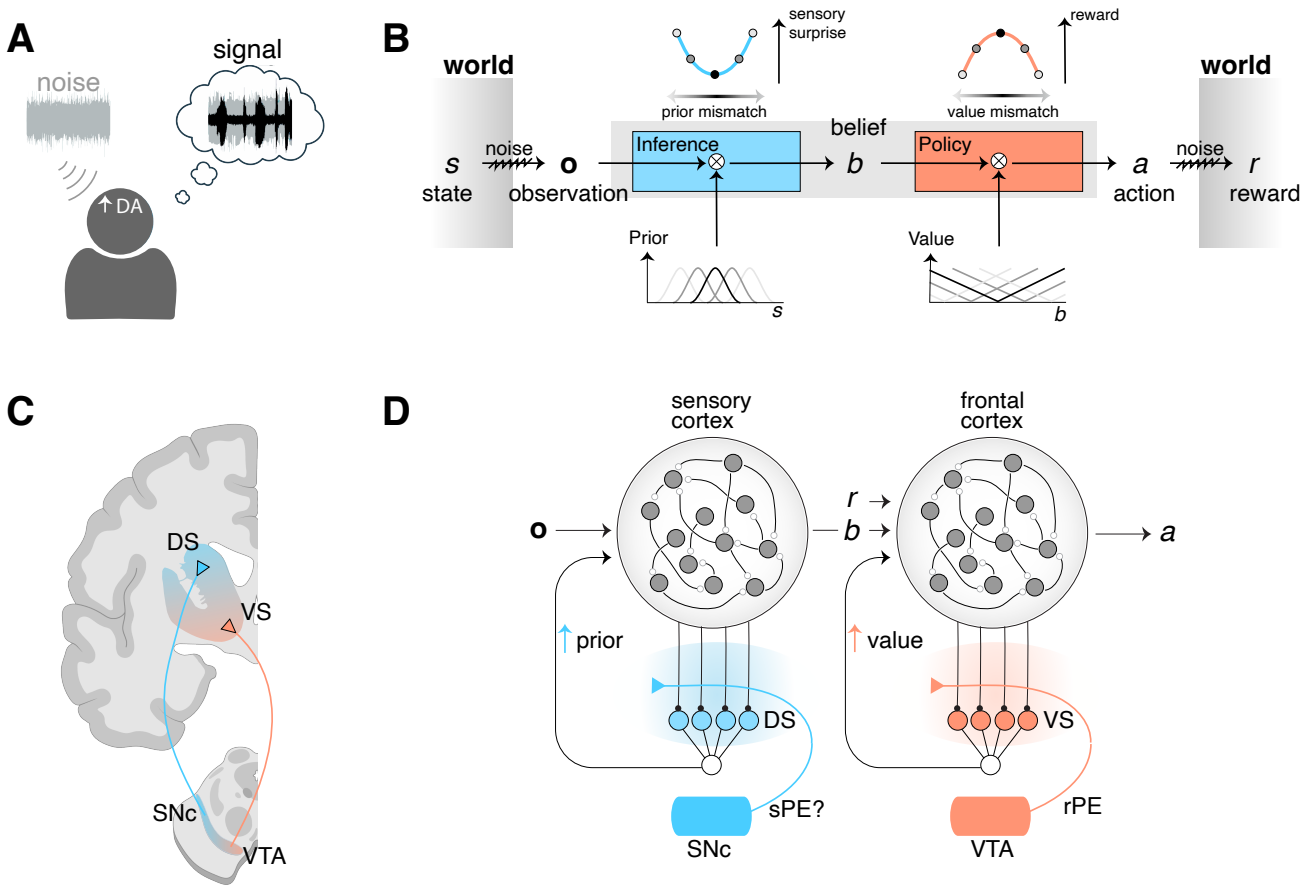

Figure 1: **Modeling framework for auditory hallucinations**. **A.** Excess striatal dopamine is linked to hallucinations, which manifests as an increase in confident false alarms in signal detection tasks. **B.** Perceptual decision-making entails inference and action selection stages that must fulfill distinct desiderata – minimizing sensory surprise and maximizing reward, objectives that can be met by learning the true expectation over states (i.e., prior) and rewards (i.e., value) respectively. **C.** Distinct dopaminergic pathways target the ventral striatum and the sensorimotor areas of the dorsal striatum. **D.** Circuit-level mechanisms of the computations in (B) can be implemented in a biologially plausible manner via plasticity mediated by distinct dopamergic prediction error signals in parallel corticostriatal loops. VS: ventral striatum, DS: Dorsal striatum, VTA: ventral tegmental area, SNc: Substantia nigra pars compacta, DA: dopamine, rPE: reward prediction error, sPE: sensory prediction error.

topographical organization of corticostriatal domains (Peters et al., 2021). Sensory and frontal cortices are modeled as recurrent neural networks, optimized respectively for perceptual inference and value-based decision-making (Methods). Concretely, sensory observations are provided as external input to the model sensory cortex, which outputs a graded estimate of the posterior probability of the latent state i.e., belief to the model frontal cortex, which in turn outputs a binary yes/no response. Critically, both cortical computations depend on learning in the striatum. Accurate perceptual inference relies on learning of sensory expectations in the SS while accurate value-based decision depends on value learning in the VS. We assume that plasticity in the corticostriatal synapses in SS is modulated by the sPE signal derived above, where the momentary change in internal belief was calculated by taking the time derivative of the

output of the sensory cortex (Methods). Plasticity in VS is modulated by a standard temporal-difference rPE signal following previous works (Rao, 2010; Hennig et al., 2023) (Methods).

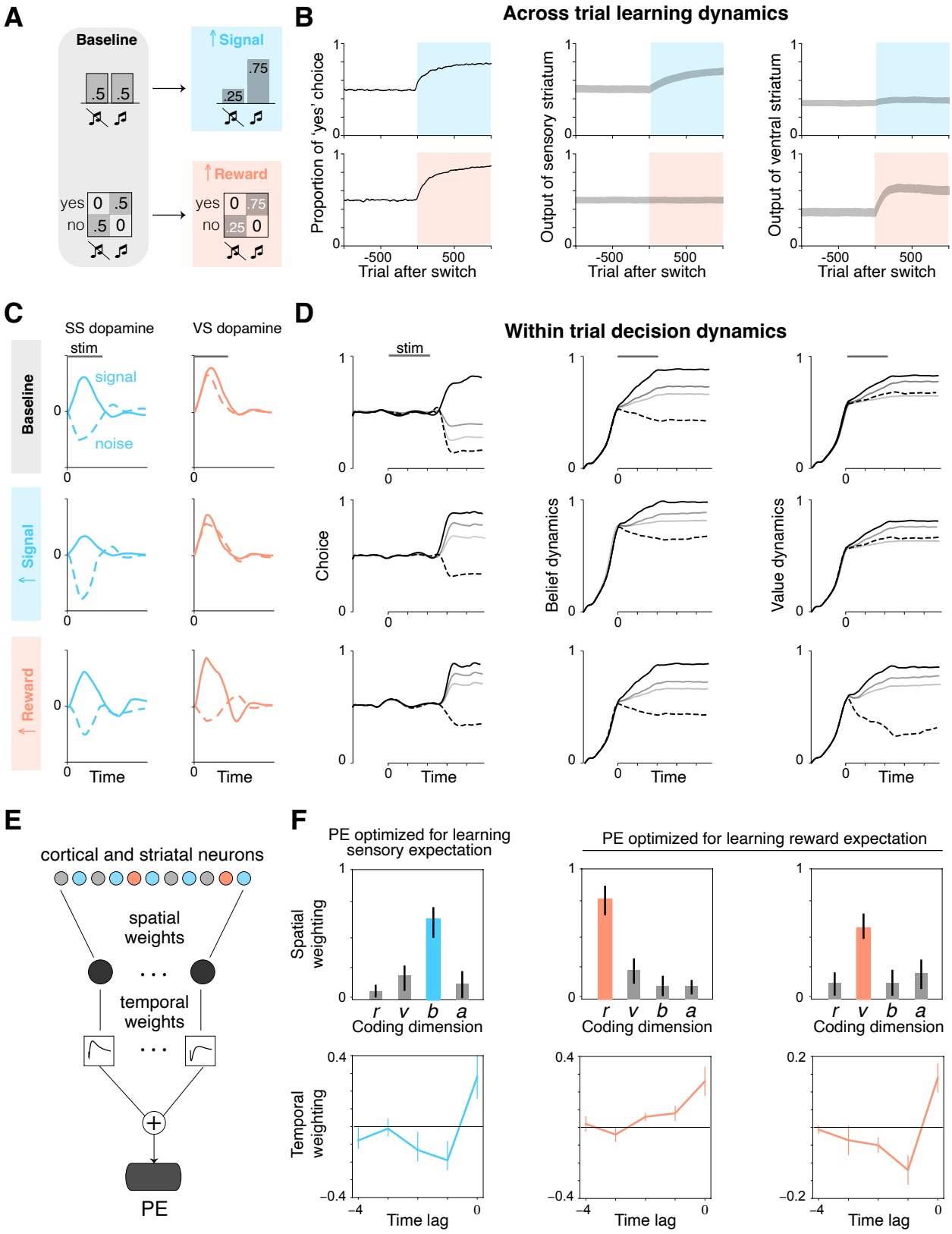

Figure 2: (Caption next page.)

Figure 2: **Distinct computational mechanisms of adaptation to signal and reward statistics**. **A.** Starting from a block of baseline trials (signal/noise stimuli presented in equal proportion and correct yes/no rewarded in equal proportion), stimulus and reward statistics were separately manipulated by increasing the fraction of signal trials (top) and increasing the probability of rewarding a correct 'yes' response (bottom) respectively. **B.** Evolution of average reward (left), sensory expectation reflected in the output of model SS (middle), and the reward expectation reflected in the output of model VS (right) across trials before and after manipulation of stimulus (top row) and reward (bottom row) statistics. **C.** Stimulus-evoked model dopamine transients in signal (solid lines) and noise (dashed lines) trials. Rows correspond to different conditions and columns correspond to different striatal sub-regions. **D.** Average within-trial dynamics of choice (left), belief (middle), and value (right) under different conditions (rows). Dashed line corresponds to noise trials while lines with gray hues correspond to signal trials, with darker hues denoting a higher signal-to-noise ratio. **E.** Prediction error (PE) was optimized by expressing it as a sum of multiple spatiotemporal filters. Each spatial filter corresponds to a weighted sum of the activity of units in the cortical and striatal regions and each temporal filter corresponds to a weighted sum of the output of the spatial filter across a moving window. **F.** Left panel: PE optimized for learning sensory expectation – magnitude of alignment of the optimized spatial weights with activity dimension encoding reward (r), value (v), belief (b), and action (a), and the corresponding temporal weights. Right two panels: Similar to the left panel, but showing weights optimized for learning reward expectation. Error bars denote 95% confidence intervals obtained by bootstrapping across simulations.

To test the learning performance of this model, we varied stimulus or reward statistics in a signal-detection task by manipulating either the proportion of 'signal' trials or the fraction of rewarded correct 'yes' responses across blocks (**Figure 2A**; Methods). Although both manipulations initially disrupted performance following block transitions, the model adapted to them by adjusting its choices (**Figure 2B** – left). Critically, adaptation to manipulations of stimulus and reward statistics was mediated almost exclusively by changes in the output of SS (encoding sensory expectations) and VS (encoding value), respectively (**Figure 2B** – middle vs. right). Dopamine signaling in these regions revealed a similar dissociation wherein stimulus and reward manipulations induced lasting changes in stimulus-induced dopamine transients within the model SS and VS, respectively, albeit in opposite directions. Increasing the base rate of signal trials decreased the amplitude of signal-evoked SS dopamine after learning (**Figure 2C** – left), whereas increasing the reward probability associated with a correct 'yes' response increased signal-evoked VS dopamine (**Figure 2C** – right). This dichotomy in striatal dopamine was reflected in the model neural activity dynamics that emerged after learning. Although both types of manipulations produced comparable changes in choice-related response dynamics leading to similar levels of adaptation (stimulus $\rightarrow$ action; **Figure 2D** – left), manipulating stimulus statistics exclusively modified the transformation implemented by the sensory corticostriatal loop i.e., perceptual inference (stimulus $\rightarrow$ belief; **Figure 2D** – middle), while reward manipulation modified value dynamics that gate policy computation in the frontal corticostriatal loop (belief $\rightarrow$ action; **Figure 2D** – right).

Above results suggest that sPE constructed from momentary changes in internal belief can support a biologically plausible form of self-supervised learning of sensory expectations, alongside learning of reward expectations from temporal-difference rPE. We aimed to determine whether learning sensory expectations using the specific type of teaching signal assumed above, namely the time-derivative of internal belief, is normative. To address this, we expressed the prediction error encoded by striatal dopamine as an

arbitrary spatiotemporal transformation of cortical and striatal activity and then used gradient-descent to optimize the spatial and temporal kernels or weights (**Figure 2E**; Methods). Since this procedure seeks to optimize the teaching signal rather than task performance, it corresponds to meta-learning or learning-to-learn (Hospedales et al., 2022), a technique that has previously been applied to many neural systems (Wang et al., 2018; Jiang & Litwin-Kumar, 2021; Lakshminarasimhan et al., 2024). When we optimized the prediction error for learning sensory expectation, the resulting spatial weights were aligned primarily with the dimension of neural activity encoding the belief that the stimulus contained a signal and the temporal weights resembled a derivative operator (**Figure 2F** – left; Methods). Alignment with dimensions that encoded reward, value (i.e., expected reward), and action was much weaker. This suggests that sPE constructed using momentary changes in internal belief optimizes biologically plausible learning of sensory expectations in this circuit architecture. In contrast, a similar approach to optimize learning of reward expectation learning revealed two spatiotemporal filters that together resembled the temporal-difference rPE signal: one for selecting the current reward, and the other for extracting momentary changes in value (**Figure 2F** – middle and right).

3.2 Signatures of sensory prediction errors in sensory striatal dopamine

We asked whether model VS and SS dopamine signals can inform the interpretation of dopamine signals recorded in those areas across different tasks. We considered two mouse studies in which dopamine recordings were performed both in the SS (specifically, the tail of striatum or TS for the remainder of this section) and VS under different conditions, allowing for a direct comparison of dopamine dynamics between the model and data.

In the first of these studies, (Menegas et al., 2017) tracked dopamine dynamics across the TS and VS during associative learning. They tested whether repeated training affected the observed pattern in TS and VS dopamine by introducing a new odor paired with a reward every day, and then measured dopamine activity while learning a new odor-reward association (**Figure 3A**; Methods). TS dopamine exhibited a response to novel odors which decreased across repeated exposures. In contrast, VS dopamine initially responded to rewards and only later on to odors in a manner that reflected the learned association between odors and rewards (**Figure 3B** – left). Simulating the model under associative learning yielded a strikingly similar pattern where odor-induced TS and VS dopamine dynamics decreased and increased respectively across trials (**Figure 3B** – right). The computational role of dopamine in the model offers an interpretation for these experimental results. Because model VS dopamine encodes rPE, the trajectory of dopamine signals reflects the backward progression of this temporal-difference error from rewards to odors. In contrast, TS dopamine encodes sPE and reflects the animal's sensory surprise upon encountering an unexpected stimulus, which gradually reduces with repeated presentations as the animal comes to expect this stimulus. Note that we treat reward delivery itself as another sensory input in our simulations (Methods) since the weak reward-related dopamine signals in TS have been attributed to the click of the water valve used to deliver reward (Menegas et al., 2018).

In a different study, (Schmack et al., 2021) recorded dopamine dynamics in the same striatal regions but

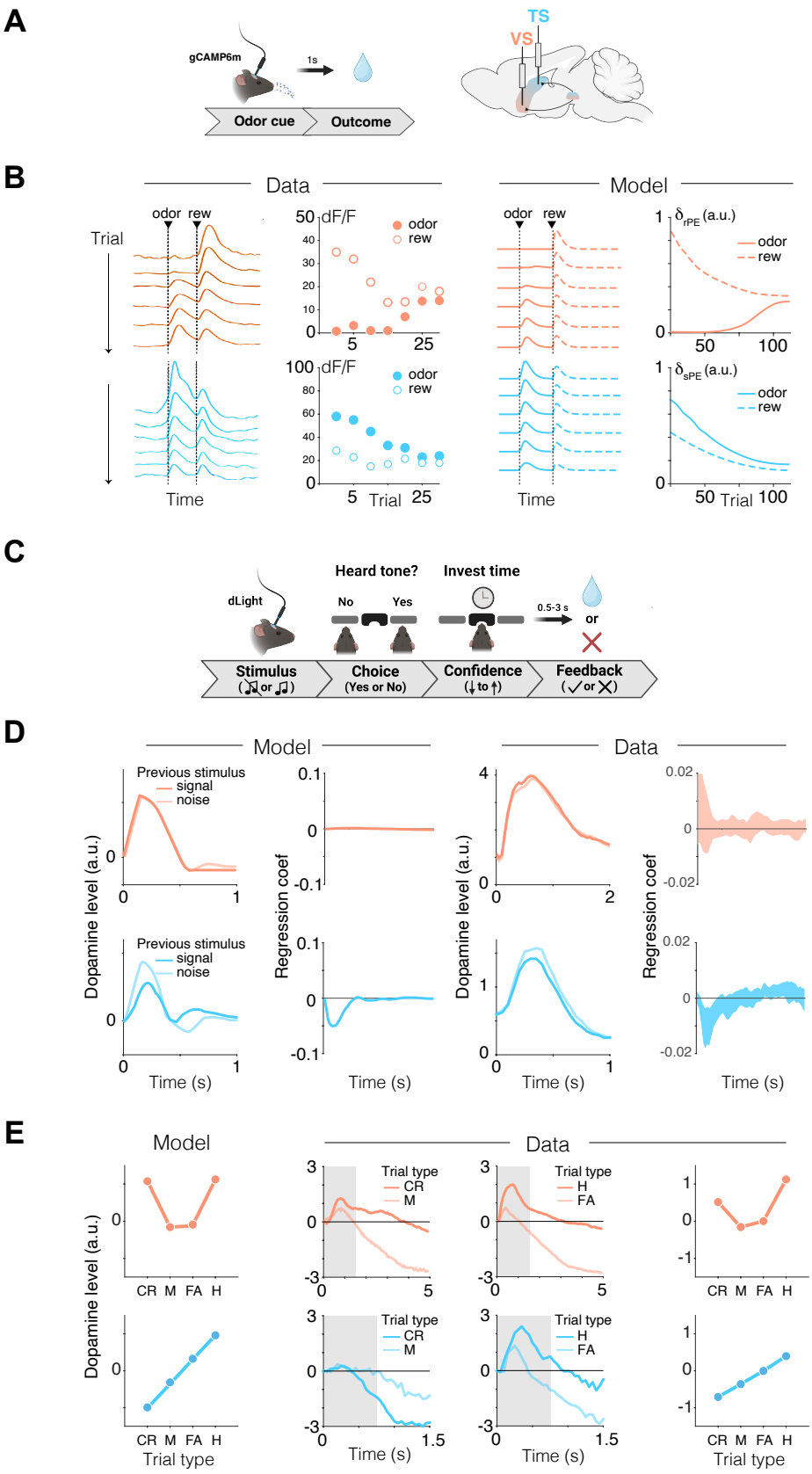

Figure 3: (Caption next page.)

Figure 3: **Signatures of distinct prediction errors in associative learning and signal detection**. **A.** Schematic of the associative learning task. Mice were exposed to repeated pairings of a novel odor and reward delivery. **B. Left**: Dopamine dynamics in the mouse VS (top, orange) and TS (bottom, cyan) across trials. **Left middle**: Evolution of the amplitude of dopamine transients evoked by odor (open circles) and reward (filled circles). **Right**: Similar to the left panels but showing the evolution of rPE (top, orange) and sPE (bottom, cyan) in model simulations. **C.** Schematic of the signal detection task. Following stimulus presentation, mice reported their choice and confidence, estimated as the amount of time invested in the choice, and received feedback in the form of a reward. **D. Left**: Model-predicted stimulus-evoked rPE (top) and sPE (bottom), conditioned on the previous trial's stimulus. **Left middle**: Regression weights quantifying the influence of previous trial's stimulus on the prediction errors at each various time points relative to stimulus onset. **Right**: Similar to the left panels but showing average dopamine levels in VS (top, orange) and TS (bottom, cyan) across mice ($n = 6$). Error bars denote 95% confidence intervals. **E. Left**: Time-averaged stimulus-evoked rPE (top, orange) and sPE (bottom, cyan) in the model simulations during trials classified based on stimulus-response relationship (CR: correct reject, M: miss, FA: False alarm, H: Hit). **Left middle**: Stimulus-evoked dopamine dynamics in VS (top) and TS (bottom) averaged across mice during CR trials and M trials. **Right middle**: Similar to the left middle panel, but during H and FA trials. **Right**: Time-averaged dopamine levels during the time windows denoted by the gray shaded regions in the middle panels, where window lengths are matched to the timescale of autocorrelation in dopamine signals in the two regions.

during auditory signal detection (**Figure 3C**). In this task, mice initiated trials with a center nosepoke, followed by a nosepoke at a left or right port depending on whether it heard a tone. To look for signatures of sPE and rPE in the data, we first examined dopamine signals in the model trained to perform signal detection. We found that the amplitude of signal-evoked model SS dopamine, but not VS dopamine, was modulated by the previous trial's stimulus identity – in addition to the scaling with signal strength (SNR), which has been previously documented. Specifically, encountering a signal trial increases the expectation of a tone, decreasing the magnitude of signal-evoked response in SS dopamine (e.g., the sPE) in the subsequent trial (**Figure 3D** – left). We found this precise pattern of modulation by stimulus history in the mouse TS. In contrast, mouse VS dopamine was not significantly modulated by stimulus history (**Figure 3D** – right): regression analyses showed that the previous trial's stimulus identity (signal/noise) negatively modulated dopamine signal during the stimulus window in TS but not VS across all trials and mice (s_{t-1} main effect; VS: $t_{30486} = 0.645$, $p = 0.519$; TS: $t_{21752} = -2.59$; $p = 9.61 \times 10^{-3}$) and that the regional difference in this effect was significant ($s_{t-1} \times$ region: $t_{52241} = -6.31$, $p = 2.86 \times 10^{-10}$). While the stimulus-history effect in TS dopamine is consistent with sPE signaling, we asked whether this signal specifically reflects fluctuations in subjective belief about the presence of a signal i.e., signed sensory surprise, as predicted by the model. When trials were grouped by stimulus-response relationship in the order of increasing confidence about the presence of a signal (CR < M < FA < H), the amplitude of model stimulus-evoked TS dopamine, but not VS dopamine, increased monotonically (**Figure 3E** – left). We found a similar trend in the mouse data: on average, correct rejections (CR) and hits (H) were associated with the lowest and highest levels of TS dopamine (**Figure 3E** – right). In contrast, VS dopamine was relatively high in both cases, consistent with rPE signaling.

The above results show that dopamine dynamics in the mouse TS but not VS resembles model sPEs. We demonstrated earlier that this prediction error signal is optimized for learning sensory expectations via

plasticity mechanisms. If this is the case, then perturbing TS dopamine should have a qualitatively similar effect on behavior as that on the model performance. This leads to two specific predictions. First, such perturbation should induce a bias in signal expectation rather than reward expectation. Second, the effect of this perturbation should gradually build up over time. We tested both predictions by re-analyzing the data from optogenetic stimulation experiments carried out during the same paradigm. Briefly, dopamine terminals in the TS were chronically stimulated via a laser across blocks of fifty trials, interleaved with baseline blocks in which the laser was turned off (**Figure 4A**; Methods).

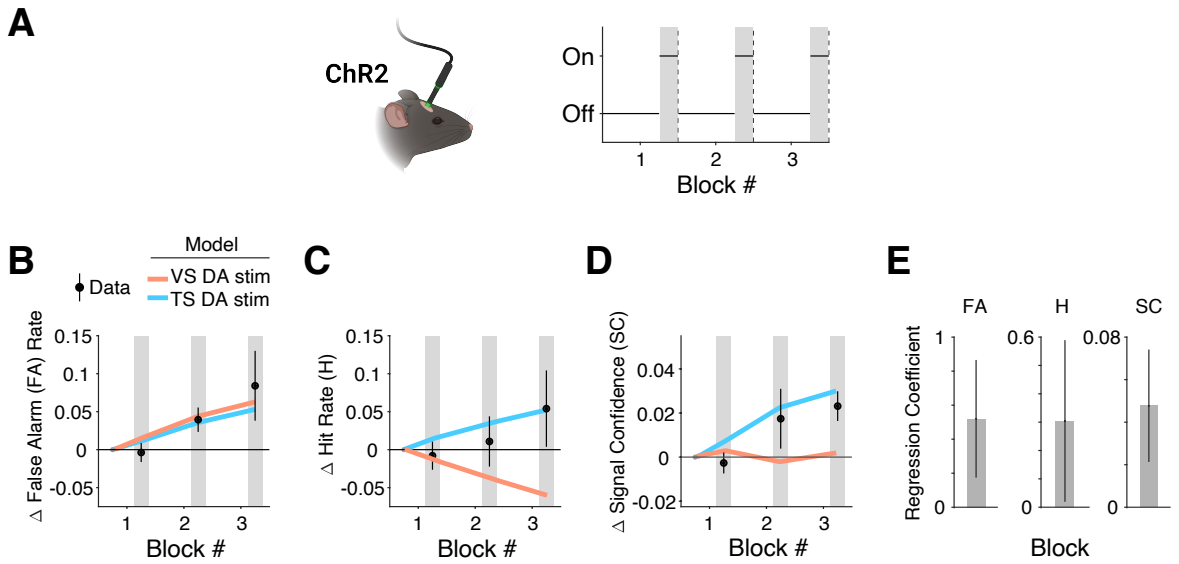

Figure 4: **Evolution of signal expectation from dopamine stimulation**. **A.** Schematic of the optogenetic stimulation protocol. Laser was ON or OFF throughout the respective blocks. **B.** Change in the rate of false alarms with respect to the baseline (i.e., the first block of the experimental session) in each block. **C.** Similar to B, but showing the change in the rate of hits. **D.** Similar to B, but showing the change in signal confidence.**E.** Regression coefficient quantifying the effect of the block number on the false alarm rate (FA), hit rate (H), and signal confidence (SC). For B, C, and D, error bars denote standard error of the mean across mice and for E error bars denote 95% confidence intervals.

Model simulations indicated that VS dopamine stimulation increases the rate of false alarms and misses, whereas TS stimulation increases false alarms, hits, as well as the confidence that a signal is present, i.e., signal confidence (**Figure 4B-E** – orange vs cyan solid lines). In the mouse stimulation data, we found that, on average, stimulation significantly increased the rate of false alarms, as reported previously (Schmack et al., 2021) (stimulation main effect: $t_{14086} = 3.70; p = 2.18 \times 10^{-4}$). Moreover, stimulation also increased the signal confidence across all trials (stimulation main effect: $t_{9803} = 2.76; p = 5.79 \times 10^{-3}$), consistent with the hypothesis that TS dopamine stimulation influences behavior by increasing signal rather than reward expectation. To test whether these effects were due to a learning rather than an instantaneous effect (driving role) of dopamine, we estimated how the above response measures evolved across blocks (47 ± 9 trials per block). We found that false alarms, hits, signal confidence all exhibited a significant increase from baseline as a function of blocks within a stimulation session (**Figure 4B-E**; false alarms across session, $t_{2875} = 2.98; p = 2.93 \times 10^{-3}$; hits, $t_{2850} = 2.11; p = 0.0351$; signal confidence, $t_{5728} = 3.58; p = 3.49 \times 10^{-4}$). Notably, increasing model VS dopamine predicts an increase not only in

the rate of false alarms but also misses, the latter of which is contradicted by data (**Figure 4C** – orange). Taken together, these analyses suggest that dopamine-mediated plasticity mechanisms in the TS play a role in learning sensory expectation.

3.3 Perceptual decision-making engages a two-stage learning process in humans

The results from the optogenetic stimulation study demonstrate distinct behavioral signatures of excess dopamine in the sensory striatal regions beyond changes in false alarm rates. Previous studies have noted that some of these signatures, such as signal confidence, correlate with hallucination propensity and clinical hallucinations in humans (Powers et al., 2017; Schmack et al., 2021). However, the extent to which hallucinatory experiences may be caused by an increase specifically in sPE signaling and whether they reflect an altered learning process rather than a constant perceptual bias remain unclear. This is in part because previous studies have typically investigated behavioral adaptation to either sensory or reward statistics but not both.

Motivated by the modular architecture in which stimulus and reward statistics are learned in distinct corticostriatal pathways, we developed a variant of a signal-detection task to identify unique signatures of learning sensory and reward expectations (**Figure 5A** – left). On each trial, participants experienced an ambiguous auditory stimulus (i.e., 1 kHz tone of varying SNR) and reported whether they heard a tone as well as the confidence associated with their judgement. Participants then received feedback about whether a tone was present and whether their choice resulted in a reward bonus. To engage both sensory and reward-based learning mechanisms, we manipulated both whether signals are more likely to be present or absent (i.e., sensory prior) and whether a 'yes' choice is likely to produce feedback indicating more or less reward (i.e., action value) (**Figure 5A** – right).

The model predicts that each of the two manipulations will drive distinct patterns of behavior in this task. As we show later, the manipulations gradually modify the pattern of behavioral responses, implicating a slow process in which stimulus and reward statistics are learned in an incremental manner. Sensory and reward learning manifest as distinct trial-history effects in the model. First, since both sensory and reward learning pathways influence choice via feedback (striatal output) to the cortex, the model predicts that choices would shift in accordance to the previous trial's stimulus and reward (**Figure 5B,C** – left, solid lines). Indeed, human participants' choices were biased by the previous trial's stimulus and reward (**Figure 5B,C** – left, filled circles; $s_{t-1} : t_{66116} = 22.9; p = 2.70 \times 10^{-115}$; $r_{t-1} : t_{66116} = 19.7; p = 5.99 \times 10^{-86}$). Second, since only the sensory-learning pathway (SS) influences belief estimates by conveying the sensory expectation, the model predicts that signal confidence would shift in alignment with the previous stimulus (**Figure 5B** – middle, solid lines) but not previous reward feedback (**Figure 5C** – middle, solid lines). Consistently, participants' signal confidence was more strongly influenced by the previous trial stimulus than previous reward feedback (**Figure 5B,C** – middle, filled circles; $F_{1,66116} = 557.13; p = 1.14 \times 10^{-122}$). Third, only the reward-learning pathway (VS) influences the policy, defined as the action taken in a given belief state and estimated as the fraction of yes choices binned by signal confidence for each participant. The model predicts that policy should shift in congruence with the

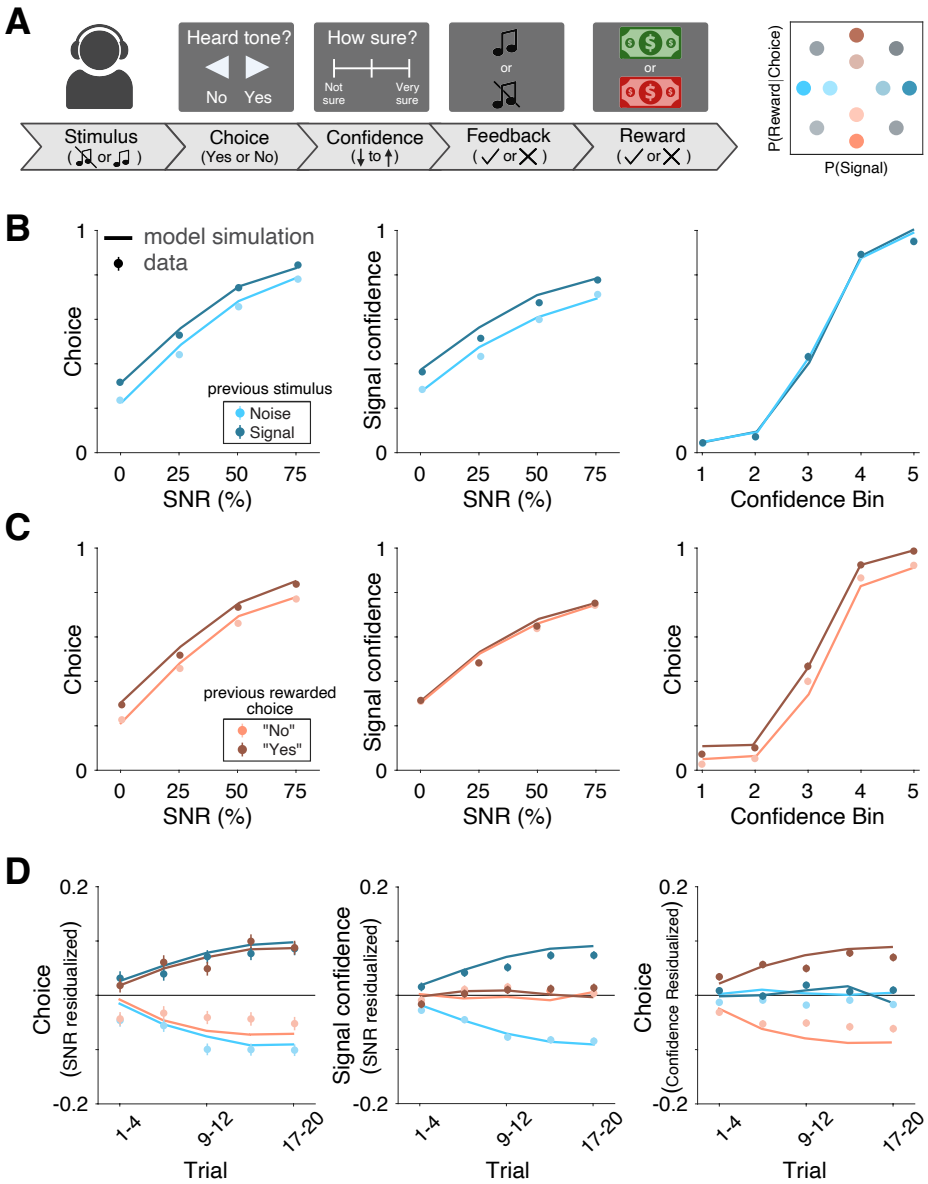

Figure 5: **Distinct behavioral signatures of perceptual and reward learning**. **A.** Schematic of the task. **Left**: Human participants made a binary judgement about whether they heard a tone, reported the associated confidence on a continuous scale, and received feedback. **Right**: The fraction of signal trials and the fraction of rewarded yes responses were manipulated across blocks. **B.** Effect of signal history. **Left**: The fraction of yes choices as a function of the signal to noise ratio (SNR). **Middle**: Signal confidence as a function of SNR. **Right**: Participants' policy, quantified as the fraction of yes choices as a function of confidence. **C.** Similar to B, but showing the effect of reward history. **D.** The evolution of choice, signal confidence, and policy across trials of the block for different manipulations. Error bars denote standard error of the mean across participants.

previous reward feedback (**Figure 5C** – right, solid lines) but not previous stimulus (**Figure 5B** – right, solid lines). In line with this, participants' policy was more strongly influenced by previous reward feedback than previous trial stimulus (**Figure 5B,C** – right, filled circles; $F_{1,66116} = 439.47; p = 2.93 \times 10^{-97}$). Overall, this alignment with model predictions suggests that human participants use distinct mechanisms

to adapt to changes in sensory and reward statistics.

In the model, adaptation to statistics is achieved by dopamine-mediated plasticity mechanisms. To further confirm that the human participants engaged mechanisms that relied on a gradual learning process, we examined the evolution of all three variables—choice, confidence, and policy—across trials. We found that all quantities shifted gradually, plateauing towards the end of the block, suggesting that adaptations to each of the two manipulations is the outcome of slow learning processes (**Figure 5D** –left; P(Signal)×Trial: $t_{69593} = 7.33; p = 2.38 \times 10^{-13}$; P(Reward|Choice) × Trial: $t_{69593} = 7.93; p = 2.15 \times 10^{-15}$ –middle; P(Signal) × Trial: $t_{69593} = 12.36; p = 4.56 \times 10^{-35}$; P(Reward|Choice) × Trial: $t_{69593} = 2.11; p = 0.035$ – right; P(Signal) × Trial: $t_{69593} = -1.71; p = 0.0873$; P(Reward|Choice) × Trial: $t_{69593} = 9.90; p = 4.28 \times 10^{-23}$).

3.4 Hallucination propensity is linked to increased sensory prediction errors

We now turn to the key question that motivated this study. What is the computational mechanism underlying the robust empirical link between hallucination severity and elevated striatal dopamine release? To address this, we combine insights from our analysis of the physiology of dopamine signaling in mouse striatum and history-dependent behaviors in humans. Recall that stimulating dopamine terminals carrying signals that bear signatures of sPE in mice led to a slow build-up of signal confidence over the course of the experimental session. In humans, both stimulus and reward history biased future choices to a comparable extent but only the former also biased the confidence associated with the presence of a signal. Based on these findings, we reasoned that if hallucinations primarily result from increased sPE, then participants with increased hallucination proneness would also exhibit greater stimulus-history effects.

To test this, using the data from human participants we asked whether the behavioral effects of stimulus history interacted with the participants' hallucination proneness. In a sample (n=290) pre-screened to ensure a substantial proportion of individuals with high hallucination propensity (Methods), we found that higher hallucination propensity (CAPS) score was associated with a greater tendency to report hearing a signal (including on noise trials, i.e., false alarms) following signal trials. (**Figure 6A,C** – left; s_{t-1} × CAPS: $t_{66108} = 2.48; p = 0.0130$). Critically, this differential effect of stimulus history on choice was mirrored by a similar effect on signal confidence following signal trials (**Figure 6A,C** – middle; s_{t-1} × CAPS: $t_{66108} = 4.01; p = 6.00 \times 10^{-5}$), but not on the policy (**Figure 6A,C** – right; s_{t-1} × CAPS: $t_{66112} = 1.60; p = 0.111$). These results are readily explained by adding a positive offset to the dopamine signals specifically in the model SS (Figure 6A – solid lines), but not in the VS, suggesting that biased perception driving hallucination proneness could arise from elevated sPE in the SS.

In contrast, we did not find a differential influence of reward history across hallucination propensity on the pattern of choices (**Figure 6B** – left; r_{t-1} × CAPS: $t_{66108} = 0.446; p = 0.655$), the pattern of signal confidence (**Figure 6B** – middle; r_{t-1} × CAPS: $t_{66108} = -0.736; p = 0.461$), or the policy (**Figure 6B** – right; r_{t-1} × CAPS: $t_{66112} = 0.389; p = 0.697$) and this was recapitulated by simulations in which dopamine in the model VS faithfully conveyed reward prediction errors in both groups (**Figure 6B** –

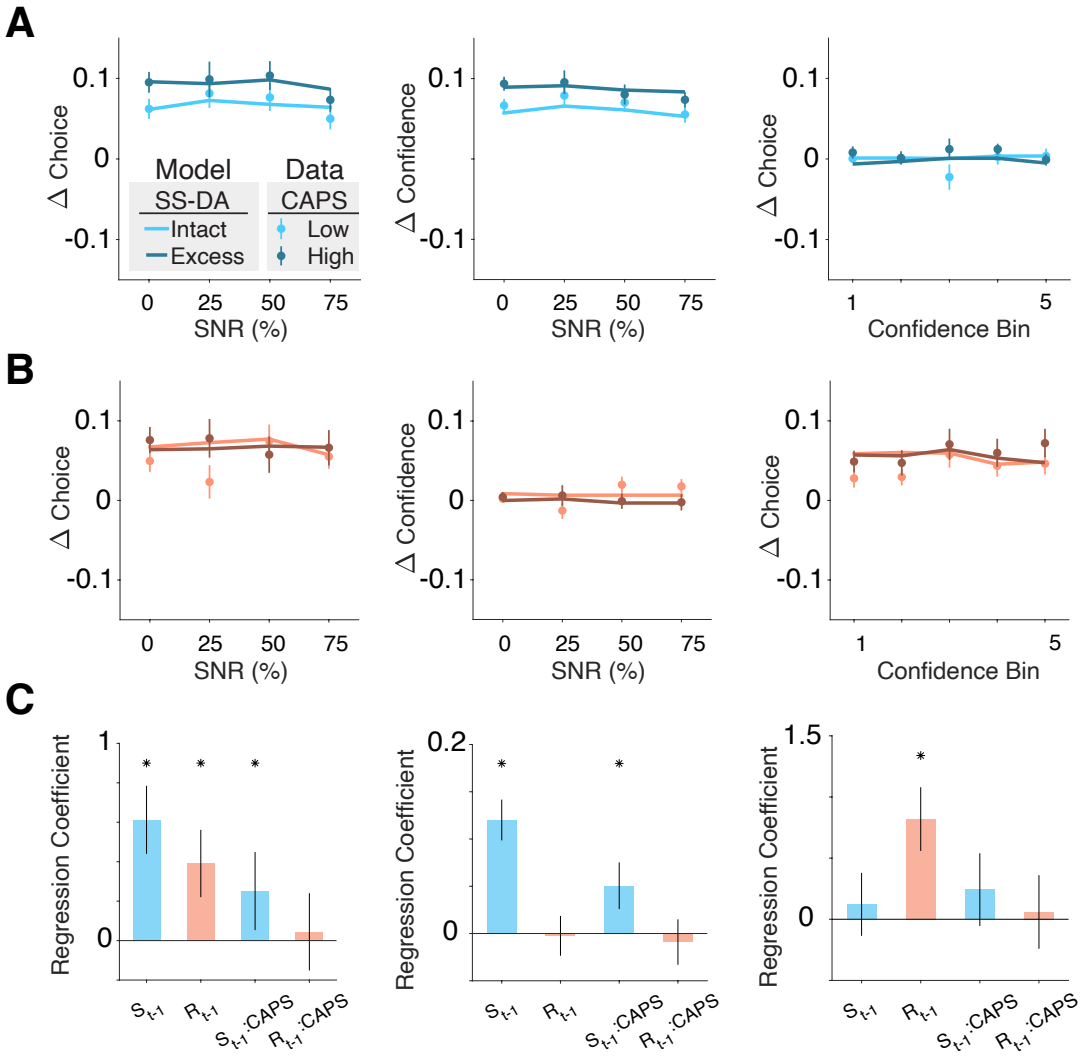

Figure 6: **Hallucination proneness exaggerates stimulus-history effects**. **A. Left**: The change in the rate at which participants report hearing a signal as a function of signal to noise ratio, after a signal trial, relative to after a noise trial. **Middle**: Similar to the left panel, but showing the change in the reported perceptual confidence. **Right**: Similar to the left panel, but showing the change in policy i.e., rate at which participants report hearing a signal as a function of confidence. Participants are grouped by their CAPS scores (filled circles, Methods). Solid lines show model simulations in which dorsal striatal dopamine was either intact (light) or scaled up (dark; Methods). **B.** Similar to A, but showing the change in response measures due to the previous trial's reward feedback. **C.** Coefficients of a mixed effects model, showing the effect of previous trial's stimulus (s_{t-1}) and reward (r_{t-1}), hallucination proneness (CAPS), and the interactions ($s_{t-1} \times$ CAPS, $r_{t-1} \times$ CAPS) on choice (left), confidence (middle) and policy (right). Asterisks denote p <0.05. For A and B, error bars denote standard error of the mean across participants and for C error bars denote 95% confidence intervals.

solid lines). Altogether, these results suggest that biased hallucination-like percepts are consistent with alterations in learning and SS corticostriatal plasticity due to elevated sPE signaling.

4 Discussion

We developed an anatomically constrained, normative circuit model of corticostriatal loops in which dopaminergic prediction-error signals in the ventral and dorsal sensory striatum are optimized to respectively support value learning and sensory learning. Dopamine dynamics in the mouse dorsal sensory (tail) striatum mirrored sensory prediction errors essential for learning sensory expectation, and dopamine stimulation in this region gradually adjusted perceptual estimates indicating a plasticity-based mechanism. In an auditory detection task, human behavior showed both sensory and reward learning, but only changes in the former correlated with hallucination proneness. These findings reveal a potential computational circuit mechanism underpinning the link between excess striatal dopamine and auditory hallucinations—and in doing so address a crucial gap in the development of computational circuit-level models of the altered subjective experiences characteristic of psychiatric illness (here, the emergence of biased internal beliefs driving hallucinatory percepts) (Fletcher & Frith, 2009; X.-J. Wang & Krystal, 2014).

Although the complex pathophysiology of psychosis points to a dysfunction in a broad range of synaptic mechanisms involving dopamine, glutamate, GABA, etc., almost all effective antipsychotic medications target striatal dopamine receptors and dopamine stimulants induce psychotic symptoms in healthy individuals (Curran et al., 2004; Sommer et al., 2012; Lally & MacCabe, 2015). Unsurprisingly, some of the earliest theories of psychosis focused on dopamine (Stevens, 1973; Seeman & Kapur, 2000; Kapur, 2003). Reflecting the dominant computational views at the time, these theories appealed to the role of dopamine in reward learning and motivational salience. However, the problems with these early theories were twofold. First, they did not directly model how excess dopamine leads to psychotic symptoms like hallucinations, instead relying on paradigms like associative and motor learning for empirical validation (Walter et al., 2009; Gradin et al., 2011). Second, they were not neurobiologically grounded in a circuit model and therefore carried limited mechanistic and translational relevance—with few notable exceptions (Maia & Frank, 2017). Consequently and despite established convergent support, state-of-the-art theoretical frameworks focused on pathological inference like those based on predictive processing have sidelined the 'dopamine hypothesis' in favor of models of cortical imbalance between excitation and inhibition (Kehrer, 2008; Jardri & Denève, 2013; Sterzer et al., 2018; Keller & Sterzer, 2024). However, an emphasis on inferential mechanisms does not preclude a role for dopamine (Fletcher & Frith, 2009; Horga & Abi-Dargham, 2019). In fact, a growing body of basic research using modern molecular tools points to anatomically segregated and physiologically distinct dopamine signaling pathways, including a role in perception for the sensory striatum and dopamine therein (L. Wang et al., 2018; Guo et al., 2018; Schmack et al., 2021; Chen et al., 2022). This development has paralleled proposals that extend the canonical, reward-centric computational theories by rethinking dopamine signaling as a generalized prediction error (Gardner et al., 2018). Informed by these empirical and theoretical results, we developed a corticostriatal model in which dopamine contributes to the inferential process by enabling self-supervised learning of sensory expectations in the sensory striatum.

4.1 Novelty and implications

The proposed model of the sensory striatum differs from alternative models employing different types of outcome-specific prediction errors (Gardner et al., 2018; Lee, Sagiv, Engelhard, Witten, & Daw, 2024) in that it allows for learning exclusively from internal beliefs, i.e., in the absence of actions and/or outcomes. This is in line with empirical work showing that humans (and other animals) can learn stimulus statistics in the absence of feedback (Petrov et al., 2006; Zylberberg et al., 2018; Fleming et al., 2019; Loewenstein et al., 2021; Schmid et al., 2024). A previous algorithmic account proposed that in the absence of feedback, sensory expectations must be updated on the basis of counterfactual beliefs evaluated using uniform priors to avoid self-reinforcement or double counting (Zylberberg et al., 2018). While a uniform prior is a natural choice for binary decision-making tasks in which the states e.g., motion directions are a priori equally likely, it is less intuitive when learning expectations for rare sensory stimuli. Furthermore, computing counterfactual estimates of belief places additional computational burden on the learning system. Learning via sPE also avoids self-reinforcement and offers a biologically plausible alternative for self-supervised learning without invoking the computation of counterfactuals.

Another advantage of sPE-gated corticostriatal plasticity is that it enables post-synaptic neurons, i.e., neurons in the sensory striatum to develop action-independent sensory representations, allowing for flexible reuse of these signals in different contexts. Action-independent coding has indeed been reported recently in the sensory striatum of rodents (Guo et al., 2018) and is consistent with earlier studies indicating that optogenetic stimulation of neurons in this region biases perception rather than choice (L. Wang et al., 2018). Because striatal learning influences cortical belief representations in the model, the proposed mechanism also explains history- and context-dependent reorganization of belief signals in the associative cortex (Akrami et al., 2018; Noel et al., 2022; Lakshminarasimhan et al., 2023). While the model postulates sPE-based learning in sensory neurons in the striatum in general, we validated it using data from the rodent tail of striatum. Previous studies demonstrated that dopamine signal in this region is consistent with other types of prediction errors such as action prediction errors (Miller et al., 2019; Greenstreet et al., 2022; Lakshminarasimhan, 2024) and threat prediction errors (Menegas et al., 2018; Akiti et al., 2022). We note that these proposals are not mutually exclusive as they pertain to different task dimensions – sensory input, motor command, and outcome valence – and dopamine dynamics could multiplex these signals to enable learning along different dimensions. Functional heterogeneity within subregions of the tail striatum is also possible.

Additionally, recent works have revealed that dopamine responses in the ventral striatum are influenced by the internal model of the environment (Starkweather et al., 2017; Babayan et al., 2018; Krausz et al., 2023). But the extent to which dopamine responses underlie the learning of those internal models remains unclear. The proposed model suggests a plausible mechanism by which model-free, self-supervised updates in the sensory striatum facilitates learning of the internal model, which subsequently influences dopamine signaling for reward-based learning.

The modeling approach represents an integration of normative and mechanistic considerations and is particularly appealing from the perspective of preclinical translational research since it allows for testing

both granular predictions about neural dynamics and behavioral predictions using data across species. The proposed link between dopamine and sensory learning has major implications for understanding the neural basis of auditory hallucinations. First, we identified signatures of sensory prediction errors in the sensory striatum but not in the ventral striatum, a result that paralleled the significant interaction between hallucination proneness and stimulus-history but not reward-history in humans. This reconciles findings from PET imaging studies that localize elevated extracellular dopamine and synthesis capacity to the human dorsal striatum (Howes et al., 2009; Kegeles et al., 2010; Mizrahi et al., 2012) rather than ventral (limbic) striatum as anticipated by earlier theories of dopamine and psychosis (Kapur, 2003). Second, a specialized role for dopamine in this subregion vis-a-vis the non-region-specific antidopaminergic (D2-receptor blockade) effect of most antipsychotic drugs provides a potential explanation of their wide-ranging side effects in terms of off-target effects (Huhn et al., 2019). Lastly, our analysis of the time course of evolution of perceptual bias in response to dopamine stimulation was consistent with a plasticity-based mechanism. A role for sensory-striatal dopamine in sensory learning via a plasticity mechanism (rather than an instantaneous driving effect of dopamine on decisions) potentially explains the slow clinical response to antipsychotic medications (Maia & Frank, 2017) as well as the biological basis of strong priors identified by Bayesian models of hallucinations (Corlett et al., 2019; Duhamel et al., 2023). More generally, a key insight from our work is highlighting the relevance of SS corticostriatal plasticity as a more proximal driver of psychotic symptoms (hallucinations), with increased SS dopamine being a mediator of plasticity changes (which also depend on glutamatergic cortical pre-synaptic input and GABAergic striatal post-synaptic output captured by the three-factor plasticity rule in our model). This suggests that pharmacological agents directly targeting corticostriatal plasticity (via dopaminergic or non-dopaminergic pathways) could be explored as a potential treatment for psychosis. Understanding the interaction between dopaminergic, glutamatergic, and cholinergic systems in cortico-striatal synapses may provide a way to target this plasticity (Reynolds et al., 2022; Gallo et al., 2022; Chantranupong et al., 2023), which could also explain the efficacy of an emerging class of antipsychotic drugs based on muscarinic agonism (Mirza et al., 2003; Kaul et al., 2024).

SS dopamine may contribute to hallucinatory percepts through plasticity-based mechanisms while also playing a more direct role in modulating sensory expectations, as suggested by previous studies (Schmack et al., 2021). However, an exclusive role in encoding sensory expectations cannot account for the modulation of stimulus-evoked dopamine by stimulus history or the gradual impact of dopamine stimulation noted here. Likewise, while threat prediction errors and action prediction errors can gradually bias choices, they do not readily explain the characteristic perceptual qualities of hallucinations (e.g., why people with psychosis *hear* specific, well-defined sounds and voices) or the strong conviction of the associated beliefs. The proposed role for sensory striatal dopamine in learning sensory expectations thus provides a parsimonious link between excess dopamine and hallucinations.

4.2 Predictions and extensions

Our analyses of striatal dopamine revealed signatures of sensory prediction errors in the tail of the rodent striatum, but the model makes concrete predictions linking dopamine to the learning of sensory

expectations, that could be validated using signal-detection paradigms with changing signal probability. First, stimulus-induced SS dopamine levels should decrease as signal probability increases, since higher probability makes signals less surprising. Second, this decrease in dopamine should be accompanied by an increase in cortical activity within the dimension that encodes belief, driven by input from the basal ganglia prior to stimulus onset. Third, discrepancies between signal probability decoded from the initial cortical state and the true signal probability should correlate with deviations of the sensory striatal dopamine from the optimal sensory prediction error signal.

One challenge in testing the model predictions is the need to precisely target recordings to the corticostriatal circuit involved in learning and inference of the variable of interest. While the loop involving the primary auditory cortex is an appropriate choice in simple auditory signal-detection tasks, future extensions of the model should expand this approach to learn complex tasks using architectures that incorporate hierarchical processing of rich sensory inputs, informed by detailed wiring diagrams of the cortico-basal-ganglia loop (Foster et al., 2021). This will prove crucial to testing whether the pathophysiology of complex hallucinations in patients with psychosis is limited to dopamine signals that encode prediction errors about specific sensory features or if extends more broadly to semantic content. A related assumption was our treatment of dopamine as a scalar prediction-error signal. While this is justified and supported by empirical work in the reward domain, sensory inputs inherently comprise multiple feature dimensions, each of which requires a separate learning channel. More theoretical work is needed to explore the scalability of the proposed model to settings with naturalistic inputs using vector-valued sensory prediction errors similar to those used in reinforcement learning (Lee et al., 2024). Our emphasis on dopamine as a prediction error signal for learning does not preclude complementary roles that stem from sources of heterogeneity not considered here. Previous studies in the reward domain have proposed separate roles for both D1- and D2-expressing striatal neurons (Collins & Frank, 2014) and for tonic and phasic components of dopamine (Niv et al., 2007) in learning and instantaneous choice behavior. This spatiotemporal heterogeneity has recently been shown to have interesting consequences for learning due to differential dopamine affinities of D1 and D2 receptors (Pinto & Uchida, 2023). Incorporating such features into the proposed account of perceptual learning will be crucial for understanding the specific cell-type-specific mechanisms by which excess striatal dopamine leads to hallucinations and for developing safer treatments. Furthermore, incorporating known alterations in GABAergic/glutamatergic function within the proposed circuit model presents an opportunity to develop a unified computational theory of psychosis that integrates the dopamine hypotheses with the E/I imbalance suggested by predictive processing models (Keller & Sterzer, 2024).

5 Methods

5.1 Model

We developed a model of parallel corticostriatal loops for learning sensory and reward statistics via a local plasticity rule in corticostriatal synapses, modulated by heterogeneous dopamine signaling in the striatum.

Dynamics. The model comprised two parallel corticostriatal loops – sensory and limbic – to enable decision-making by respectively learning sensory and reward statistics. The sensory loop included the sensory cortex and the sensory portion of the dorsal striatum i.e., the sensory striatum (SS), while the limbic loop included the frontal cortex and the ventral striatum (VS). The sensory cortex was modeled as a recurrent neural network with N units which combined time-varying external sensory inputs with the output of the sensory striatum to track the momentary belief state i.e., the posterior probability of the underlying state of the world given the observations, while M units in the sensory striatum are themselves driven by those in the sensory cortex. The frontal cortex was also modeled as a recurrent neural network which combined the belief state with the output of the ventral striatum to output an action that maximized reward. Analogous to the sensory striatum, M ventral striatal units are driven by N units in the frontal cortex. The dynamics of units in cortical and striatal regions in both loops are therefore given by the same set of equations where subscript k is used to index the corticostriatal loop (sensory or limbic):

$$\tau \dot{\mathbf{x}}_k(t) = -\mathbf{x}_k(t) + \phi[J_k \mathbf{x}_k(t) + I_k(t) + y_k(t)] \quad , \tag{1a}$$

where t denotes time, τ is the intrinsic time-constant of each unit, nonlinearity $\phi(\cdot)$ is the tanh function, $\mathbf{x}_k \in \mathbb{R}^N$ denotes the population activity in cortical region indexed by k, J_k is the recurrent connectivity in that region, I_k is the scalar external input (sensory observations or posterior belief depending on the cortical region), and $y_k(t) = \sigma[\mathbf{1}^T \mathbf{z}_k(t)]$ denotes the summed output of the striatal units in region k passed through a sigmoidal nonlinearity. Activity of striatal units $\mathbf{z}_k \in \mathbb{R}^M$ is given by:

$$\mathbf{z}_k(t) = \phi[W_k(t)\mathbf{x}_k(t)] \quad , \tag{1b}$$

where $W_k \in \mathbb{R}^{M \times N}$ denotes the corticostriatal synaptic weights.

Neural signatures of evidence accumulation for perceptual and value computations have been found in associative and frontal cortical regions respectively (de Lafuente et al., 2015; Lin et al., 2020). Accordingly, recurrent connectivity J within the cortical regions are pre-trained to perform perceptual inference i.e., belief computation (sensory cortex) or value-based decision-making i.e., action selection (frontal cortex) via backpropagation-through-time and then held fixed after this initial pre-training. Concretely, recurrent

weights within the sensory cortex are pre-trained such that the summed activity encodes the posterior belief that the stimulus contains a signal i.e., to minimize the loss $\int \{\mathbf{1}^T \mathbf{x}_k(t) - b(t)\}^2 \, \mathrm{d}t$ where $b(t) = P(s = 1|I(1), ..., I(t))$ and s denotes the latent state ($s = 1$ for signal and $s = 0$ for noise). In contrast, recurrent weights within the frontal cortex are pre-trained such that the summed activity encodes the binary choice (yes/no) that maximizes reward i.e., to minimize the loss $\int \{\mathbf{1}^T \mathbf{x}_k(t) - a(t)\}^2 \, \mathrm{d}t$ where $a(t) = \arg\max_a P(r = 1|b(t_d), a) \;\; \forall \;\; t > t_d$ where t_d is the decision time. Critically, optimal performance depends on accurate estimation of sensory and reward expectations provided by the striatal output to the cortex i.e., $y_k()$, which need to be continually learned in non-stationary environments. To enable this, the elements of corticostriatal connectivity matrix W in both loops are initially random (gaussian, sampled from $\mathcal{N}(0, 1/N)$) but modified over time by local plasticity as described below. Since the objective is to model the effect of striatal learning on cortical computation, we are agnostic to the precise temporal structure of individual striatal units and thus ignore recurrent inhibition within the striatum.

Plasticity rule. We assume a standard three-factor learning rule (Łukasz Kuśmierz et al., 2017; Gerstner et al., 2018) in which teaching signals encoded by dopamine serve as a third factor to gate plasticity in corticostriatal synapses according to:

$$\dot{W}_k^{ij}(t) = \alpha \rho_k^{ij}(t) \delta_k(t) \quad , \tag{2}$$

where W_k^{ij} denotes the synaptic weight from cortical unit j to striatal unit i, ρ_k^{ij} denotes the eligibility trace of that synaptic weight estimated as the covariance between the pre-synaptic and the (derivate of) post-synaptic activity: $\tau \dot{\rho}_k^{ij}(t) = -\rho_k^{ij}(t) + \phi'[z_k^i(t)] x_k^j(t)$, δ_k serves as the teaching signal (the third factor) to gate plasticity at that synapse, and α denotes the learning rate. Note that the prediction error signals are global in that they are independent of the identity of pre- and post-synaptic neuron (indexed by superscripts i and j). However, they are assumed to be region-specific (indexed by the subscript k) in keeping with the observed heterogeneity in dopamine signaling across striatal regions.

Optimization. Standard neural network approaches optimize network parameters (e.g., synaptic weights W) to maximize accuracy. Because our objective here is to develop a model that can adapt to changes in task statistics (i.e., stimulus and reward statistics), we instead optimize the teaching signal δ_k such that the local plasticity rule in Equation 2 above can enable the network to learn either the stimulus or the reward statistics. Because we optimize the model for learning rather than performance, this approach corresponds to learning-to-learn or metalearning. Concretely, we expressed it as a spatiotemporal transformation of the activity of all units in the network:

$$\delta_k(t) = \sum_{\tau=1}^{T} H_k(\tau) A_k \mathbf{u}(t - \tau) \quad , \tag{3}$$

where $\mathbf{u} = (\mathbf{x}, \mathbf{z})^T$ denotes the concatenated P-dimensional activity of all units in the network where

$P = 2N + 2M$, $A_k \in \mathbb{R}^{P\times K}$ and $H_k \in \mathbb{R}^{K\times T}$ respectively denote a rank-K spatial filter and temporal filter where T is the length of the filter (number of timesteps). To determine the teaching signal for learning sensory expectations, we optimized the elements of A and H using gradient descent in a signal detection paradigm in which the fraction of signal trials was resampled from a standard uniform distribution $U(0, 1)$ every 20 trials for 10 blocks i.e., a total of 200 trials. To determine the teaching signal for learning reward expectations, we similarly optimized A and H but in a setting in which the fraction of trials with rewarded 'yes' responses was resampled from a standard uniform distribution $U(0, 1)$ every 20 trials. In both cases, optimization (i.e., meta-learning) via gradient-descent was performed in an outer loop while the inner learning loop was used to update corticostriatal synapses according to Equation 2. We found that a single dimension ($P = 1$) was sufficient to optimize learning of sensory expectations while learning reward expectations required a two dimensional subspace ($P = 2$), and the simulations in the main text correspond to these choices. The composition of the optimized teaching signal was characterized by estimating the alignment between the columns of A_k and the directions of neural activity $\mathbf{u}$ encoding the belief (b), value (v), reward (r), and action (a).

Teaching signals. In addition to using the teaching signals derived via the optimization approach described above, we simulated the model using idealized teaching signals for learning sensory and reward expectations. In these simulations, we used sensory prediction error (sPE) and reward prediction error (rPE) as teaching signals for learning sensory expectation and reward expectation in the SS and VS respectively. sPE was defined as:

$$\delta_{sPE}(t) = b(t) - b(t-1) \quad , \tag{4}$$

where $b(t) = \mathbf{1}^T \mathbf{x}_k(t)$ denotes the posterior belief encoded in the summed activity of the model sensory cortex as described earlier. Note that the teaching signal defined in this manner corresponds to the derivative of the subjective belief. rPE was defined according to the standard temporal difference learning rule:

$$\delta_{rPE}(t) = r(t) + v(t) - v(t-1) \quad , \tag{5}$$

where $v(t)$ denotes the expected reward associated with choosing 'yes' response in the current belief state. $r(t) = \delta_{ay} r_y + \delta_{an} r_n$ for $t = T$ denotes the reward feedback where r_y and r_n denote the reward when action a corresponds to correct yes (y) and correct no (n) choice respectively.

Simulations. The cortical modules each comprise N=128 units while the striatal modules comprise M=16 units in keeping with convergent corticostriatal projections in the mammalian brain (Kincaid, Zheng, & Wilson, 1998). Simulations were performed using timesteps dt = 10ms, the intrinsic time constant of individual units was set to $\tau = 50$ms (5 timesteps), and learning rate $\alpha = 0.1$.

For simulating the associative learning task, the odor and reward were both modeled as brief 100ms pulses, spaced 1s apart. Since reward delivery coincided with the click of a water valve (Menegas et al., 2018), the sensory cortex received two sensory inputs in sequence. To qualitatively match the empirical findings, we initialized the sensory expectation for the water-valve click to be slightly higher than the odor, resulting in weaker sensory prediction errors. This assumption is justified since animals are likely to be more familiar with the former due to prior exposure to the experimental setup.

For simulating the signal detection task, the stimuli were 500ms long and comprised 10 samples of 50ms each. Samples were drawn from a gaussian with mean $\mu \in \{0, 0.1\}$ and variance $\sigma \in \{0.1, 0.2, 0.3\}$ where mean and variance controlled the trial type (signal or noise) and signal to noise ratio respectively. For simulations used to test the model's performance on signal detection, we used two blocks of 500 trials each. In the first block, the fraction of signal trials was 0.5 and rewards were delivered with a probability of 0.5 both for correct 'yes' (true positive) and correct 'no' (true negative) responses. In the subsequent block, we either increased the fraction of signal trials to 0.75 (signal manipulation) or the fraction of rewarded 'yes' trials to 0.75 (reward manipulation).

5.2 Mouse Behavioral Tasks

Mouse Associative Learning. As described in greater detail in the original work (Menegas et al., 2017), bulk calcium signals were recorded from dopamine axons projecting to the ventral striatum (N=9) or tail striatum (N=11) while mice learned associations between an odor and an outcome. Importantly, the mice had no prior experience with the odor. Each trial, an odor was presented followed by a one second delay and then either a water reward (90% of trials) or nothing (10% of trials).

Mouse Auditory Signal Detection. As described in greater detail in the original work (Schmack et al., 2021), on each trial water-restricted mice reported their auditory percept (either a 1 kHz tone or white noise) by pressing a lever. Tone trials were presented on 50% of trials. Correct decisions were rewarded with water after a random delay period. Trials were self-initiated, which created tension between waiting for a reward and initiating another trial. This willingness to wait for reward has been previously validated as a measure of statistical decision confidence (Hangya et al., 2016).

5.3 Human Data Collection

Sample Information. All participants provided informed consent and all study procedures were approved by local Institutional Review Boards. For both samples, participants were recruited through Prolific and the experiment was hosted on Pavlovia.org (Peirce et al., 2019).

Participant Pre-screening and Retained Sample The distribution of subclinical hallucination propensity is heavily skewed in the general population, so participants with high hallucination propensity

are poorly represented in random samples. To ensure a wide range of variability of hallucination propensity in our samples, we used a pre-screening procedure following our previous work (Ashinoff, Buck, Woodford, & Horga, 2022). We first collected large initial samples (N=499 and N=494 for Sample 1 and Sample 2, respectively) which completed the Cardiff Anomalous Perception Scale (CAPS) and the Beck Depression Inventory (BDI). Then, using previous published norms (Schmack et al., 2021) we derived unbiased CAPS-score cutoffs to determine three groups: low, medium, and high (under 30.1 for low and over 85.7 for high). All participants from the high CAPS group and matched participants in the low and medium CAPS groups (according to their BDI score) were invited to complete the behavioral task. 152 and 148 participants completed the full experiment in Sample 1 and Sample 2, respectively.

Human Auditory Signal Detection. To maximize comparisons with the mouse data, core features of the human task were matched to the mouse task. White noise was constantly played in the background and on some trials, a pure 1 kHz tone was presented. Detection difficulty was standardized across participants by using a separate auditory thresholding procedure (Powers et al., 2017) to determine the volume at which the tone was detected on 25%, 50% and 75% of presentations. To further standardize the delivery of auditory stimuli, participants were required to wear headphones which was enforced using a headphone check procedure based on prior work (Huggins Pitch and Binaural Beats, (Milne et al., 2021)).

To evaluate the influence of sensory and reward learning on high-confidence false alarms, two block-wise manipulations were introduced in the human task (**Figure 5A** – right): 1) the probability of a tone being presented (i.e., signal probability), and 2) the probability of receiving additional reward for a signal decision (i.e., reward probability). Since using these contingencies maximized overall reward receipt, participants were incentivized to engage in both sensory and reward learning and incorporate their learned expectations into their choices and confidence reports. Importantly, we selected block probabilities to minimize the correlation between perceptual correctness and reward magnitude which allowed us to infer distinct effects of both signal and reward probability on behavior. The task consisted of 12 blocks of 20 trials each. There were three categories of blocks: signal, reward, and hybrid manipulation blocks. In signal blocks, reward probability is fixed at 50% and signal probability is either 10%, 30%, 70%, or 90% (4 blocks total). In reward blocks, signal probability is fixed at 50% and reward probability either 10%, 30%, 70%, or 90% (4 blocks total). In hybrid blocks, signal and reward probability are either 20% or 80% (all combinations result in 4 blocks total).

To further separate the influence of sensory and reward contributions to choice, participants also reported how certain they were that a tone was present (i.e., signal confidence) on a visual analog scale. At the end of each trial, participants were told whether the tone was indeed present or absent and whether their choice resulted in an additional reward (**Figure 5A** – left).

Questionnaires and Additional Cognitive Tasks In the same session as the auditory signal detection task, participants repeated the CAPS and BDI and also completed the Beck Anxiety Inventory, RAVENS progressive matrices, and the forward and backward digit span.

Data Quality To ensure task comprehension, participants were required to complete several practice sessions and a quiz before completing the signal detection task. For participants who completed the full task, we also used self-report and task-based screening measures to promote high data quality (Zorowitz, Solis, Niv, & Bennett, 2023). Each self-report questionnaire included a catch question which should elicit a standard response from attentive participants. Participants were excluded if two or more responses to catch questions were incorrect (N=1 and N=1 for Sample 1 and Sample 2, respectively). To promote the reliability of hallucination propensity measures, participants were also excluded for improbably large changes from the first to second administration of CAPS (changes in group assignment based on normed cutoffs from low to high or high to low; N=2 and N=1 for Sample 1 and Sample 2, respectively). Finally, participants were excluded for suspicious median response times (<200 ms or >3s) during the signal detection task (N=4 and N=2 for Sample 1 and Sample 2, respectively). With these exclusions, the final sample sizes used for analyses were N=146 and N=144 for Sample 1 and Sample 2, respectively.

5.4 Data Analysis

In general, we use mixed-effects regression models to evaluate the influence of variables of interest on behavior or neural data at the group-level: $Y_{i,j} = \beta_0 + \beta_1 X_{ij} + u_{0j} + \epsilon_{ij}$ where the outcome Y for subject j on trial i is the determined by the predictor X_{ij}, a group-level intercept β_0, a group level slope β_1, and subject-specific random intercepts u_{0j}. When modeling choice, we use mixed-effects logistic regression, while for confidence and neural activity, we use mixed-effects linear regression.

Mouse Fiber Photometry Analysis (Data from Schmack et al., 2021). A prediction error should be composed of a positive response to an outcome and a dissociable negative response that scales with the expectation of the outcome (Rutledge, Dean, Caplin, & Glimcher, 2010). So, to test whether dopamine release in the striatum was consistent with a sensory prediction error, we evaluated whether these signals were sensitive to both stimulus signal-to-noise ratio and stimulus history (an index of sensory expectation) independent of reward and choice history. Specifically, we predict dopamine signal in the sensory and ventral striatum as a function of SNR, previous trial stimulus, previous trial reward, and previous trial choice. For quantification, we predicted average dopamine signal in a 1 second window following stimulus presentation. For exploratory visualization, we used a moving window regression with 500 ms windows incremented by 250 ms.

Optogenetic stimulation (Data from Schmack et al., 2021). (Schmack et al., 2021) reported that optogenetic stimulation of dopamine release in the tail striatum increased false alarm rate and confidence. To assess more granular temporal dynamics of this effect, we predicted choice and confidence as a function of experiment progress (i.e., session number), session progress (i.e., block number), and block progress (i.e., trial number within a block). Since the number of sessions varied across animals and the number of blocks varied across animals and sessions, we normalized predictors by taking median splits for each animal. In other words, we considered whether the effect of dopamine stimulation changed from the first

to the second half of sessions, blocks, or trials.

Human Learning. Participants are incentivized to learn the sensory and reward contingencies of the task to develop expectations about the sensory environment and which decisions yield additional reward. If participants are incorporating their learned expectations into their behavior, recent sensory or reward evidence should measurably bias choice and confidence. To quantify this effect, we predicted current trial choice and confidence as a function of SNR, previous trial sensory feedback, and previous trial reward outcome. To evaluate whether learning effects are moderated by hallucination proneness, we also tested a model that included interactions of each predictor with total scores on the Cardiff Anomalous Perceptions Scale (CAPS) into the mixed-effects regression models.

Choices are influenced by perceptual beliefs and the mapping between these beliefs and actions (i.e., policy). Since confidence reports are a readout of perceptual beliefs we can estimate participants' policy by evaluating how our reward manipulation influences choice after accounting for the effect of confidence. To this end, we predicted choice as a function of SNR, previous trial sensory feedback, previous trial reward outcome, and current trial confidence. Since confidence is a separate predictor in this model, we interpret any influence of reward history on choice as an effect of a change in policy.

6 Acknowledgements

The authors thank Arturo Torres and Katharina Schmack for helpful discussions. This work was supported by NIH grants R01MH136672 to G.H. and C.K. and F31MH134617 to J.B., the Kavli Foundation, the Gatsby Charitable Foundation GAT3780, and a NARSAD Young Investigator Grant #31985 from the Brain & Behavior Research Foundation to K.L.

References

Abi-Dargham, A., Gil, R., Krystal, J., Baldwin, R. M., Seibyl, J. P., Bowers, M., ... Laruelle, M. (1998). Increased Striatal Dopamine Transmission in Schizophrenia: Confirmation in a Second Cohort. *American Journal of Psychiatry*, *155*(6), 761–767. doi: 10.1176/ajp.155.6.761

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*. doi: 10.3389/fpsyt.2013.00047

Akiti, K., Tsutsui-Kimura, I., Xie, Y., Mathis, A., Markowitz, J. E., Anyoha, R., ... Watabe-Uchida, M. (2022). Striatal dopamine explains novelty-induced behavioral dynamics and individual variability in threat prediction. *Neuron*, *110*. doi: 10.1016/j.neuron.2022.08.022

Akrami, A., Kopec, C. D., Diamond, M. E., & Brody, C. D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature*, *554*. doi: 10.1038/nature25510

Alderson-Day, B., Moffatt, J., Lima, C. F., Krishnan, S., Fernyhough, C., Scott, S. K., ... Evans, S. (2022). Susceptibility to auditory hallucinations is associated with spontaneous but not directed modulation of top-down expectations for speech. *Neuroscience of Consciousness*, *2022*. doi: 10.1093/nc/niac002

Ashinoff, B. K., Buck, J., Woodford, M., & Horga, G. (2022). The effects of base rate neglect on sequential belief updating and real-world beliefs. *PLOS Computational Biology*, *18*(12), e1010796. (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1010796

Babayan, B. M., Uchida, N., & Gershman, S. J. (2018). Belief state representation in the dopamine system. *Nature Communications*, *9*. doi: 10.1038/s41467-018-04397-0

Cassidy, C. M., Balsam, P. D., Weinstein, J. J., Rosengard, R. J., Slifstein, M., Daw, N. D., ... Horga, G. (2018). A perceptual inference mechanism for hallucinations linked to striatal dopamine. *Current Biology*. doi: 10.1016/j.cub.2017.12.059

Cassidy, C. M., Zucca, F. A., Girgis, R. R., Baker, S. C., Weinstein, J. J., Sharp, M. E., ... Horga, G. (2019). Neuromelanin-sensitive mri as a noninvasive proxy measure of dopamine function in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *116*. doi: 10.1073/pnas.1807983116

Chantranupong, L., Beron, C. C., Zimmer, J. A., Wen, M. J., Wang, W., & Sabatini, B. L. (2023). Dopamine and glutamate regulate striatal acetylcholine in decision-making. *Nature*, *621*(7979), 577–585. doi: 10.1038/s41586-023-06492-9

Chen, A., Chen, L., Kim, T. A., & Xiong, Q. (2021). Integrating the roles of midbrain dopamine circuits in behavior and neuropsychiatric disease. *Biomedicines*, *9*, 647. doi: 10.3390/biomedicines9060647

Chen, A., Malgady, J., Chen, L., Shi, K., Cheng, E., Plotkin, J., ... Xiong, Q. (2022). Nigrostriatal dopamine pathway regulates auditory discrimination behavior. *Nature Communications*, *13*.

Collins, A. G., & Frank, M. J. (2014). Opponent actor learning (opal): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review*, *121*. doi: 10.1037/a0037015

Corlett, P. R., Horga, G., Fletcher, P. C., Alderson-Day, B., Schmack, K., & Powers, A. R. (2019). *Hallucinations and strong priors* (Vol. 23). doi: 10.1016/j.tics.2018.12.001

Curran, C., Byrappa, N., & McBride, A. (2004). Stimulant psychosis: Systematic review. *British Journal*

of Psychiatry, *185*. doi: 10.1192/bjp.185.3.196

de Lafuente, V., Jazayeri, M., & Shadlen, M. N. (2015). Representation of accumulating evidence for a decision in two parietal areas. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *35*, 4306-18.

Dosher, B., & Lu, Z. L. (2017). *Visual perceptual learning and models* (Vol. 3). doi: 10.1146/annurev -vision-102016-061249

Duhamel, E., Mihali, A., & Horga, G. (2023). Effects of hallucination proneness and sensory resolution on prior biases in human perceptual inference of time intervals. *Journal of Neuroscience*, *43*. doi: 10.1523/JNEUROSCI.0692-22.2023

Fleming, G., Wright, B. A., & Wilson, D. A. (2019). The value of homework: Exposure to odors in the home cage enhances odor-discrimination learning in mice. *Chemical Senses*, *44*. doi: 10.1093/ chemse/bjy083

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*. doi: 10.1038/nrn2536

Foster, N. N., Barry, J., Korobkova, L., Garcia, L., Gao, L., Becerra, M., ... Dong, H. W. (2021). The mouse cortico–basal ganglia–thalamic network. *Nature*, *598*. doi: 10.1038/s41586-021-03993-3

Friston, K. J. (2005). Hallucinations and perceptual inference. *Behavioral and Brain Sciences*, *28*. doi: 10.1017/S0140525X05290131

Gallo, E. F., Greenwald, J., Yeisley, J., Teboul, E., Martyniuk, K. M., Villarin, J. M., ... Kellendonk, C. (2022). Dopamine d2 receptors modulate the cholinergic pause and inhibitory learning. *Molecular Psychiatry*, *27*. doi: 10.1038/s41380-021-01364-y

Gardner, M. P. H., Schoenbaum, G., & Gershman, S. J. (2018). Rethinking dopamine as generalized prediction error. *Proceedings of the Royal Society B: Biological Sciences*, *285*, 20181645. doi: 10.1098/rspb.2018.1645

Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D., & Brea, J. (2018). Eligibility traces and plasticity on behavioral time scales: Experimental support of neohebbian three-factor learning rules. *Frontiers in Neural Circuits*, *12*. doi: 10.3389/fncir.2018.00053

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*, 585-595. doi: 10.1016/j.neuron.2010.04.016

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., ... Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, *134*. doi: 10.1093/brain/awr059

Greenstreet, F., Vergara, H. M., Pati, S., Schwarz, L., Wisdom, M., Marbach, F., ... Stephenson-Jones, M. (2022). Action prediction error: a value-free dopaminergic teaching signal that drives stable learning. *bioRxiv*, 2022.09.12.507572. doi: 10.1101/2022.09.12.507572

Guo, L., Walker, W. I., Ponvert, N. D., Penix, P. L., & Jaramillo, S. (2018). Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nature Communications*, *9*. doi: 10.1038/s41467-018-03994-3

Hangya, B., Sanders, J. I., & Kepecs, A. (2016). A Mathematical Framework for Statistical Decision Confidence. *Neural Computation*, *28*(9), 1840–1858. doi: 10.1162/NECO_a_00864

Heinz, A., & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: Salience attribution revisited. *Schizophrenia Bulletin*, *36*. doi: 10.1093/schbul/sbq031

Hennig, J. A., Pinto, S. A., Yamaguchi, T., Linderman, S. W., Uchida, N., & Gershman, S. J. (2023). Emergence of belief-like representations through reinforcement learning. *PLoS Computational Biology*, *19*. doi: 10.1371/journal.pcbi.1011067

Horga, G., & Abi-Dargham, A. (2019). An integrative framework for perceptual disturbances in psychosis. *Nature Reviews Neuroscience*, *20*. doi: 10.1038/s41583-019-0234-1

Hospedales, T., Antoniou, A., Micaelli, P., & Storkey, A. (2022). Meta-learning in neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*. doi: 10.1109/TPAMI.2021.3079209

Howes, O. D., Montgomery, A. J., Asselin, M. C., Murray, R. M., Valli, I., Tabraham, P., ... Grasby, P. M. (2009). Elevated striatal dopamine function linked to prodromal signs of schizophenia. *Archives of General Psychiatry*, *66*. doi: 10.1001/archgenpsychiatry.2008.514

Huhn, M., Nikolakopoulou, A., Schneider-Thoma, J., Krause, M., Samara, M., Peter, N., ... Leucht, S. (2019). Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. *The Lancet*, *394*. doi: 10.1016/S0140-6736(19)31135-3

Jardri, R., & Denève, S. (2013). Circular inferences in schizophrenia. *Brain*, *136*, 3227-3241. doi: 10.1093/brain/awt257

Jiang, L., & Litwin-Kumar, A. (2021). Models of heterogeneous dopamine signaling in an insect learning and memory center. *PLoS Computational Biology*, *17*. doi: 10.1371/journal.pcbi.1009205

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*. doi: 10.1176/appi.ajp.160.1.13

Kaul, I., Sawchak, S., Correll, C. U., Kakar, R., Breier, A., Zhu, H., ... Brannan, S. K. (2024). Efficacy and safety of the muscarinic receptor agonist karxt (xanomeline–trospium) in schizophrenia (emergent-2) in the usa: results from a randomised, double-blind, placebo-controlled, flexible-dose phase 3 trial. *The Lancet*, *403*. doi: 10.1016/S0140-6736(23)02190-6

Kegeles, L. S., Abi-Dargham, A., Frankle, W. G., Gil, R., Cooper, T. B., Slifstein, M., ... Laruelle, M. (2010). Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. *Archives of General Psychiatry*, *67*. doi: 10.1001/archgenpsychiatry.2010.10

Kehrer, C. (2008). Altered excitatory-inhibitory balance in the nmda-hypofunction model of schizophrenia. *Frontiers in Molecular Neuroscience*, *1*. doi: 10.3389/neuro.02.006.2008

Keller, G. B., & Sterzer, P. (2024). Predictive processing: A circuit approach to psychosis. *Annual Review of Neuroscience*, *47*. doi: 10.1146/annurev-neuro-100223-121214

Kesby, J. P., Eyles, D. W., McGrath, J. J., & Scott, J. G. (2018). Dopamine, psychosis and schizophrenia: The widening gap between basic and clinical neuroscience. *Translational Psychiatry*. doi: 10.1038/s41398-017-0071-9

Kincaid, A. E., Zheng, T., & Wilson, C. J. (1998). Connectivity and convergence of single corticostriatal axons. *Journal of Neuroscience*, *18*. doi: 10.1523/jneurosci.18-12-04722.1998

Krausz, T. A., Comrie, A. E., Kahn, A. E., Frank, L. M., Daw, N. D., & Berke, J. D. (2023). Dual credit

assignment processes underlie dopamine signals in a complex spatial environment. *Neuron*, *111*. doi: 10.1016/j.neuron.2023.07.017

Lakshminarasimhan, K. (2024). A computational principle of habit formation. *bioRxiv*. doi: 10.1101/2024.10.12.618033

Lakshminarasimhan, K., Avila, E., Pitkow, X., & Angelaki, D. E. (2023). Dynamical latent state computation in the male macaque posterior parietal cortex. *Nature Communications*, *14*, 1832. doi: 10.1038/s41467-023-37400-4

Lakshminarasimhan, K., Xie, M., Cohen, J. D., Sauerbrei, B. A., Hantman, A. W., Litwin-Kumar, A., & Escola, S. (2024). Specific connectivity optimizes learning in thalamocortical loops. *Cell Reports*, *43*, 114059. doi: https://doi.org/10.1016/j.celrep.2024.114059

Lally, J., & MacCabe, J. H. (2015). Antipsychotic medication in schizophrenia: A review. *British Medical Bulletin*, *114*. doi: 10.1093/bmb/ldv017

Laruelle, M., Abi-Dargham, A., Dyck, C. H. V., Gil, R., D'Souza, C. D., Erdos, J., ... Innis, R. B. (1996). Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *Proceedings of the National Academy of Sciences of the United States of America*, *93*. doi: 10.1073/pnas.93.17.9235

Lee, R. S., Sagiv, Y., Engelhard, B., Witten, I. B., & Daw, N. D. (2024). A feature-specific prediction error model explains dopaminergic heterogeneity. *Nature Neuroscience*, *27*, 1574-1586. doi: 10.1038/s41593-024-01689-1

Lin, Z., Nie, C., Zhang, Y., Chen, Y., & Yang, T. (2020). Evidence accumulation for value computation in the prefrontal cortex during decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *117*. doi: 10.1073/pnas.2019077117

Loewenstein, Y., Raviv, O., & Ahissar, M. (2021). Dissecting the roles of supervised and unsupervised learning in perceptual discrimination judgments. *Journal of Neuroscience*, *41*. doi: 10.1523/JNEUROSCI.0757-20.2020

Maia, T. V., & Frank, M. J. (2017). An integrative perspective on the role of dopamine in schizophrenia. *Biological Psychiatry*, *81*, 52-66. doi: 10.1016/j.biopsych.2016.05.021

McCutcheon, R., Beck, K., Jauhar, S., & Howes, O. D. (2018). Defining the locus of dopaminergic dysfunction in schizophrenia: A meta-analysis and test of the mesolimbic hypothesis. *Schizophrenia Bulletin*, *44*. doi: 10.1093/schbul/sbx180

McKetin, R., Lubman, D. I., Baker, A. L., Dawe, S., & Ali, R. L. (2013). Dose-related psychotic symptoms in chronic methamphetamine users: evidence from a prospective longitudinal study. *JAMA psychiatry*, *70*(3), 319–324. doi: 10.1001/jamapsychiatry.2013.283

Menegas, W., Akiti, K., Amo, R., Uchida, N., & Watabe-Uchida, M. (2018). Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nature Neuroscience*, *21*, 1421-1430. doi: 10.1038/s41593-018-0222-1

Menegas, W., Babayan, B. M., Uchida, N., & Watabe-Uchida, M. (2017). Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *eLife*, *6*, e21886. doi: 10.7554/eLife.21886.001

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, *126*. doi: 10.1037/rev0000120

Milne, A. E., Bianco, R., Poole, K. C., Zhao, S., Oxenham, A. J., Billig, A. J., & Chait, M. (2021). An online headphone screening test based on dichotic pitch. *Behavior Research Methods*, *53*(4), 1551–1562. doi: 10.3758/s13428-020-01514-0

Mirza, N. R., Peters, D., & Sparks, R. G. (2003). Xanomeline and the antipsychotic potential of muscarinic receptor subtype selective agonists. *CNS Drug Reviews*, *9*. doi: 10.1111/j.1527-3458.2003.tb00247.x

Mizrahi, R., Addington, J., Rusjan, P. M., Suridjan, I., Ng, A., Boileau, I., ... Wilson, A. A. (2012). Increased stress-induced dopamine release in psychosis. *Biological Psychiatry*, *71*. doi: 10.1016/j.biopsych.2011.10.009

Moseley, P., Smailes, D., Ellison, A., & Fernyhough, C. (2016). The effect of auditory verbal imagery on signal detection in hallucination-prone individuals. *Cognition*, *146*. doi: 10.1016/j.cognition.2015.09.015

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

Noel, J.-P., Balzani, E., Avila, E., Lakshminarasimhan, K. J., Bruni, S., Alefantis, P., ... Angelaki, D. E. (2022). Coding of latent variables in sensory, parietal, and frontal cortices during closed-loop virtual navigation. *eLife*, *11*, 1-25. doi: 10.7554/elife.80280

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. doi: 10.3758/s13428-018-01193-y

Peters, A. J., Fabre, J. M., Steinmetz, N. A., Harris, K. D., & Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature*, *591*. doi: 10.1038/s41586-020-03166-8

Petrov, A. A., Dosher, B. A., & Lu, Z. L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, *112*. doi: 10.1037/0033-295X.112.4.715

Petrov, A. A., Dosher, B. A., & Lu, Z. L. (2006). Perceptual learning without feedback in non-stationary contexts: Data and model. *Vision Research*, *46*. doi: 10.1016/j.visres.2006.03.022

Pinto, S. R., & Uchida, N. (2023). Tonic dopamine and biases in value learning linked through a biologically inspired reinforcement learning model. *bioRxiv*, 2023.11.10.566580. doi: 10.1101/2023.11.10.566580

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science*, *357*. doi: 10.1126/science.aan3458

Rao, R. P. (2010). Decision making under uncertainty: A neural model based on partially observable markov decision processes. *Frontiers in Computational Neuroscience*, *4*. doi: 10.3389/fncom.2010.00146

Reynolds, J. N., Avvisati, R., Dodson, P. D., Fisher, S. D., Oswald, M. J., Wickens, J. R., & Zhang, Y. F. (2022). Coincidence of cholinergic pauses, dopaminergic activation and depolarisation of spiny projection neurons drives synaptic plasticity in the striatum. *Nature Communications*, *13*. doi: 10.1038/s41467-022-28950-0

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *Journal of Neuroscience*, *30*(40), 13525–13536. doi: 10.1523/JNEUROSCI.1747-10.2010

Schmack, K., Bosc, M., Ott, T., Sturgill, J. F., & Kepecs, A. (2021). Striatal dopamine mediates

hallucination-like perception in mice. *Science*. doi: 10.1126/science.abf4740

Schmid, C., Haziq, M., Baese-Berk, M. M., Murray, J. M., & Jaramillo, S. (2024). Passive exposure to task-relevant stimuli enhances categorization learning. *eLife*, *13*. doi: 10.7554/eLife.88406

Seeman, P., & Kapur, S. (2000). Schizophrenia: More dopamine, more d2 receptors. *Proceedings of the National Academy of Sciences of the United States of America*, *97*. doi: 10.1073/pnas.97.14.7673

Sohn, H., & Jazayeri, M. (2021). Validating model-based bayesian integration using prior–cost metamers. *Proceedings of the National Academy of Sciences of the United States of America*, *118*. doi: 10.1073/pnas.2021531118

Sommer, I. E., Slotema, C. W., Daskalakis, Z. J., Derks, E. M., Blom, J. D., & Gaag, M. V. D. (2012). The treatment of hallucinations in schizophrenia spectrum disorders. *Schizophrenia Bulletin*, *38*. doi: 10.1093/schbul/sbs034

Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature Neuroscience*, *20*. doi: 10.1038/nn.4520

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., ... Corlett, P. R. (2018, 11). The predictive coding account of psychosis. *Biological Psychiatry*, *84*, 634-643. doi: 10.1016/j.biopsych.2018.05.015

Stevens, J. R. (1973). An anatomy of schizophrenia? *Archives of General Psychiatry*, *29*. doi: 10.1001/archpsyc.1973.04200020023003

Walter, H., Kammerer, H., Frasch, K., Spitzer, M., & Abler, B. (2009). Altered reward functions in patients on atypical antipsychotic medication in line with the revised dopamine hypothesis of schizophrenia. *Psychopharmacology*, *206*. doi: 10.1007/s00213-009-1586-4

Wang, Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., ... Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*. doi: 10.1038/s41593-018-0147-8

Wang, L., Rangarajan, K. V., Gerfen, C. R., & Krauzlis, R. J. (2018). Activation of striatal neurons causes a perceptual decision bias during visual change detection in mice. *Neuron*, *97*, 1369-1381.e5. doi: 10.1016/j.neuron.2018.01.049

Wang, X.-J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, *84*, 638-654. doi: https://doi.org/10.1016/j.neuron.2014.10.018

Xiong, Q., Znamenskiy, P., & Zador, A. M. (2015). Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature*, *521*, 348-351. doi: 10.1038/nature14225

Yun, S., Yang, B., Anair, J. D., Martin, M. M., Fleps, S. W., Pamukcu, A., ... Parker, J. G. (2023). Antipsychotic drug efficacy correlates with the modulation of D1 rather than D2 receptor-expressing striatal projection neurons. *Nature Neuroscience*, *26*(8), 1417–1428. (Publisher: Nature Publishing Group) doi: 10.1038/s41593-023-01390-9

Zorowitz, S., Solis, J., Niv, Y., & Bennett, D. (2023). Inattentive responding can induce spurious associations between task behaviour and symptom measures. *Nature Human Behaviour*, *7*(10), 1667–1681. (Publisher: Nature Publishing Group) doi: 10.1038/s41562-023-01640-7

Zylberberg, A., Wolpert, D. M., & Shadlen, M. N. (2018). Counterfactual reasoning underlies the learning of priors in decision making. *Neuron*, *99*, 1083-1097.e6. doi: https://doi.org/10.1016/j.neuron.2018.07.035

Łukasz Kuśmierz, Isomura, T., & Toyoizumi, T. (2017). Learning with three factors: modulating hebbian plasticity with errors. *Current Opinion in Neurobiology*, *46*. doi: 10.1016/j.conb.2017.08.020