Published in final edited form as:
Am J Psychiatry. 2025 September 01; 182(9): 819–829. doi:10.1176/appi.ajp.20240556.

20 years of aberrant salience in psychosis: What have we learned?

Philip R. Corlett, Ph.D.[1,*], **Kurt M. Fraser, Ph.D.**[2]

[1] Wu Tsai Neuroscience Institute, Departments of Psychiatry & Psychology Yale University, New Haven, CT 06520

[2] Department of Psychology, University of Minnesota, Minneapolis, MN 55455

Abstract

Twenty years ago Shitij Kapur's "***Psychosis as a state of aberrant salience***" captured the attention of clinicians, cognitive, and behavioral neuroscientists. It has become the *de facto* way of talking about delusion formation in labs and clinics. Here, evidence for this theory is critically evaluated in consideration of evolving data since its publication. A particular focus is placed on its specific predictions, regarding the neural and behavioral loci of dopamine dysfunction in psychosis and find them lacking. This examination is informed by recent advances in understanding of the function of the dopamine system and its impacts on behavior following the explosion of new tools and probes for precise measurement and manipulation of dopaminergic circuits. Contemporary theories that have developed since Kapur, which suggest a role for dopamine in belief formation, belief updating under uncertainty, and abductive inference to the best explanation for some set of circumstances – are argued to form a more cogent theory that fits better with the work in patients with delusions and hallucinations, how they behave and what is known about the function of their dopamine system. The original salience hypothesis has been influential as it attempted to unite neurochemical dysfunction to clinical phenomenology through computational cognitive neuroscience, which has led to the development of novel predictions which the authors highlight as future directions for the field.

Keywords

Psychosis; Hallucinations; Delusions; Dopamine; Salience; Motivation; Incentive; Prediction Error

"***Psychosis as a state of aberrant salience***" was published 20 years ago in the *American Journal of Psychiatry* (1). In the intervening years it has been cited more than 3000 times. It is a work of brilliance – elegantly weaving connections between clinical phenomenology, psychopharmacology, and behavioral neuroscience, via a concept from addiction – incentive salience (2). It aims to answer the question, oft posed by patients:

---

[*] Corresponding author: Philip R. Corlett, PhD. philip.corlett@yale.edu.

“Why, when I take antipsychotic drugs, do my delusions and hallucinations go away?”

The piece provides a perspicuous answer: Psychosis (hallucinations and delusions) is caused by excessive dopamine levels, which cause events, either internal or external, that are surreptitiously coincident with a dopamine peak to be imbued with aberrant incentive salience – such that when re-encountered they will command attention and drive goal-directed behavior. These experiences manifest ultimately as psychotic symptoms; attempts to explain away the aberrant salience (1). When treated with drugs which block dopamine receptors, the aberrant salience abates, and symptoms resolve (1).

Here we review the work that has gone on considering the theory, how it has been received, and what our field continues to get wrong about aberrant salience, in light of our evolving understanding of dopamine function, incentive salience attribution, and clinical computational neuroscience.

What predictions does the theory offer, and how have they fared under empirical scrutiny? We think there are 2 critical predictions made by the theory:

1. Positive symptoms (largely delusions) should relate to the neurobehavioral ascription of incentive salience
2. The dopamine abnormality in psychosis should be in the ventral striatum

What have we learned in the intervening years?

1. Delusions should relate to the neurobehavioral ascription of incentive salience

Throughout this piece we will mostly focus on incentive salience attribution and delusions. Kapur offered a briefer explanation of hallucinations as the attribution of incentive salience to one’s own inner speech, such that it becomes so salient it must have emanated from another agent (1). We shall return to hallucinations later in our piece. Kapur’s ideas about delusions and incentive salience attribution have received more empirical attention.

Consider the Salience Attribution Task, a probabilistic reward-learning task featuring compound cue stimuli that vary along two dimensions, one task-relevant and one task-irrelevant. Performance in this task can be described as ‘adaptive’ or ‘maladaptive’ reward learning depending on whether one exhibits greater differences along task-relevant or -irrelevant dimensions as assessed by differences in explicit measures (such as liking ratings of, and reward expectations following, task stimuli) and implicitly assessed by reaction times (RTs). If salience processing deficits are thought to be related to aberrant attribution of incentive salience then one would expect that patients with delusions should exhibit more ‘maladaptive’ learning of task-irrelevant features than control subjects. This is not the case (3). There are no differences in aberrant salience between patients with delusions and healthy controls (3). However, there was a correlation between aberrant salience attribution and delusions in patients. It is hard to conclude, based on these data, that aberrant incentive salience causes delusions – because the people with delusions should have more incentive salience misattribution than the people without. People at high risk for psychosis have aberrant salience attribution, and this is associated with a dearth of nucleus accumbens activation during the task (4). However, these differences do not persist in follow-up, and have not, as far as we are aware, been predictive of conversion to psychosis and delusions. Indeed, in treatment refractory patients (with delusions) there was no evidence of excessive attribution of salience to the irrelevant cues in people with refractory delusions (5). At best, these data suggest a role for impaired incentive salience attribution during the delusional mood preludial to psychosis, but the data do not suggest the particular link to delusions that Kapur posited.

The monetary incentive delay task was initially developed to examine responses to reward predicting cues with human fMRI and determine whether humans showed the same reward prediction error signals as observed in midbrain dopamine cells in monkeys (6). In this task, participants are presented with instructive cues indicating if reward is available on this trial and if a response is required or not. After a variable delay, for trials requiring a response, a very brief imperative stimulus is presented during which the participant must respond to earn reward. However, the cues are pre-trained at length prior to scanning, and they indicate the magnitude of potential reward (or loss) given a speeded response (that is fast enough) following an imperative stimulus. In this way, the MID may involve incentive salience attribution to the trial-type cues and imperative stimulus. Patients with schizophrenia show attenuated striatal responses to the trial-type cues. These changes do not correlate with the severity of delusions (7). They may instead be a function of dopamine receptor blockade by antipsychotic drugs (8). When participants are re-scanned on the task having been switched from typical ($D_{2/3}$ blocking) to atypical antipsychotic drugs, their striatal cue response normalizes (8). However, aberrant prediction error responses at the time of the outcome (rather than the incentive salient reward predicting cue) in prefrontal cortex correlate with delusion severity (9). It appears there is more to the pathophysiology of delusions than the striatum and incentive salience.

2. The dopamine abnormality in psychosis should be in the ventral striatum

Kapur did not use the word striatum in his 2003 paper, nor did he make specific predictions about the limbic or ventral striatum being the site of dysfunction in psychosis. However, the originators of the motivational and incentive salience views of dopamine dysfunction, which Kapur applied to psychosis, did locate incentive salience attribution to dopamine function in the ventral striatum. (10). They did so on the basis of studies in humans which related ventral striatal dopamine release with increases in wanting of L-Dopa (11) and other drugs (12). Since the concept that Kapur co-opted to explain psychosis involves ventral striatal dopamine, then ventral striatal dopamine ought to relate to psychosis. At the time Kapur wrote the salience attribution article, striatal dopamine release had clearly been linked to schizophrenia and psychosis (13). Amphetamine stimulated striatal dopamine release, as measured with Positron Emission Tomography (PET) and a 11C-raclopride tracer, correlates with the severity of amphetamine induced psychotic symptoms in patients. However, with the advent of different tracers, and more precise spatial localization with modern imaging, it appears that the site of aberrant dopamine function in schizophrenia is the primarily the head of caudate or associative striatum (14). Accumbens dopamine may be unchanged in patients with schizophrenia (14). Furthermore, Fluoridopa (F-Dopa) scans measure dopamine synthesis capacity in pre-synaptic terminals (rather than dopamine in the synaptic cleft that competes with tracers at post-synaptic receptor sites). Dopamine synthesis capacity is higher in people at risk for psychosis and for those who are currently psychotic. This effect is particularly pronounced in the associative striatum, and not the accumbens (15) (see Figure 1). The regional location of the dopamine dysfunction in psychosis with most relevance to delusions appears to be the head of caudate/associative striatum. Intriguingly, aberrant prediction errors in this part of the striatum, and in DLPFC (which sends projections to associative striatum, and participates in the associative corticostriatal loop with dorsomedial striatum, see Figure 2), correlate with delusions in patients (16), and perceptual aberrations induced by ketamine (17). In the original paper, Kapur distinguished aberrant salience from aberrant prediction errors (on page 14, he dismissed the prediction error idea because prediction error dopamine experiments have not dealt well with aversive learning or the longer-term modulatory role of dopamine on behavior). In brief, prediction errors are surprises, they represent this mismatch between expectation and experience (18). They are encoded in dopamine neurons, and they drive learning directly through the strengthening of associations and indirectly through the allocation of attention (19). Surprising events are salient events, but not necessarily incentively salient.

Falsification?

We have scrutinized some predictions of incentive salience theory of psychosis – particularly as applied to delusions. Some (more or less auxiliary) predictions of the original theory have been found lacking. We believe that a strong incentive salience theory of delusions (and hallucinations) cannot hold. We make that claim based on cognitive neuroimaging data and PET data on dopamine release capacity. We acknowledge though that our understanding of dopamine function and salience attribution continues to evolve, and, for each novel idea about dopamine function (causal maps, uncertainty, see below) there is ample room for incentive salience. The dopamine data do not point to ventral striatum as a locus for psychosis. However, the dorsal striatum may still be involved in ascribing incentive salience, albeit evidence for this notion is lacking in the preclinical literature. A more parsimonious idea perhaps is that aberrant prediction errors increase uncertainty which may, under some circumstances, engage incentive salience attribution. This is particularly possible under a Bayesian predictive processing explanation for dopamine, uncertainty, perception, and belief. This account has been criticized for being unfalsifiable, especially since one can conceive of a set of prior beliefs that someone must have had to behave in any particular way. However, the commitments of a specific predictive processing view and their instantiation in the brain can be examined empirically and thus falsified. If we can measure prior overweighting in a task, we can test whether that metric relates to hallucinations (for example).

Evolution of the theory

How then have prediction error and incentive salience accounts of dopamine signaling and explaining psychosis converged and diverged in the last 20 years? The incentive salience theory of psychosis has been immensely generative. Over the past 20 years it has spawned countless empirical investigations and inspired numerous alternative theories (including the application of associative learning and predictive processing to explain psychosis). The theory itself has also evolved. In 2005, an updated version of the theory appeared – now focusing on novelty ascription to objects and associations (20). This version owes much to the earlier work of Jeffrey Gray and David Hemsley who identified a hippocampal-striatal circuit for novelty ascription, associability and learning as a key locus in generating psychosis (21, 22). Robert Miller before them argued that dopamine set the threshold for concluding two elements (thoughts, percepts, events, etc) were related, and that excess dopamine would garner spurious associations in people with psychosis (23). Next, Kapur and colleagues formalized the associative learning concept by evoking temporal difference learning and prediction error (see below).

However, Smith, Kapur and colleagues extended the temporal difference prediction error formalism to incorporate a model-based component encoded by tonic dopamine, based solely on theory crafting, which embodied a look-ahead process that guided the vigor of responding – this was still essentially equivalent to incentive salience albeit now a switch from phasic bursts to steady state release of dopamine (24–26). In 2008, Kapur’s team showed aberrant ventral striatal prediction error responses to neutral cues during fear conditioning in patients with schizophrenia, however, these did not correlate with delusions or other psychotic symptoms (27). Dysregulated prediction errors observed in patients in schizophrenia had previously observed in the midbrain during reward learning (28), and during causal learning, but critically this dysregulation in prediction error was correlated with delusions (16). Crucially, this task does not entail incentive salience attribution, but rather attentional salience misattribution, secondary to prediction errors ((29–31) see below).

This idea – of attentional salience misattribution- is mentioned in subsequent theoretical pieces by Kapur and Howes, however, the specter of incentive salience attribution remains – these are still incentive salience theories (32). In what follows, we will discuss what is now known about the dopamine dysfunction in psychosis, namely that it resides in dorsal rather than ventral striatum, and we will describe what is known about the functions of dopamine in those subregions of striatum. We will clarify the type of salience misattribution that characterizes psychosis and discuss its impact on belief formation and updating. Finally, we will examine hallucinations, and we will speculate on the roles of other neuromodulators considering these refined ideas about salience.

Salience?

Incentive salience was proposed as an explanation for how dopamine facilitates behaving, rather than learning (33, 34). It captures the desires which motivate us to pursue goals. Incentive salience is attributed to environmental stimuli (cues, contexts, settings, etc.). This can be learned, however learning is not causal for the attribution of incentive salience (35). In substance use disorders, excessive attribution of incentive salience to drug related cues can usurp goal-directed behavior via a desire to use drugs triggered by those cues, leading to relapse even after long periods of abstinence. For Kapur, the process of incentive salience attribution provided a means to link the phenomenology of psychosis (in particular, the delusional mood that is a prelude to delusion formation - wherein the world becomes pregnant with new meaning and significance) to what is presumed to be the underlying neurobiological dysfunction within the dopamine system, implied by the success of dopamine blocking antipsychotic medication in treating psychotic symptoms.

However, it is important to note that incentive salience is not about “salience” per se (as in, the quality of being particularly noticeable or important). Rather, incentive salience attribution involves the transformation of stimuli from “cold” to “hot” targets of attention via increasing their motivational value which, in turn, co-opts approach behaviors (33, 34). Obviously, not all salience is incentive salience. This disparity between salience (broadly construed) and incentive salience was raised soon after the original piece was published by Jeffrey Gray, David Hemsley, and their colleagues (36), who had written earlier about the role of dopamine and salience attribution in psychosis (21), but sadly the confusion has lasted.

There are numerous ways in which an event can be salient that have nothing to do with desire. Surprising events are salient (19). Events with unexpected, rather than expected, consequences are also salient. We learn from and about all these types of events - which Gray et al (21) and Grossberg before them recognized (37). Incentive salience does not capture this form of general surprise or attention-grabbing.

This is not merely quibbling about semantics. The failure to distinguish incentive salience from other types of salience has been problematic for both Kapur’s conceptualization and the field at large (not least those of us who study prediction error, whose work is often mistakenly conflated with Kapur’s) – since it has led to a focus on particular types of empirical investigation (focused on rewards) and it has stifled others – under the misapprehension that we understand how psychotic symptoms arise in the brain. These issues are further compounded by the difficulty in operationalizing motivational salience and capturing it with computational models that make unique predictions, and further, how those models of motivation intersect with computational models of learning (38–42). There are examples (38–42), but this area is less well developed empirically and theoretically. More broadly, we suggest that the link between motivational salience and positive psychotic symptoms (either in terms of neural or behavioral manifestations) has not been made empirically, despite the elegance of the hypothesis.

As we look back, it is increasingly apparent that Kapur’s theory as originally proposed is not sufficient and in some cases at odds with the associations between prediction error, uncertainty, and psychosis (43). This is not to say that dopamine is not involved in delusions, but that the argument that dopamine makes delusions relevant via amplifying their motivational value and turning them into targets of desire is insufficient to explain the phenomenology of delusion and incompatible with contemporary empirical neurobehavioral data.

What else might dopamine do, and how?

In the time since Kapur’s initial proposal there has been a renaissance in the preclinical dopamine field. We can now record dopamine release with carbon-fiber electrodes or genetically modified biosensors and manipulate dopamine neurons with optogenetics, both in real-time (in preclinical studies). While debate regarding the precise function or computation conveyed by dopamine activity and dopamine release is hotly contested, there is a general consensus that dopamine acts in one, or more, of three primary ways : 1) signaling and updating the value of cues, actions, and outcomes in a manner analogous to temporal-difference reinforcement learning (44–46), 2) promoting and instantiating cues and actions associated with appetitive and aversive outcomes with incentive motivational value that dynamically goads behavior (47–50) and 3) constructing an associative map with which one can aptly and nimbly infer the causal structure of the world (51, 52).

These are not mutually exclusive, since midbrain dopamine cells project to multiple different cortical and subcortical regions which may implement different computations. Others have attempted to find unifying computational motifs that may span the different influences dopamine has; these include encoding a domain-general surprise signal (irrespective of event valence) (53), and implementing an uncertainty computation over beliefs (54), again, not mutually exclusive.

While there is still robust evidence that dopamine within the nucleus accumbens promotes and encodes incentive salience attribution (49, 55), there is also strong data that mesolimbic dopamine can directly promote learning (consistent with reward prediction) and can even drive the learning of “value-less” stimulus-stimulus associations comprising a rudimentary cognitive map of causal relationships (56–59). This is sometimes referred to as a model-based computation, which implies a degree of understanding of the structure of the world or task (48, 51). In preclinical experiments evidence for such a process derives from situations where an update in behavior or an inference of change are demanded. Surprisingly, learning produced by mesolimbic dopamine activation is sensitive to reward devaluation, contingency degradation, and supports chains of actions (52, 58, 60–62). This theoretical notion, of an associative map, disconnects dopamine signaling from behavioral change, and behavioral change from value or incentive salience (52).

However, in many cases, incentive salience still provides an accurate distillation of the ways in which dopamine supports flexible behavior in a changing world. For example, mesolimbic dopamine is necessary for representation-mediated devaluation, where gastrointestinal distress paired with a previously food-paired stimulus in the devaluation of that food (63). This type of rapid revaluation requires active representation of the precise sensory properties of that food when you merely perceive a cue that predicts it (64). This representation-mediated process may represent a preclinical model of hallucinations (65). Notably, ketamine can facilitate representation mediated devaluation (66) (see below).

So far, we have discussed dopamine’s role in stimulus-reward association, incentive motivation, and causal inference. One framework that may span these functions in a manner consilient with dopamine’s role in psychosis involves dopamine encoding specific beliefs, in particular beliefs about uncertainty (67, 68).

The dopamine system is sensitive to probability and adapts to experiences of environmental stochasticity (69–72). There are unexpected increases in reward responses to long-delayed rewards, when there is uncertainty if reward will be delivered at all (67, 73). The dopamine system appears to track a probability distribution over states, which could serve as evidence of tracking which state of the world you are in currently. Regarding action-outcome contingency and agency attribution, such dopamine driven computations could impinge on our understanding of the structure of the world, and our ability to have agency within it, which may lead to psychotic symptoms like passivity delusions (74).

Dopamine neurons may not necessarily directly encode these beliefs about uncertainty, but rather the inputs and outputs of the dopamine system may facilitate a role for such uncertainties in the computational architecture of what the neurons are calculating and what is conveyed. Inputs to the dopamine system and their targets in the striatum from sources like the prefrontal cortex can themselves convey state uncertainty whereas action uncertainty could arise directly from dopaminoceptive striatal neurons themselves (75).

This theory of dopamine function certainly has face validity for psychosis, since delusions are beliefs about the self, the world, and other agents. However, it does not deal well with sensory prediction errors (which may be key to hallucinations, and to the perceptual aberrations that characterize the delusional mood). Furthermore, uncertainty has also been empirically linked to incentive salience attribution, and as such, a dopamine system that encodes uncertainty aberrantly may also encode aberrant incentive salience. In brief, uncertain (i.e. partially reinforced) cues act as incentive magnets; reward uncertainty associated with a conditioned stimulus may generate greater incentive motivation (or ‘wanting’) toward that reward cue (76, 77).

One account that could incorporate uncertainty, prediction error, and incentive motivation is Bayesian brain theory (78–80). Here, dopamine encodes the precision-weighted trade-off between sensory inputs and beliefs with regards to some world model. This account can explain the impact of dopamine disruptions on perception and belief, since such disruptions could bias the trade-off towards inputs (garnering spurious beliefs) or towards ones internal model and prior beliefs, leading to rigid delusions and perhaps hallucinations (81) (see below). In this view uncertainty garners incentive motivation because of the epistemic value that understanding the cue might unveil. Minimizing uncertainty is the goal, and as such, approaching uncertain cues, attending to and learning about them, may minimize more uncertainty in the long run (78–80). Next, we will examine the data in psychotic patients and what they might imply about dopamine functioning and psychosis – with particular focus on the neuroanatomical locus of dopamine dysfunction in dorsal striatum and how that aligns with the cognitive and neural mechanisms that underwrite delusions.

What is the state of the art in clinical computational cognitive neuroscience of delusions, hallucinations, and psychosis?

Psychosis seems to involve both aberrant prediction errors AND overweighted priors Which manifests in a specific study or dataset may depend on the type of task used to solicit the behavior/decision/belief AND on the symptom-experiences reported by particular study populations (82). They may also interact; such that aberrant prediction errors militate compensatory strengthening of higher-level prior beliefs. In a landmark study, Schmack and colleagues used optogenetic methods to target and stimulate specific dopamine pathways in the rodent dorsolateral striatum during auditory perceptual decision-making (83). In their task, the strength of the perceptual stimulus they were trained to detect varied such that their certainty of having perceived it varied too. The animals held their nose in a nose-port for longer the more confident they were in their choice (presence or absence of stimulus) (83). They found that increasing dopamine release in the rodent tail of striatum induced a pattern of high-confidence false alarms. Computational modeling indicated that the dopamine transients were influencing the weighting of a perceptual expectation - or prior belief. As dopamine increased in the pathway, animals more heavily weighted their perceptual inferences, whether the cue they were trained to respond to was actually presented or not, toward their prior expectation. This is consistent with observations of excess prior weighting in humans who hallucinate (84), which has also been correlated with dopamine levels in the associative striatum measured with PET (85).

A conciliatory proposition

Might it be possible to reconcile these ideas and data into a coherent account of how dopamine dysfunction elicits psychotic symptoms? We believe so. While some aspects of what we propose are inspired by the aberrant salience theory, our conciliatory proposition goes beyond incentive salience. Aberrant prediction errors have been observed in the mesocorticolimbic dopamine system in people with psychosis with fMRI (17, 28). These signals typically correlate with delusions (17) (although not always (28)). Aberrant dopamine signaling could cause a prediction error to be registered when none ought to be. Such a signal would drive the allocation of attention toward potentially explanatory cues and events (31). Under a causal inference theory of dopamine, this is a process of retrospective causal inference - towards the best explanation for why a particular stimulus or event feels significant (52). This notion has precedent in earlier critiques of RPE theory, such as Redgrave and Gurney (74), who suggest that midbrain dopamine cells are responsible for coincidence detection between intentions and outcomes - such that important (salient) events can be attributed to ones own actions (or not) (74). Perturbations in the timing of prediction errors might lead to over ascription of agency (as in experiences of apparent mental causation, or grandiosity). They might also lead to under-ascription of agency, as in delusions of passivity, where thoughts and actions are attributed to (often malevolent) external agents (74).

There are multiple potential sources of this disruption that could be manifest within the corticostriatal circuitry and its dopaminergic input (see Figure 3). We hypothesize that in schizophrenia and psychosis there is a combination of potential deficits in striatal circuitry and their associated computations that lead to symptomatology: First, per the meta-analytic evidence (15), upregulation of dopamine function in the associative and sensorimotor loops that promotes excessive goal-monitoring and action preference but in the absence of the equivalent upregulation for the assignment of meaning or valuation that would underlie such preferences. Typically these systems would achieve parity via cross compartment communication, for example, a disconnect between the limbic loop and the associative and sensorimotor loops may result in a lack of causal meaning for actions executed and needing to post-hoc assign meaning to one’s own behavior that apparently occurs without valid meaning. These compartment specific effects in striatum may ultimately arise from deficits in prediction error broadcasting by midbrain dopamine neurons which in turn can lead to over- and under-updating of specific priors or predictions in these circuits (83).

Taken together, these data suggest that delusions might be an inference to the best explanation for aberrant dopamine dynamics - where different types of neural noise are interpreted as real signal, indicating malintent (in the case of paranoia) or aberrant significance (in the case of non-paranoid delusions) (86). We note, of course, that the preclinical data on which we are speculating were gathered in the mesoaccumbal circuit. We do not know whether its principles translate to mesocortical circuits or a more general deficit in the compartmentalization of computational functions (87). However, we do know that dopamine in the dorsomedial striatum encodes volatility and change during reversal learning. And we know that dopamine in the dorsolateral striatum can encode motivational salience (88) and habit learning (89), but also perceptual predictions and the overweighting of such predictions to produce high confidence false alarms (hallucinations) in rodents (83). Dorsomedial and Dorsolateral Striatum are precisely the loci of dopamine elevation in psychosis, rather than ventral striatum (see Figure 1)

Future directions

We see four avenues that demand further exploration:

1. **Continued theorizing about dopamine** to accommodate our rapidly expanding knowledge base – both clinically and preclinically. For example, errant surprise signals render events coincident with them (e.g. stimuli, thoughts, percepts) more motivationally salient. How might this happen? More broadly, a consilient explanation of dopamine function, encompassing surprise, causal inference, perception, cognition (social and non-social), and motivation is desirable.
2. **Translation of tasks that engage dopamine function and relate to psychosis, back and forth across species**. There has been notable recent progress in this regard with animal models of hallucinations (66, 83, 90) and paranoia (91, 92) developing around a computational understanding of dopamine and glutamate signaling (93). See (94) for a recent example of new data on dopamine driving a new theory about schizophrenia. However, much work is needed to translate these advances into tasks that can predict psychotic symptom onset or severity. A key outstanding question here is whether one task that addresses all possible hypotheses about the impacts of dopamine on cognition and perception is possible or even desirable (given challenges around task motivation and broader cognition in people with psychosis). An early success with midbrain signaling (95) did show aberrant midbrain responses to perceptual and emotional surprise which correlated with delusions (96), but also negative symptoms (suggesting a symptom burden explanation). Further a great deal of data were lost due to fatigue and movement. Shorter tasks aimed at separate mechanisms and symptoms (e.g. predictions versus prediction errors) have yielded more success.
3. **Exploration of how other neuromodulators contribute to psychosis**. Kapur carefully acknowledged that dopamine is not the only neuromodulator or

neurotransmitter implicated in psychosis. He noted that glutamate appears to be relevant to the pathophysiology. Bayesian predictive coding accounts suggest that glutamate (and GABA) may carry the predictions and prediction errors, and that slower neuromodulators (like dopamine and serotonin) may instantiate their precision-weighting (97). In this way, serotonergic hallucinogens may overweight perceptual expectancies - without any intermediary change in incentive salience (98). Pimavanserin is a 5HT2a partial agonist that is used to treat hallucinations and delusions in Parkinson’s disease, and is being tested in schizophrenia. Likewise cholinergic interventions may change the weighting of perceptual and doxastic expectations (99); through effects on acetylcholine release on the weighting of prior expectations (100) and the impact of prediction errors (101).

Examination of how psychosis relates to different domains of cognition and perception and the circuitry that underlies them.

There are mechanisms of learning and belief updating outside of frontostriatal circuits which may prove crucial in understanding delusions. For example, the orbitofrontal cortex and hippocampus seem key for representing conceptual and social spaces (102). Perturbations in that mapping, and inferences based on that mapping may contribute to the genesis and maintenance of psychosis. Again, incentive salience attribution need not be directly involved in these. Finally, these maps contribute to the genesis of predictions in the subcortical dopamine system, and may allow for more abstract combination of maps and integration of disparate experiences (103) - all of which may go awry to underwrite the errant conclusions of people with delusional beliefs. We note with interest that representation of maps and rules appears to be disrupted in people with schizophrenia (104), and that one proposed mechanism of action of atypical (less dopaminergic) drugs is the serotonergic and glutamatergic projections from orbitofrontal cortex to midbrain dopamine cells (105). We note too that more complex propositional mechanisms do not appear to be related to delusion-like beliefs in healthy people (106), however counterfactual processes and cue combination mechanisms may be impaired in those at clinical high risk for psychosis (86).

Following these four paths whislt also attending to their role(s) in the genesis, maintenance, and treatment of delusions are what we consider critical next steps for the field.

Conclusion

One theme that has emerged from this discussion; there are multiple ways in which events can be salient - by being predictive of important outcomes, by being unpredictable, by being biologically significant for our survival (related to human voices or social relationships). This was a key insight from Kapur - that the biology of such significance was central to the pathophysiology of psychosis, and that we should seek to unite the phenomenology of psychosis with that biology. We have challenged the specifics of salience theory’s construction and predictions, but we deeply appreciate the effort and have offered what we hope are constructive alternatives. Our modest additions are also worthy of critique. For example, we have favored prediction errors and abductive inference to the best explanation as a grounds for delusion formation. Neither we nor Kapur attempted to address delusion fixity. One possibility, in the spirit of aberrant salience theory, is that aberrant dopamine signals induce a partial reinforcement extinction effect - minor pieces of evidence that support a given belief (perhaps even coincidentally) will render that belief less sensitive to disconfirmation, especially under uncertainty (76). Delusions might become overhypotheses, overfit to data, and capable of explaining anything (107). This has been modeled computationally with a Drichelet process, inspired by category learning, and awaits empirical confirmation (107). Something akin to it has been observed recently (86). Models that can account for and integrate the various ways in which events can be salient and in which our beliefs can become circumscribed and inflexible will be key to furthering this work and, ultimately, using it to help patients.

Funding

Kurt Fraser was supported by funding from the National Institutes of Health (MH127782). Phil Corlett was supported by the Yale University Department of Psychiatry, the Connecticut Mental Health Center (CMHC) and Connecticut State Department of Mental Health and Addiction Services (DMHAS). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing Interests

PRC is a cofounder of Tetricus Labs, a precision-psychiatry company, who did not fund this work. KMF has no competing interests.

References

1. Kapur S Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiat. 2003;160:13–23. [PubMed: 12505794]
2. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev. 1998;28:309–369. [PubMed: 9858756]
3. Roiser JP, Stephan KE, den Ouden HE, Barnes TR, Friston KJ, Joyce EM. Do patients with schizophrenia exhibit aberrant salience? Psychol Med. 2009;39:199–209. [PubMed: 18588739]
4. Schmidt A, Antoniades M, Allen P, Egerton A, Chaddock CA, Borgwardt S, Fusar-Poli P, Roiser JP, Howes O, McGuire P. Longitudinal alterations in motivational salience processing in ultra-high-risk subjects for psychosis. Psychol Med. 2017;47:243–254. [PubMed: 27697078]
5. Abboud R, Roiser JP, Khalifeh H, Ali S, Harrison I, Killaspy HT, Joyce EM. Are persistent delusions in schizophrenia associated with aberrant salience? Schizophr Res Cogn. 2016;4:32–38. [PubMed: 27284531]
6. Knutson B, Westdorp A, Kaiser E, Hommer D. FMRI visualization of brain activity during a monetary incentive delay task. Neuroimage. 2000;12:20–27. [PubMed: 10875899]
7. Arrondo G, Segarra N, Metastasio A, Ziauddeen H, Spencer J, Reinders NR, Dudas RB, Robbins TW, Fletcher PC, Murray GK. Reduction in ventral striatal activity when anticipating a reward in depression and schizophrenia: a replicated cross-diagnostic finding. Front Psychol. 2015;6:1280. [PubMed: 26379600]
8. Juckel G Inhibition of the reward system by antipsychotic treatment. Dialogues Clin Neurosci. 2016;18:109–114. [PubMed: 27069385]
9. Waltz JA, Schweitzer JB, Ross TJ, Kurup PK, Salmeron BJ, Rose EJ, Gold JM, Stein EA. Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. Neuropsychopharmacology. 2010;35:2427–2439. [PubMed: 20720534]
10. Berridge KC. The debate over dopamine’s role in reward: the case for incentive salience. Psychopharmacology. 2007;191:391–431. [PubMed: 17072591]

11. Evans AH, Pavese N, Lawrence AD, Tai YF, Appel S, Doder M, Brooks DJ, Lees AJ, Piccini P. Compulsive drug use linked to sensitized ventral striatal dopamine transmission. Ann Neurol. 2006;59:852–858. [PubMed: 16557571]
12. Leyton M, Boileau I, Benkelfat C, Diksic M, Baker G, Dagher A. Amphetamine-induced increases in extracellular dopamine, drug wanting, and novelty seeking: a PET/[11C]raclopride study in healthy men. Neuropsychopharmacology. 2002;27:1027–1035. [PubMed: 12464459]
13. Laruelle M, Abi-Dargham A. Dopamine as the wind of the psychotic fire: new evidence from brain imaging studies. Journal of psychopharmacology. 1999;13:358–371. [PubMed: 10667612]
14. Kegeles LS, Abi-Dargham A, Frankle WG, Gil R, Cooper TB, Slifstein M, Hwang DR, Huang Y, Haber SN, Laruelle M. Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. Archives of general psychiatry. 2010;67:231–239. [PubMed: 20194823]
15. McCutcheon R, Beck K, Jauhar S, Howes OD. Defining the Locus of Dopaminergic Dysfunction in Schizophrenia: A Meta-analysis and Test of the Mesolimbic Hypothesis. Schizophr Bull. 2018;44:1301–1311. [PubMed: 29301039]
16. Corlett PR, Murray GK, Honey GD, Aitken MR, Shanks DR, Robbins TW, Bullmore ET, Dickinson A, Fletcher PC. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain : a journal of neurology. 2007;130:2387–2400. [PubMed: 17690132]
17. Corlett PR, Honey GD, Aitken MR, Dickinson A, Shanks DR, Absalom AR, Lee M, Pomarol-Clotet E, Murray GK, McKenna PJ, Robbins TW, Bullmore ET, Fletcher PC. Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. Archives of general psychiatry. 2006;63:611–621. [PubMed: 16754834]
18. Rescorla RA, Wagner AR: A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. in Classical conditioning II: Current research and theory. Edited by Black AH, Prokasy WF New York, Appleton-Century-Crofts; 1972.
19. Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol Rev. 1980;87:532–552. [PubMed: 7443916]
20. Kapur S, Mizrahi R, Li M. From dopamine to salience to psychosis--linking biology, pharmacology and phenomenology of psychosis. Schizophr Res. 2005;79:59–68. [PubMed: 16005191]
21. Gray JA, Feldon J, Rawlins JNP, Hemsley DR The neuropsychology of schizophrenia. Behav Brain Sci. 1995;14:1–20.
22. Hemsley DR: Perceptual and cognitive abnormalities as the basis for schizophrenic symptoms. in The Neuropsychology of Schizophrenia. Edited by David AS, Cutting, Hove JC, UK, Laurence Erlbaum Associates; 1994. pp. 97–118.
23. Miller R Schizophrenic psychology, associative learning and the role of forebrain dopamine. Med Hypotheses. 1976;2:203–211. [PubMed: 9558]
24. Smith A, Li M, Becker S, Kapur S. Dopamine, prediction error and associative learning: a model-based account. Network. 2006;17:61–84. [PubMed: 16613795]
25. Smith AJ, Becker S, Kapur S. A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. Neural Comput. 2005;17:361–395. [PubMed: 15720772]
26. Smith AJ, Li M, Becker S, Kapur S. Linking animal models of psychosis to computational models of dopamine function. Neuropsychopharmacology. 2007;32:54–66. [PubMed: 16710321]
27. Jensen J, Willeit M, Zipursky RB, Savina I, Smith AJ, Menon M, Crawley AP, Kapur S. The formation of abnormal associations in schizophrenia: neural and behavioral evidence. Neuropsychopharmacology. 2008;33:473–479. [PubMed: 17473838]
28. Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, Jones PB, Bullmore ET, Robbins TW, Fletcher PC. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Molecular psychiatry. 2008;13:239, 267–276.
29. Corlett PR, Fletcher PC. Delusions and prediction error: clarifying the roles of behavioural and brain responses. Cogn Neuropsychiatry. 2015:1–11. [PubMed: 25078663]

30. Corlett PR, Honey GD, Fletcher PC. From prediction error to psychosis: ketamine as a pharmacological model of delusions. Journal of psychopharmacology. 2007;21:238–252. [PubMed: 17591652]
31. Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH. Toward a neurobiology of delusions. Progress in neurobiology. 2010;92:345–369. [PubMed: 20558235]
32. Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III--the final common pathway. Schizophr Bull. 2009;35:549–562. [PubMed: 19325164]
33. Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Res Brain Res Rev. 1993;18:247–291. [PubMed: 8401595]
34. Robinson TE, Berridge KC. The psychology and neurobiology of addiction: an incentive-sensitization view. Addiction. 2000;95 Suppl 2:S91–117. [PubMed: 11002906]
35. Robinson MJ, Berridge KC. Instant transformation of learned repulsion into motivational "wanting". Curr Biol. 2013;23:282–289. [PubMed: 23375893]
36. Gray JA. On biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry. 2004;161:377; author reply 377–378. [PubMed: 14754801]
37. Grossberg S Processing of expected and unexpected events during conditioning and attention: a psychophysiological theory. Psychol Rev. 1982;89:529–572. [PubMed: 7178332]
38. Dayan P, Balleine BW. Reward, motivation, and reinforcement learning. Neuron. 2002;36:285–298. [PubMed: 12383782]
39. McClure SM, Daw ND, Montague PR. A computational substrate for incentive salience. Trends in neurosciences. 2003;26:423–428. [PubMed: 12900173]
40. Smith BJ, Read SJ. Modeling incentive salience in Pavlovian learning more parsimoniously using a multiple attribute model. Cognitive, Affective, & Behavioral Neuroscience. 2022:1–14.
41. Zhang J, Berridge KC, Aldridge JW. Computational models of incentive-sensitization in addiction: Dynamic limbic transformation of learning into motivation. Computational neuroscience of drug addiction. 2012:189–203.
42. Zhang J, Berridge KC, Tindell AJ, Smith KS, Aldridge JW. A neural computational model of incentive salience. PLoS computational biology. 2009;5:e1000437. [PubMed: 19609350]
43. Feeney EJ, Groman SM, Taylor JR, Corlett PR. Explaining Delusions: Reducing Uncertainty Through Basic and Computational Neuroscience. Schizophr Bull. 2017;43:263–272. [PubMed: 28177090]
44. Schultz W, Dickinson A. Neuronal coding of prediction errors. Annu Rev Neurosci. 2000;23:473–500. [PubMed: 10845072]
45. Dabney W, Kurth-Nelson Z, Uchida N, Starkweather CK, Hassabis D, Munos R, Botvinick M. A distributional code for value in dopamine-based reinforcement learning. Nature. 2020;577:671–675. [PubMed: 31942076]
46. Watabe-Uchida M, Eshel N, Uchida N. Neural Circuitry of Reward Prediction Error. Annu Rev Neurosci. 2017;40:373–394. [PubMed: 28441114]
47. Berridge KC. From prediction error to incentive salience: mesolimbic computation of reward motivation. European Journal of neuroscience. 2012;35:1124–1143. [PubMed: 22487042]
48. Dayan P, Berridge KC. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. Cognitive, Affective, & Behavioral Neuroscience. 2014;14:473–492.
49. Saunders BT, Richard JM, Margolis EB, Janak PH. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat Neurosci. 2018;21:1072–1083. [PubMed: 30038277]
50. de Jong JW, Fraser KM, Lammel S. Mesoaccumbal dopamine heterogeneity: what do dopamine firing and release have to do with it? Annual review of neuroscience. 2022;45:109–129.
51. Langdon AJ, Sharpe MJ, Schoenbaum G, Niv Y. Model-based predictions for dopamine. Curr Opin Neurobiol. 2018;49:1–7. [PubMed: 29096115]
52. Jeong H, Taylor A, Floeder JR, Lohmann M, Mihalas S, Wu B, Zhou M, Burke DA, Namboodiri VMK. Mesolimbic dopamine release conveys causal associations. Science. 2022;378:eabq6740. [PubMed: 36480599]

53. Kutlu MG, Zachry JE, Melugin PR, Cajigas SA, Chevee MF, Kelly SJ, Kutlu B, Tian L, Siciliano CA, Calipari ES. Dopamine release in the nucleus accumbens core signals perceived saliency. Curr Biol. 2021;31:4748–4761 e4748. [PubMed: 34529938]
54. Gardner MP, Schoenbaum G, Gershman SJ. Rethinking dopamine as generalized prediction error. Proceedings of the Royal Society B. 2018;285:20181645.
55. Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H. A selective role for dopamine in stimulus–reward learning. Nature. 2011;469:53–57. [PubMed: 21150898]
56. Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci. 2013;16:966–973. [PubMed: 23708143]
57. Sadacca BF, Jones JL, Schoenbaum G. Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. Elife. 2016;5.
58. Keiflin R, Pribut HJ, Shah NB, Janak PH. Ventral tegmental dopamine neurons participate in reward identity predictions. Current Biology. 2019;29:93–103. e103. [PubMed: 30581025]
59. Sharpe MJ, Batchelor HM, Mueller LE, Yun Chang C, Maes EJP, Niv Y, Schoenbaum G. Dopamine transients do not act as model-free prediction errors during associative learning. Nat Commun. 2020;11:106. [PubMed: 31913274]
60. Sharpe MJ, Chang CY, Liu MA, Batchelor HM, Mueller LE, Jones JL, Niv Y, Schoenbaum G. Dopamine transients are sufficient and necessary for acquisition of model-based associations. Nat Neurosci. 2017;20:735–742. [PubMed: 28368385]
61. Fraser KM, Pribut HJ, Janak PH, Keiflin R. From prediction to action: dissociable roles of ventral tegmental area and substantia nigra dopamine neurons in instrumental reinforcement. Journal of Neuroscience. 2023;43:3895–3908. [PubMed: 37185097]
62. Garr E, Cheng Y, Jeong H, Brooke S, Castell L, Bal A, Magnard R, Namboodiri VMK, Janak PH. Mesostriatal dopamine is sensitive to changes in specific cue-reward contingencies. Science Advances. 2024;10:eadn4203. [PubMed: 38809978]
63. Fry BR, Fex V, Sawa A, Niwa M, Johnson AW. The role of midbrain dopamine cells projecting to the insular cortex in mediated performance: Implications for animal models of reality testing. Behav Neurosci. 2024;138:164–177. [PubMed: 38934920]
64. Holland PC. Acquisition of representation-mediated conditioned food aversions. Learning and Motivation. 1981;12:1–18.
65. McDannald MA, Whitt JP, Calhoon GG, Piantadosi PT, Karlsson R-M, O'Donnell P, Schoenbaum G. Impaired reality testing in an animal model of schizophrenia. Biol Psychiatry. 2011;70:1122–1126. [PubMed: 21798517]
66. Fleming LM, Jaynes FB, Thompson SL, Corlett PR, Taylor JR. Targeted effects of ketamine on perceptual expectation during mediated learning in rats. Psychopharmacology. 2022.
67. Babayan BM, Uchida N, Gershman SJ. Belief state representation in the dopamine system. Nat Commun. 2018;9:1891. [PubMed: 29760401]
68. Gershman SJ, Uchida N. Believing in dopamine. Nat Rev Neurosci. 2019;20:703–714. [PubMed: 31570826]
69. Singer B, Scott-Railton J, Vezina P. Unpredictable saccharin reinforcement enhances locomotor responding to amphetamine. Behavioural brain research. 2012;226:340–344. [PubMed: 21924296]
70. Zack M, Featherstone RE, Mathewson S, Fletcher PJ. Chronic exposure to a gambling-like schedule of reward predictive stimuli can promote sensitization to amphetamine in rats. Frontiers in behavioral neuroscience. 2014;8:36. [PubMed: 24574987]
71. Zeeb FD, Li Z, Fisher DC, Zack MH, Fletcher PJ. Uncertainty exposure causes behavioural sensitization and increases risky decision-making in male rats: toward modelling gambling disorder. Journal of Psychiatry and Neuroscience. 2017;42:404–413. [PubMed: 28832319]
72. Mascia P, Wang Q, Brown J, Nesbitt KM, Kennedy RT, Vezina P. Maladaptive consequences of repeated intermittent exposure to uncertainty. Prog Neuropsychopharmacol Biol Psychiatry. 2020;99:109864. [PubMed: 31952958]

73. de Jong JW, Liang Y, Verharen JP, Fraser KM, Lammel S. State and rate-of-change encoding in parallel mesoaccumbal dopamine pathways. Nat Neurosci. 2024;27:309–318. [PubMed: 38212586]
74. Redgrave P, Gurney K. The short-latency dopamine signal: a role in discovering novel actions? Nat Rev Neurosci. 2006;7:967–975. [PubMed: 17115078]
75. Starkweather CK, Gershman SJ, Uchida N. The medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. Neuron. 2018;98:616–629. e616. [PubMed: 29656872]
76. Anselme P Incentive salience attribution under reward uncertainty: A Pavlovian model. Behav Processes. 2015;111:6–18. [PubMed: 25444780]
77. Anselme P, Robinson MJ, Berridge KC. Reward uncertainty enhances incentive salience attribution as sign-tracking. Behav Brain Res. 2013;238:53–61. [PubMed: 23078951]
78. FitzGerald TH, Dolan RJ, Friston K. Dopamine, reward learning, and active inference. Front Comput Neurosci. 2015;9:136. [PubMed: 26581305]
79. Friston KJ, Shiner T, FitzGerald T, Galea JM, Adams R, Brown H, Dolan RJ, Moran R, Stephan KE, Bestmann S. Dopamine, affordance and active inference. PLoS Comput Biol. 2012;8:e1002327. [PubMed: 22241972]
80. Shiner T, Symmonds M, Guitart-Masip M, Fleming SM, Friston KJ, Dolan RJ. Dopamine, Salience, and Response Set Shifting in Prefrontal Cortex. Cereb Cortex. 2015;25:3629–3639. [PubMed: 25246512]
81. Sterzer P, Adams RA, Fletcher P, Frith C, Lawrie SM, Muckli L, Petrovic P, Uhlhaas P, Voss M, Corlett PR. The Predictive Coding Account of Psychosis. Biological psychiatry. 2018.
82. Fleming LM, Lemonde AC, Benrimoh D, Gold JM, Taylor JR, Malla A, Joober R, Iyer SN, Lepage M, Shah J, Corlett PR. Using dimensionality-reduction techniques to understand the organization of psychotic symptoms in persistent psychotic illness and first episode psychosis. Sci Rep. 2023;13:4841. [PubMed: 36964175]
83. Schmack K, Bosc M, Ott T, Sturgill JF, Kepecs A. Striatal dopamine mediates hallucination-like perception in mice. Science. 2021;372. [PubMed: 34437095]
84. Powers AR, Mathys C, Corlett PR. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. Science. 2017;357:596–600. [PubMed: 28798131]
85. Cassidy CM, Balsam PD, Weinstein JJ, Rosengard RJ, Slifstein M, Daw ND, Abi-Dargham A, Horga G. A Perceptual Inference Mechanism for Hallucinations Linked to Striatal Dopamine. Curr Biol. 2018.
86. Rossi-Goldthorpe R, Silverstein SM, Gold JM, Schiffman J, Waltz JA, Williams TF, Powers AR, Woods SW, Zinbarg RE, Mittal VA, Ellman LM, Strauss GP, Walker EF, Levin JA, Castiello S, Kenney J, Corlett PR. Different learning aberrations relate to delusion-like beliefs with different contents. Brain : a journal of neurology. 2024.
87. Morales M, Margolis EB. Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. Nat Rev Neurosci. 2017;18:73–85. [PubMed: 28053327]
88. DiFeliceantonio AG, Berridge KC. Dorsolateral neostriatum contribution to incentive salience: opioid or dopamine stimulation makes one reward cue more motivationally attractive than another. European Journal of Neuroscience. 2016;43:1203–1218. [PubMed: 26924040]
89. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci. 2004;19:181–189. [PubMed: 14750976]
90. Koh MT, Ahrens PS, Gallagher M. A greater tendency for representation mediated learning in a ketamine mouse model of schizophrenia. Behav Neurosci. 2018;132:106–113. [PubMed: 29672108]
91. Reed EJ, Uddenberg S, Suthaharan P, Mathys CD, Taylor JR, Groman SM, Corlett PR. Paranoia as a deficit in non-social belief updating. Elife. 2020;9.
92. Suthaharan P, Thompson SL, Rossi-Goldthorpe RA, Rudebeck PH, Walton ME, Chakraborty S, Noonan MP, Costa VD, Murray EA, Mathys CD, Groman SM, Mitchell AS, Taylor JR, Corlett PR, Chang SWC. Lesions to the mediodorsal thalamus, but not orbitofrontal cortex, enhance volatility beliefs linked to paranoia. Cell Rep. 2024;43:114355. [PubMed: 38870010]

93. Mondragon E Mediated learning: A computational rendering of ketamine-induced symptoms. Behav Neurosci. 2024;138:178–194. [PubMed: 38635181]

94. Millard SJ, Bearden CE, Karlsgodt KH, Sharpe MJ. The prediction-error hypothesis of schizophrenia: new data point to circuit-specific changes in dopamine activity. Neuropsychopharmacology. 2022;47:628–640. [PubMed: 34588607]

95. Bunzeck N, Duzel E. Absolute coding of stimulus novelty in the human substantia nigra/VTA. Neuron. 2006;51:369–379. [PubMed: 16880131]

96. Knolle F, Ermakova AO, Justicia A, Fletcher PC, Bunzeck N, Duzel E, Murray GK. Brain responses to different types of salience in antipsychotic naive first episode psychosis: An fMRI study. Transl Psychiatry. 2018;8:196. [PubMed: 30242202]

97. Friston K A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 2005;360:815–836. [PubMed: 15937014]

98. Corlett PR, Frith CD, Fletcher PC. From drugs to deprivation: a Bayesian framework for understanding models of psychosis. Psychopharmacology. 2009;206:515–530. [PubMed: 19475401]

99. Friston KJ. Hallucinations and perceptual inference. Behavioral and Brain Science. 2005;28:764–766.

100. Warburton DM, Wesnes K, Edwards J, Larrad D. Scopolamine and the sensory conditioning of hallucinations. Neuropsychobiology. 1985;14:198–202. [PubMed: 3835496]

101. Bao S, Chan VT, Merzenich MM. Cortical remodelling induced by activity of ventral tegmental dopamine neurons. Nature. 2001;412:79–83. [PubMed: 11452310]

102. Tavares RM, Mendelsohn A, Grossman Y, Williams CH, Shapiro M, Trope Y, Schiller D. A Map for Social Navigation in the Human Brain. Neuron. 2015;87:231–243. [PubMed: 26139376]

103. Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, Schoenbaum G. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci. 2011;14:1590–1597. [PubMed: 22037501]

104. Nour MM, Liu Y, Arumuham A, Kurth-Nelson Z, Dolan RJ. Impaired neural replay of inferred relationships in schizophrenia. Cell. 2021;184:4315–4328 e4317. [PubMed: 34197734]

105. Homayoun H, Moghaddam B. Orbitofrontal cortex neurons as a common target for classic and glutamatergic antipsychotic drugs. Proc Natl Acad Sci U S A. 2008;105:18041–18046. [PubMed: 19004793]

106. Ongchoco JDK, Castiello S, Corlett PR. Excessive teleological thinking is driven by aberrant associations and not by failure of reasoning. iScience. 2023;26:107643. [PubMed: 37705957]

107. Erdmann T, Mathys C. A generative framework for the study of delusions. Schizophr Res. 2022;245:42–49. [PubMed: 33648810]

108. Plot Digitizer.

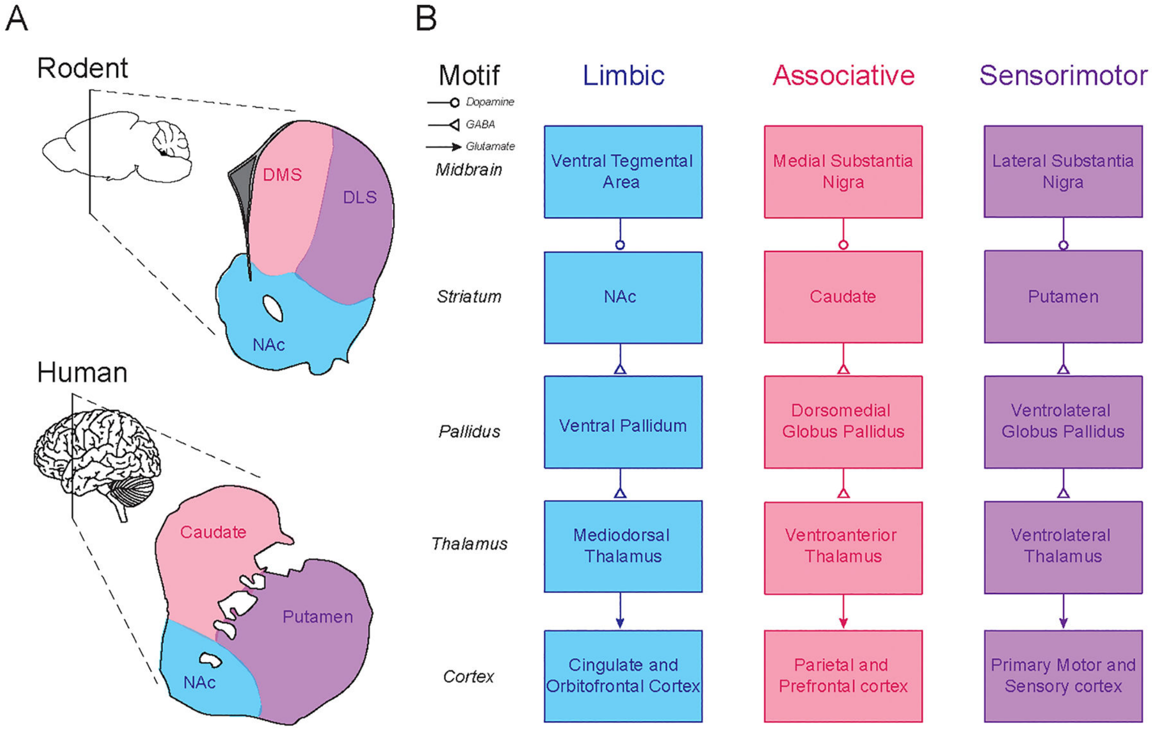

Figure 1. Dopaminergic dysregulation is system-specific in schizophrenia.

A) According to Kapur, the localization of dopamine dysregulation in schizophrenia is within those dopaminergic systems mediating the attribution of incentive motivation to environmental stimuli. In other words, Kapur argued the reason antipsychotic medication is effective is due to its normalization of dopamine function within the nucleus accumbens. **B)** In a recent metaanalysis of 21 neuroimaging studies, McCutcheon et al. (15), determined that the focal source of dopamine dysfunction is not within the nucleus accumbens, but consistently hyperactive within the caudate and putamen. Data in **B** were reproduced using plot digitizer (108).

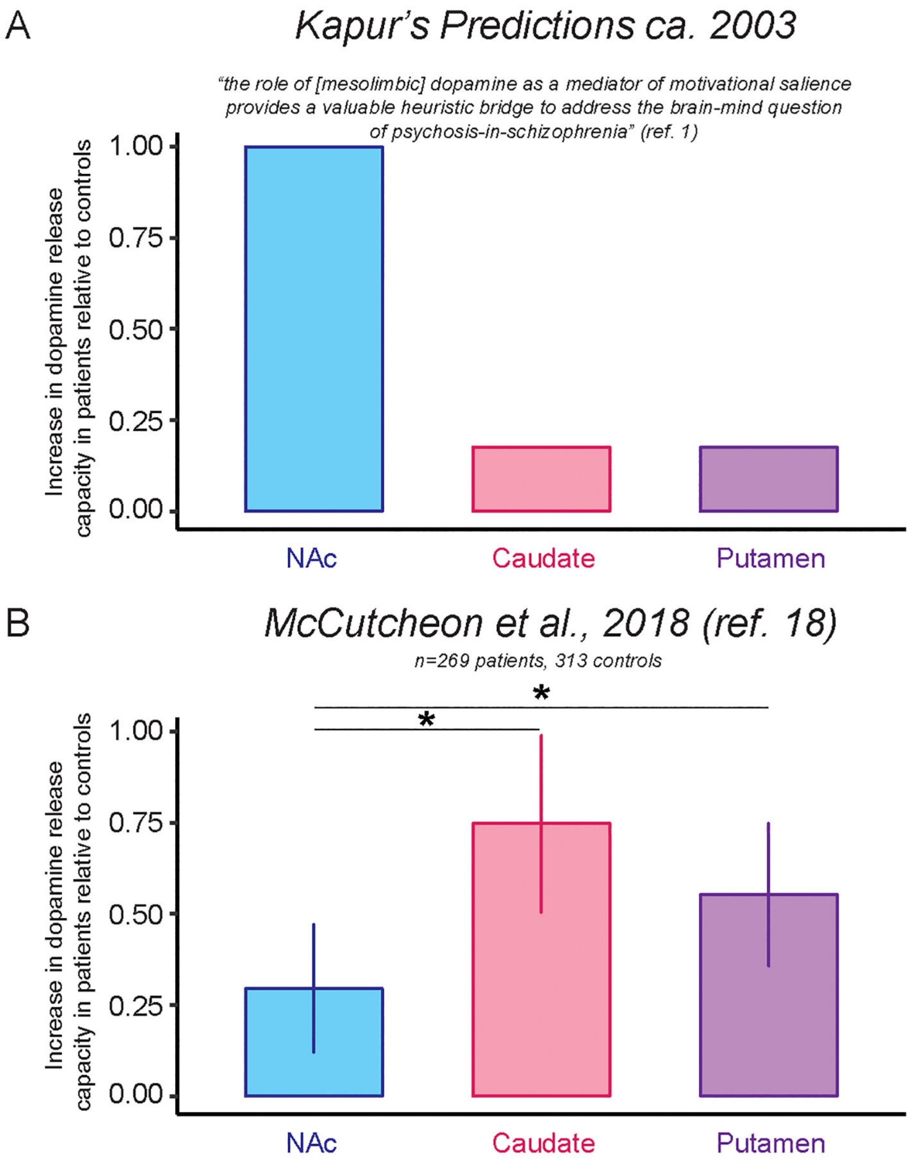

Figure 2. Striatal circuit homology across organisms.

A) The striatum is a complex forebrain structure comprised primarily of GABAergic spiny projection neurons that is divided into differing levels of subregions depending on the precise neurochemical or anatomical markers one prefers. Generally, there is a consensus that at least three general divisions exist: Putamen (dorsal lateral striatum in lower mammals), Caudate (dorsal medial striatum in lower mammals), and nucleus accumbens. In humans and non-human primates, the caudate and putamen are divided in part by the internal capsule, whereas in rodents and other lower mammals these structures are more or less contiguous. **B)** Each striatal region is part of a greater circuit motif that is common across regions. Each striatal region receives dopaminergic input from a defined set of midbrain dopamine neurons, projects to a pallidal region which each project to unique thalamic structures who ultimately project to the cortex. For simplicity, we have omitted known projections between systems (primarily the interactions between striatal regions and midbrain dopamine neurons and the existence of arkypallidal neurons which project back from pallidus to the striatum). Classically, this circuit motif has provided a useful mnemonic for understanding the systems level interactions between these structures, with nucleus accumbens circuitry being referred to as limbic, caudate circuitry as associative, and putamen circuity as sensorimotor. DMS, dorsal medial striatum; DLS, dorsal lateral striatum; NAc, nucleus accumbens.

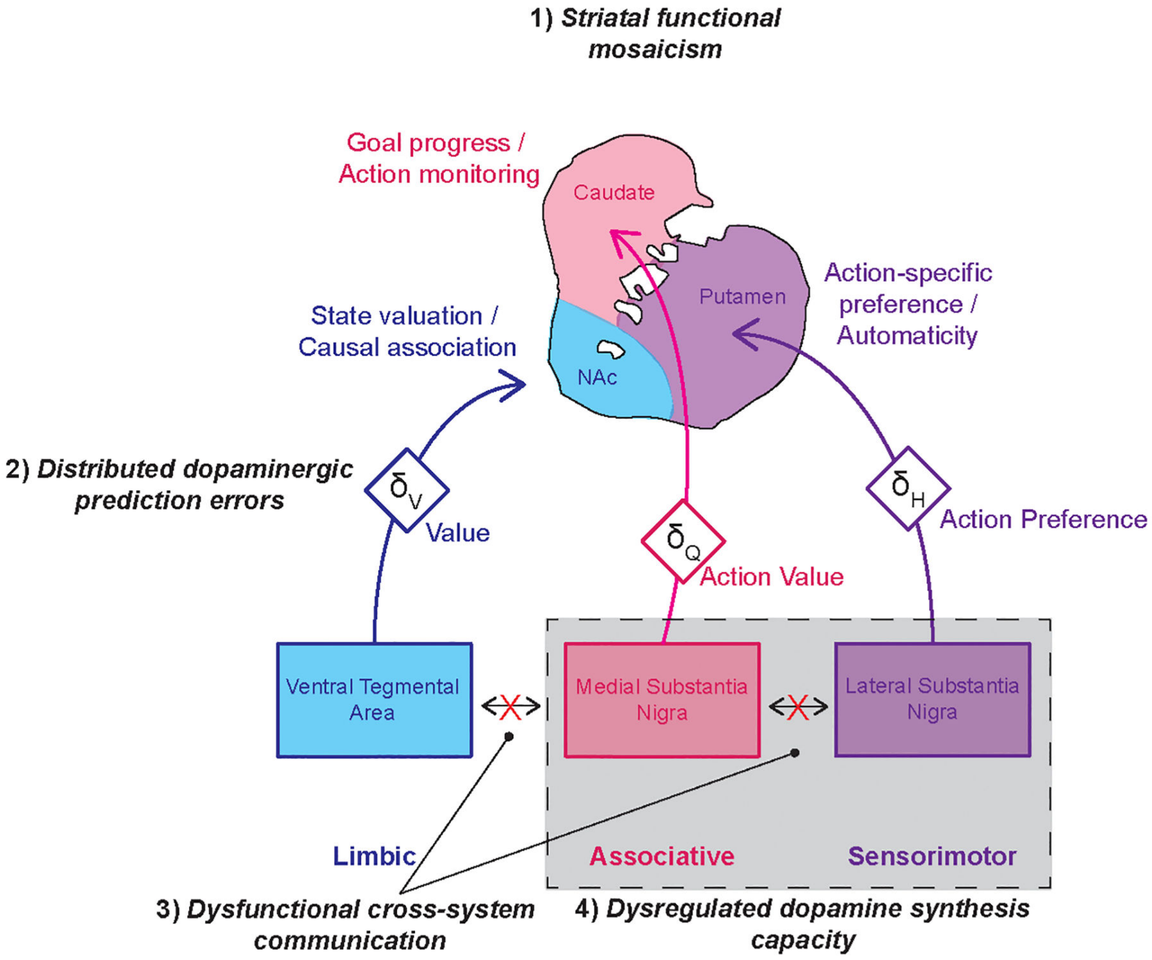

Figure 3. Sources of striatal dysfunction in schizophrenia.

We hypothesize that in schizophrenia and psychosis there is a combination of potential deficits in striatal circuitry and their associated computations that lead to symptomatology. 1) Disruption of typical processing by striatal spiny projection neurons themselves which can result in compartment-specific processing deficits of their associated cognitive functions. 2) Systems specific deficits in prediction error broadcasting by midbrain dopamine neurons which in turn can lead to over- and under-updating of specific priors or predictions that these circuits contribute to. 3) Dysfunctional cross-system communication that results in siloing of computations that are typically coherent and collaborative. For example, a disconnect between the limbic loop and the associative and sensorimotor loops may result in a lack of causal meaning for actions executed and needing to post-hoc assign meaning to one's own behavior that apparently occurs without understanding. 4) Upregulation of dopamine function in the associative and sensorimotor loops that promotes excessive goal-monitoring and action preference but in the absence of the equivalent upregulation for the assignment of meaning or valuation that would underlie such preferences.