

Perspective

Developing algorithmic psychiatry via multi-level spanning computational models

Michael M. Halassa,^{1,2,*} Michael J. Frank,³ Philippa Garety,⁴ Dost Ongur,⁵ Raag D. Airan,^{6,7} Gerard Sanacora,⁸ Kafui Dzirasa,^{9,10} Sahil Suresh,¹ Susan M. Fitzpatrick,¹¹ and Douglas L. Rothman¹²

¹Department of Neuroscience, Tufts University School of Medicine, Boston, MA, USA

²Department of Psychiatry, Tufts University School of Medicine, Boston, MA, USA

³Department of Cognitive and Psychological Sciences, Carney Institute for Brain Sciences, Brown University, Providence, RI, USA

⁴Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

⁵McLean Hospital and Harvard Medical School, Boston, MA, USA

⁶Department of Radiology, Stanford University School of Medicine, Stanford, CA, USA

⁷Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA

⁸Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

⁹Department of Neurobiology, Duke University Medical Center, Durham, NC, USA

¹⁰Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

¹¹James S. McDonnell Foundation, St. Louis, MO, USA

¹²Department of Biomedical Engineering, Radiology and Biomedical Imaging, Yale University, New Haven, CT, USA

*Correspondence: michael.halassa@tufts.edu

<https://doi.org/10.1016/j.xcrm.2025.102094>

SUMMARY

Modern psychiatry faces challenges in translating neurobiological insights into treatments for severe illnesses. The mid-20th century witnessed the rise of molecular mechanisms as pathophysiological and treatment models, with recent holistic proposals keeping this focus unaltered. In this perspective, we explore how psychiatry can utilize systems neuroscience to develop a vertically integrated understanding of brain function to inform treatment. Using schizophrenia as a case study, we discuss scale-related challenges faced by researchers studying molecules, circuits, networks, and cognition and clinicians operating within existing frameworks. We emphasize computation as a bridging language, with algorithmic models like hierarchical predictive processing offering explanatory potential for targeted interventions. Developing such models will not only facilitate new interventions but also optimize combining existing treatments by predicting their multi-level effects. We conclude with the prognosis that the future is bright, but that continued investment in research closely driven by clinical realities will be critical.

INTRODUCTION

After his monthly paliperidone injection, Robert goes back to his room and lies down for the rest of the day. He seldom interacts with others and spends much of his time pacing the hallway. Though he now engages with staff and is less internally preoccupied, he cannot leave the facility he resides in. Before hospitalization, Robert's disruptive behavior strained his family relationships, and they will not take him back. While medications have lessened his psychotic symptoms, they have not restored his functional capacity, leaving him unemployed for years. Like many with schizophrenia, he lives a life of uncertainty, unsure he will ever gain independence, hold down a job, or build meaningful relationships. Robert is all too representative of the many individuals who, despite great effort by researchers and clinicians, remain debilitated by this profound disorder.

As a medical discipline, psychiatry can claim some progress over the past century, emerging from the eras of asylum care into an era marked by the development of several classes of psychoactive pharmaceuticals.^{1,2} Major psychiatric disorders, such

as mood, anxiety, and psychotic disorders, have benefited from drug development, with many overcoming or reducing the frequency of symptomatic episodes that, untreated, would have severely compromised their ability to fully participate in society. More recently, evidence-based psychotherapies and interventions like transcranial magnetic stimulation (TMS) and deep brain stimulation have further improved therapeutic outcomes.^{3–5} These new technologies have been complemented by recent advances in precision neurophysiology and neuroimaging^{6,7} for mental health characterization, and precision noninvasive intervention and functional mapping of the brain.⁸

Despite these successes, schizophrenia remains a large unmet public health need. Schizophrenia is a condition that attacks the very notion of what it is to be a person with agency over our senses and actions; patients often suffer from hallucinations (commonly auditory) and delusions (persistent false beliefs despite contrary evidence). These features, referred to as positive symptoms, co-exist with a host of negative ones in which patients appear apathetic and withdrawn but without a clear affective expression or anxiety. Although antipsychotic



medications do provide some symptom relief (albeit less so for negative ones), almost all patients continue to exhibit cognitive deficits that prohibit their independence. In about 30% of patients, even psychotic symptoms cannot be adequately controlled, leading to “treatment resistance,” which can condemn people to homelessness, state hospital residency, or total reliance on family care. Recently, the Food and Drug Administration approved xanomeline/trospium (Cobenfy), the first non-dopaminergic treatment targeting muscarinic pathways, showing promise for negative and cognitive symptoms.⁹ While exciting, it will take several years before we’re able to evaluate the impact of this advance. Critically, availability of new medications does not fundamentally alter the arguments we present in this article.

In contrast to the slow clinical progress, brain research has advanced rapidly in recent decades, driven by technologies enabling the interrogation of the brain and mind from novel perspectives.¹⁰ The rise of approaches in animals for observing and perturbing the activity of synapses, neurons, and populations within intact circuits and networks has been nothing short of a scientific revolution.¹¹ Guided by human genomics, these studies have identified promising pharmaceutical targets.^{12–13} In addition, neuroimaging methods such as positron emission tomography (PET), electroencephalography (EEG), and magnetic resonance imaging/spectroscopy (MRI/MRS) now allow direct visualization of alterations in the brain chemistry and function in human patients, as well as the impact of new treatments.

Why hasn’t this tremendous progress in research translated to similar advances in patient care? Following the discovery of the first effective antipsychotic medication chlorpromazine, hope was placed on the notion that medications could target specific molecular alterations and correct existing brain chemical imbalances. As research identified abnormalities in dopamine (DA), glutamate, and other neurochemical systems in schizophrenia, drug development focused on these targets. Unfortunately, this approach has not produced genuine breakthroughs, prompting many to declare a crisis in schizophrenia treatment development.^{14,15}

Given these findings, critics increasingly argue that schizophrenia research has become overly reductionist, unable to address the heterogeneous, dynamic nature of mental illness.^{16,17} Such criticism is often directed at a traditional molecular-focused paradigm in which a genomic lesion propagates across multiple levels of organization to impact behavior (Figure 1A). Critics challenge the assumption that higher level abnormalities can be corrected by single molecular interventions at lower levels, emphasizing that many aspects of mental illness, especially interactions between individuals and their social environment, are beyond reductionist approaches. In contrast, advocates of the molecular (now extended to -omics) paradigm have argued that the problem is not too much emphasis on reductionist science but instead not enough, criticizing the alternate “holistic” view as being too vague to lead to useful treatments.^{18,19} In this paper, we advocate a perspective centered around algorithmic, “latent variable” models of behavior. However, unlike pure behavioral state models, we advocate integrating the behavioral description with advances in computational modeling at the level of cells and circuits to improve our

ability to predict the impact of molecular and regional targeted treatments (Figure 1B).

Such models assume a stepwise nature of human cognition as a sequence of unobservable latent states that can be decomposed into a list of computational functions (short term memory, integration, feedback, etc. ...). These models have been developed to describe behavior in various types of tasks, providing advantages over more qualitative descriptors.^{20–22} First, they provide ways to group task states via a common computation instead of simply averaging unrelated states.²³ Second, such state grouping provides a way for correlation with neural measurements, linking brain and behavior through computational language.^{24–26} Third, psychiatry already implicitly uses latent descriptors to guide treatment. For example, when evaluating depressed mood, the astute psychiatrist avoids symptom-triggered treatment using antidepressants and instead clarifies whether this is part of a bipolar syndrome that would require mood stabilizers instead. Psychiatric diagnostic entities address multiple symptoms with a single medication, reducing polypharmacy. Therefore, the adoption of algorithmic models would naturally align with current treatment optimization practices.

Algorithmic behavioral models are at the center of multi-level spanning computational models, which can incorporate data across the levels of organization relevant to psychiatry. The time is ripe for such developments, given the revolution in computational power and deep learning approaches.²⁷ By integrating all the traditional levels of organization in this framework, psychiatry can potentially achieve what it heretofore has lacked—a vertically integrated theory of brain function that links molecular structures with neural functions, neural functions with cognitive operations, and cognitive operations with real-world behavior (Figures 1B and 1C). Consider an active working memory task in which subjects must remember and update a list of items. An algorithmic model may break down this task into latent cognitive states such as maintenance, updating, and response selection. While these latent states are not directly observable, behavioral data allow us to infer them. Below this algorithmic level, we have the command level, which specifies the computational operations that implement these latent states, such as neural integration,²⁸ line,²⁹ or bump attractors.³⁰ These computational operations can be directly tested against empirical neural data and further linked to single-cell neural encoding properties at the level beneath. Critically, out of all these levels, the algorithmic level is the one that is closest to behavior, and therefore, it can have clear inroads to behavioral data fitting on one end and neural modeling on another (Figure 1C).

Because biological data are described in a manner that is not intended to inform psychological function, translating biological knowledge into computational terms is essential to this framework. In fact, the parallel revolution in computer science and artificial intelligence not only provides tools to examine how the underlying neurobiology works but also the appropriate language to link biology with psychology.^{31,32} While the higher level component descriptions of behavior derive from cognitive psychology and therefore are intrinsically describable in the language of stepwise computational processes, advanced neuroimaging studies have not found equivalent neuroanatomical

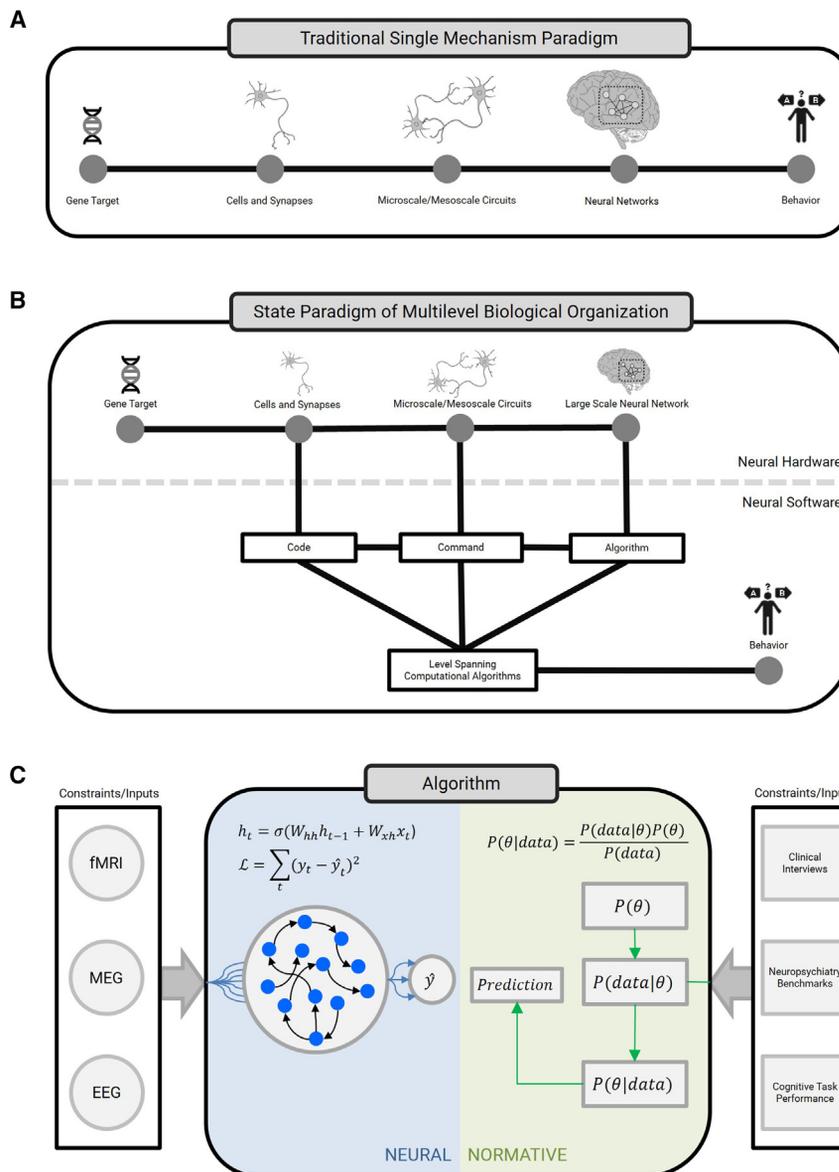


Figure 1. An algorithmic framework for connecting neurobiological substrates to thought and behavior

(A) Traditional paradigm, where genes function within cells, giving rise to circuits that operate within networks, which collectively give rise to behavior. This paradigm emphasizes that genetic lesions can be targeted by (pharmacological) approaches for treatment and implicitly assumes alterations occurring in higher levels will normalize.

(B) State paradigm for multi-level biological organization, emphasizing the parallel nature of computations across levels. The distinction between neural hardware (biological structures) and neural software (functional computations) is illustrated to further illustrate models that span across these levels.

(C) Focus on the algorithmic level, which is closest to observable behavior and offers the most immediate opportunities for altering disease trajectories. Behavior-relevant algorithms can come in a neural form, which involve task-optimized models solving various behavioral tasks and provide comparable data to human subject (or animal) data. Normative models of behavior provide a complementary perspective; they are non-neural and their Bayesian machinery can provide interpretable quantities, which also provide insight into the strategies taken by subjects. We suggest that the integration of these two types of models will be critical for the future of behavioral description in psychiatry, as well as their initial mapping onto neural substrates.

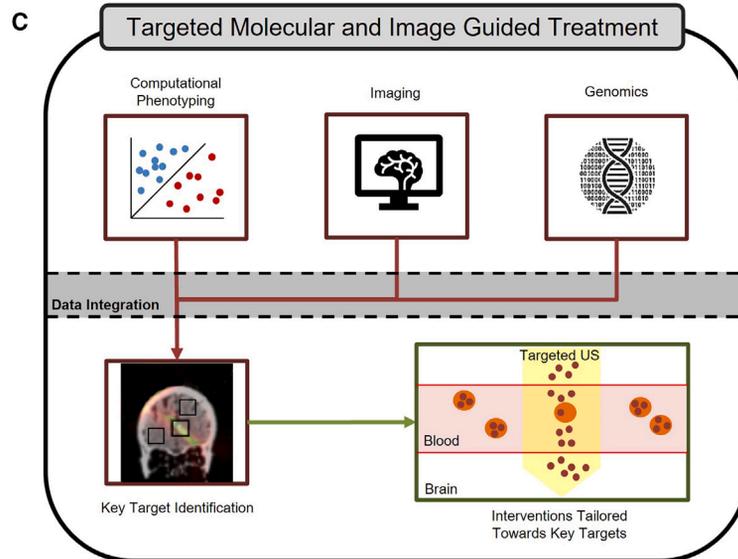
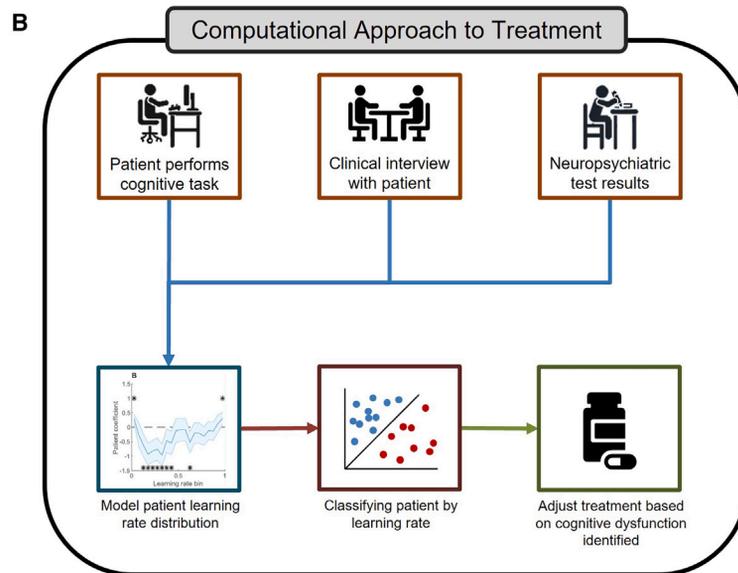
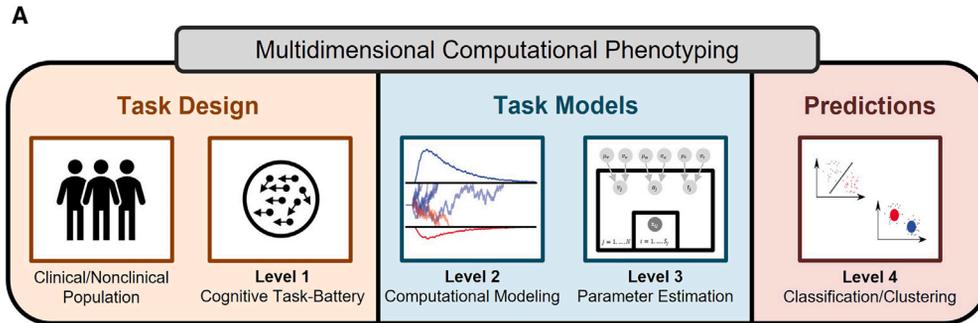
representations.³³ Therefore, even if an impaired computational process is identified by behavioral studies, the anatomical location and biological level for optimal intervention are no longer clear with the loss of a 1:1 function to anatomical location relationship. However, imaging studies have linked behaviors to distributed anatomical networks of neuronal activity. The intrinsically parallel multi-level algorithms of modern deep learning approaches are potentially ideal and likely essential for relating networks to latent variable models of behavior as well as to lower circuit and synapse levels of description (Figure 1C). Additionally, leveraging rapid advances in anatomically targeted therapies, as discussed in Conclusions and future outlook, is crucial.

In the rest of this piece, we first provide an overview for the role of formal computational theory in describing cognitive op-

erations that link to behavior. Next, we illustrate how this approach can augment an ongoing success story that takes inspiration from behavioral economics to target abnormal thinking in schizophrenia. We then move to progressively lower levels of mechanistic description, from large-scale brain networks measured by methods such as fMRI to molecular mechanisms at the level of receptors and genes to describe how new experimental approaches and statistical tools are revolutionizing these areas and building bridges between them. We end with speculations on how all these advances will inform treatment.

PRESENT-LEVEL SPANNING COMPUTATIONAL APPROACHES AND THEIR APPLICATION TO TREATMENT

In this section, we summarize several approaches using computational models to bridge across levels of analysis. The discussion is not meant to be comprehensive, but rather illustrates examples of key achievements and current limitations experienced to date.



(legend on next page)

The role of formal computational theory in schizophrenia, linking cognitive theory to behavior

While much progress has been made in identifying biological alterations in the pathophysiology of schizophrenia, such findings alone do not explain how they translate into functional outcomes or symptoms that characterize schizophrenia and mental illness more broadly. The complexity of systems neuroscience implies that even neurological diseases with “high essentiality” (i.e., well-defined etiology traced to a single or small number of mechanisms^{34,35}) can have complex interactive effects on cognition and behavior. A canonical example is Parkinson’s disease (PD), which results primarily from degeneration within midbrain DA neurons, but which percolates throughout corticostriatal circuits and thus leads to alterations across motor, affective, and cognitive functions.³⁶ While primarily characterized by a set of stereotypical motor symptoms, the cognitive, affective, and perceptual consequences of PD are heterogeneous.^{37,38} This poses a treatment challenge not unlike that seen in schizophrenia, particularly that traditional antipsychotic treatments need to be balanced against worsening motor symptoms. Thus, even a known lesion, such as the degeneration of DA neurons in an identified neuroanatomical location, can yield heterogeneous symptoms whose circuit-level explanation remains an open question and where treatment is just as complex as that of less molecularly understood disorders such as schizophrenia.

Formal computational models can help address the problem of heterogeneity and unify seemingly disparate symptom clusters in PD. They can do so by reducing the complexity of the circuits to their functional considerations shared across such circuits, with analogous effects of DA modulation therein.³⁹ These models, spanning detailed neural circuits to algorithmic descriptions of action selection,⁴⁰ have inspired studies to test their predicted implication of valuation and reinforcement learning in the progression of Parkinsonian symptoms. Such approaches have been applied to both patients with PD and to patients experiencing Parkinsonian symptoms because of antipsychotic treatment in schizophrenia. Predictions of these models have been confirmed in rodent models. For example, the models correctly predicted that motor dysfunction would advance as a function of aberrant learning (exaggerated plasticity) in striatal D2 receptor-expressing spiny neurons and that such deficits can be separated from the direct effects of DA denervation on “performance.”^{41,42} Similarly, they have accounted for analogous effects on cognitive function and decision-making via multiple corticostriatal circuits⁴³ and the paradoxical finding that DA medications can have both therapeutic and deleterious effects depending on task demands.^{40,44}

This same strategy to compress multiple interacting cellular and systems-level circuits into latent low-dimensional models that summarize their core computations (“multidimensional computational phenotyping”) has improved prediction and diagnostics in more complex mental illness,^{45,46} including schizophrenia⁴⁷ and depression⁴⁸ (Figure 2). However, it becomes increasingly difficult to link such functional considerations to a small number of reductionist mechanisms. In the language of Hitchcock et al., schizophrenia represents an illness of low to moderate “essentiality”: there are some consistent findings on the pathophysiology, but it certainly cannot be reduced to a single set of mechanisms. The essence of the illness is more likely to be found in the heterogeneity across patients.³⁴ Thus, successful predictive models have largely taken the “quantitative abductive” approach,^{22,49,50} beginning with algorithmic models of behavior, fitting them to behavioral data, and determining post hoc which model parameters best discriminate between patient populations, without drawing strong conclusions about mechanism.

Nevertheless, sometimes observed parameter differences can plausibly be linked to a neural circuit or set of mechanisms, either because neural recordings were used in the associated study and their variations were linked to model parameters,^{48,51} or because the observed parameter differences have well-established links to underlying mechanisms. As a disorder of moderate essentiality, schizophrenia has long been linked to disturbances of the prefrontal cortex (PFC)⁵² as well as the DA system.^{53,54} Formal models of corticostriatal DA systems have been extended to provide an account of the DA hypothesis of schizophrenia, considering how multiple aspects of positive and negative symptoms may arise from disruptions to DA within individual circuits and on different time scales.⁵⁵ But here one should think of this approach as “how far can one get” with a simple model rather than a comprehensive account of schizophrenia. Indeed, task paradigms, combined with formal models designed to disentangle striatal DA mechanisms from prefrontal cortical contributions, have consistently failed to show clear alterations of DA-dependent reinforcement learning processes in (chronic) schizophrenia and instead attribute alterations in learning deficits in prefrontal cortical working memory processes.^{56,57} Given the extensive interactions among these systems and with hippocampal and other circuits, drawing precise causal conclusions from any single task or patient group remains challenging.

One might argue that such studies often involve chronically medicated patients, whose intact striatal DA-dependent processes could simply reflect normalization by antipsychotics

Figure 2. A framework for a computational approach to treatment guidance

(A) Multidimensional computational phenotyping. Depiction of computational phenotyping beginning with task design where clinical and nonclinical populations undergo task batteries designed to probe specific cognitive functions and capture behavioral data. The data are subjected to modeling to simulate cognitive processes and identify key parameters in performance and output. The parameters are parsed to quantify individual differences in cognitive processes, and the individuals are classified and clustered based on the models to allow for personalized predictions.

(B) Computational approach to treatment. Illustration of how phenotyping can guide treatment decisions. Cognitive task performance from patients is complemented with additional clinical data for modeling, which allows for cognitive behavioral classification. Treatment is adjusted according to the dysfunctions identified in a personalized manner.

(C) Targeted molecular and image-guided treatment. By combining computational phenotyping with advanced functional imaging and genomics, clinicians can design integrated targeted therapies (e.g., using localized ultrasound for targeted drug release).

(D2 receptor antagonists preferential targeting in the striatum rather than the PFC). But despite decades of research and clinical trials, next-generation antipsychotics have seen limited progress. This may be understood from the perspective of next-generation computational models and systems neuroscience approaches highlighting that a single blunt manipulation of DA or other modulators is unlikely to completely ameliorate symptomatology.

Consider that within the DA system, traditional reinforcement learning models and data relied on the notion that DA signals conveying salient “reward prediction error” signals^{58,59} are *scalar* (reflecting a single number that goes up or down) and *global* (influencing all downstream targets across the striatum and elsewhere equally). Recent studies in animal models have challenged this notion however, with endogenous DA signals showing substantial heterogeneity, in part depending on their downstream targets.^{60–63} Moreover, even within localized regions of the dorsal striatum, rewards induce wave-like spatio-temporal dynamics of DA, with the direction of the waves (from medial to lateral or vice versa) dependent on task demands.⁶⁴ When the task requires *agency*—the animal controls how they progress from one state to another before receiving rewards—DA waves initiate within the dorsomedial striatum (DMS). But when task events indicating progress to reward were incongruent with the animals’ actions, the DMS was the last region to receive DA. These opponent wave dynamics were complemented by more gradual DA *ramps* in the DMS that led up to the reward as the animal progressed in the task, again in opponent directions (increasing ramps in the agency condition, decreasing ramps in the passive condition). Finally, brief DA transients in individual axon segments localized to DMS subregions as animals experienced individual events signaling discrete progress toward reward.^{64,65} A “mixture of experts” computational model synthesized how such transients, ramps, and waves interacted in the service of directing “credit assignment” to the appropriate striatal subregion, allowing the animal to learn whether they are in control of task events and how to guide and reinforce behavior accordingly. Model predictions, in terms of how the different DA dynamics were predictive of each other and of subsequent behavioral adjustment, were supported by observed experimental data.

One level up: Taking inspiration from behavioral economics to target abnormal thinking in schizophrenia

What does the aforementioned computational approach mean to Robert, who, like many others, struggles to live to his full potential as a consequence of his condition? We seek answers in biological alterations, but as noted earlier, such findings alone are insufficient to explain how they translate into symptomatic or functional outcomes. Is it useful to ask how his symptom of inactivity relates to his conscious thoughts and feelings, if at all? Suppose Robert reports that he does not go out because he is fearful, believing that he is being followed by people with malicious intent, and the reason for this is that he often hears isolated unidentifiable voices, telling him to “go home.” Staying home feels safer, despite the negative repercussions, and over the years, it has become habitual, further limiting his social functioning. The computational approach means that for individuals

like Robert, we can quantitatively identify and understand the ways in which his brain processes information abnormally.

Cognitive processes and phenomenological experiences form a crucial component of vertically integrated models, linking biological vulnerability, social factors, and the manifestation of individual psychotic symptoms.⁶⁶ Over the past two decades, a number of converging cognitive models of psychosis^{66–69} propose that beliefs and appraisals are central to determining the clinical consequence of psychotic experiences. These cognitive models of psychosis recognize biological alterations as conferring an increased vulnerability to disordered perceptions (e.g., hearing a disembodied voice). However, disordered perceptions alone are not psychotic (for example, benign voices are a relatively common human experience).⁷⁰ How individuals make sense of, and respond to, anomalous experiences that characterize psychosis can determine whether they remain benign or result in impairment. The conclusions that Robert believes the voices he hears derive from individuals who intend harm and follow him about are typically considered delusions. These are distressing and lead him to adhere to dysfunctional behaviors such as asociality. These appraisals, as such models suggest, stem from past experiences (e.g., of childhood trauma or later adversity) influencing content of appraisals (e.g., voices as malign). They are also influenced by biased reasoning, which may be due to cognitive limitations (e.g., working memory).

Reasoning processes have long been considered disrupted in schizophrenia (and were integral to the Diagnostic and Statistical Manual definition of delusions as “false beliefs *due to incorrect inference* about external reality” DSM-III and -IV). Systematic reviews and meta-analyses have detailed a large and consistent evidence base, in over 70 studies, in which the clear majority show that individuals with delusions and psychosis make decisions on the basis of limited evidence in probabilistic reasoning tasks, the so-called “jump-to-conclusions” (JTC) bias.^{71–74} This bias is posited to contribute to delusion formation; thus, it is proposed that anomalous or ambiguous information is rapidly appraised, and a delusional conclusion is drawn on the basis of limited evidence. Additionally, once a conclusion is drawn, even if implausible, it occurs without a thorough consideration of alternatives or a corrective review of the evidence, thus entrenching delusional persistence. This latter process of reviewing one’s initial impressions or beliefs has been termed “belief flexibility,” which research has also shown to be poor in people with delusions.^{73,75} We should note that there are more nuanced views on the JTC bias, highlighting basic differences in information seeking in patients with schizophrenia.⁷⁴

This empirical research finds a clear parallel in the work of the behavioral economists Tversky and Kahneman, who in numerous studies have shown that people do not always make rational or optimal decisions even when the information is available to do so.⁷⁶ They have documented many heuristics and biases in decision-making, which may lead people to irrational actions. Kahneman’s book, “Thinking Fast and Slow” popularized this work and distilled his body of research within the framework of two-process models of reasoning.⁷⁶ In cognitive psychological theory, dual-process models of human reasoning posit two parallel systems or processes underpinning decision-making: type 1, fast, high capacity, independent of working

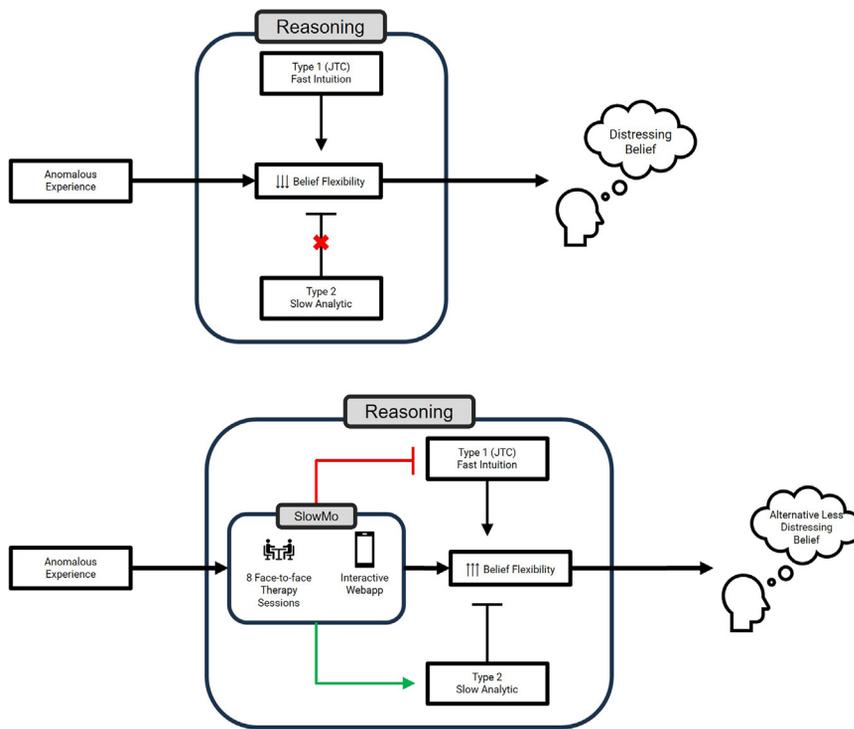


Figure 3. Dual-process model of reasoning

Top: Fast and slow thinking in schizophrenia. A schematic representation of the dominance of fast intuition over slow analytic reasoning as it applies to poor belief flexibility and thus distressing beliefs about anomalous experiences.

Bottom: Technology-assisted approach (SlowMo) for ameliorating reasoning dysfunction in schizophrenia. A schematic representation of the role of the dual-process treatment of SlowMo therapy. Through therapy sessions and a mobile application that acts as an “extended mind,” there is increased resource allocation toward slow, analytic thinking. This overrides the type 1 fast thinking and leads to alternative, less distressing beliefs, consequently mitigating cognitive dysfunction.

The success of SlowMo supports that further work on integration of formal algorithmic modeling with higher cognitive theories that interface with real-world interventions would be of tremendous benefit. For example, computational phenotyping of patients with schizophrenia based on their task behavior, movement, speech patterns, and/or brain signals

memory and cognitive ability; and type 2, slow, low capacity, heavily dependent on working memory and related to individual differences in cognitive ability.^{77–81}

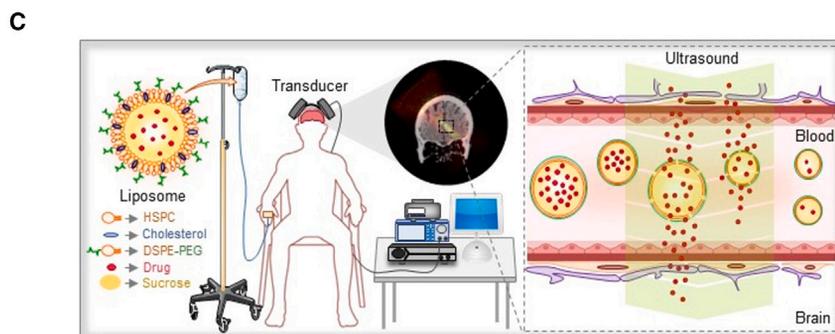
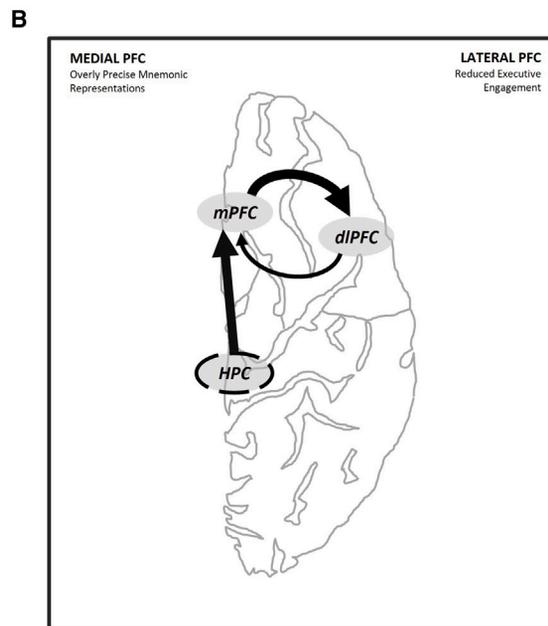
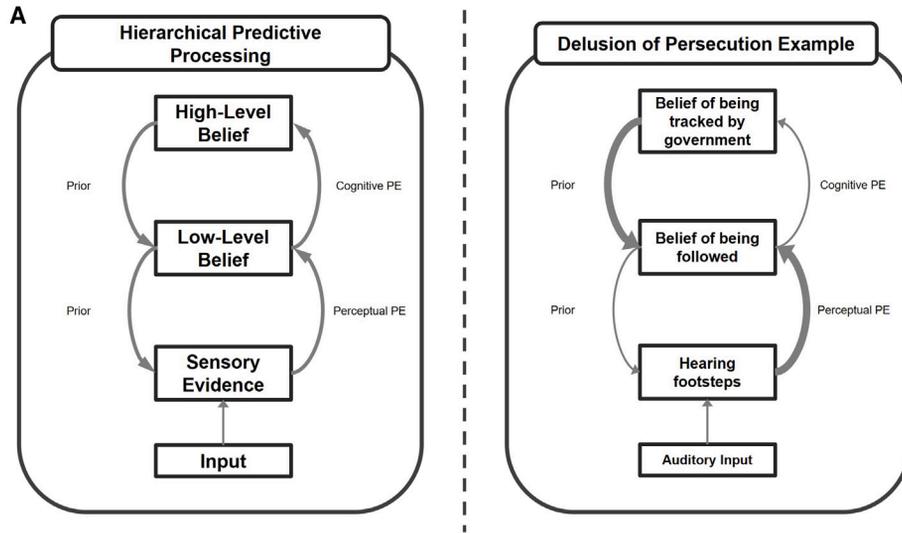
If we consider reasoning in psychosis in this context, it is apparent that JTC may be considered to reflect the operation of type 1 *fast* processes, while the documented failure in people with delusions to correct mistaken judgments from conflicting evidence is a failure of type 2 *slow thinking* (i.e., the ability to consider alternative explanations and review the evidence thoroughly).⁷⁵

Stimulated by this research, a new digitally supported psychological intervention, “SlowMo,” was developed to help patients with schizophrenia manage delusions by targeting reasoning, specifically fast and slow thinking (Figure 3A). This is grounded in cognitive models of psychosis, and cognitive behavior therapy, but focuses intensively on the mechanisms of fast and slow thinking, bringing into a person’s awareness their tendency to jump to conclusions and providing tools for “slowing down” thinking and generating more benign alternative thoughts to the person’s current and distressing persecutory delusion (Figure 3B). SlowMo is a blended therapy with eight face-to-face therapy sessions and is supported by a web app with interactive materials, vignettes, and exercises. A linked mobile app encourages the user to enhance slow thinking in everyday life: providing ready access to a repository of personalized non-delusional safety ideas, previously generated in therapy, thus supporting effective self-management during distressing experiences. These interactive elements can be modified to specifically fit each patient’s reasoning style based on behavioral feedback. A clinical trial⁸² demonstrated significant improvements in persecutory delusions, overall functioning, and well-being and that one mechanism for this was through increased type 2 slow thinking, specifically belief flexibility.

should help identify to what degree fast thinking style is exaggerated versus slow thinking impaired. By capturing the metrics of fast vs. slow thinking—such as decision thresholds, reconsideration times, and belief flexibility—the field can quantify the degree to which each reasoning style contributes to persistent delusional beliefs. These individualized metrics allow SlowMo to be tailored to target the specific cognitive processes most in need of intervention, thus resulting in the design of a personalized intervention, capable of iterative amelioration of psychotic experience. The dual-process model of reasoning can consequently be enriched by bridging these cognitive metrics with biological processes. Controlled paradigms of decision-making are essential to this integration, where subjects are presented with initially ambiguous images that are resolved over time. Under these conditions, patients with schizophrenia indeed jump to fast conclusions.⁸³ However, developing paradigms in which participants may revise their conclusions could reveal additional deficits and formalize the notion of “belief reappraisal” biases in patients with schizophrenia. More broadly, creating paradigms with dissociable individual components facilitates algorithmic modeling of the dynamics of belief formation and revision. This accomplishes two goals: the individualized patient approach outlined previously and the link to the underlying brain mechanisms, as the fast thinking style may align better with BG mechanisms whereas the slow style with prefrontal ones.⁸⁴ This has important implications for predicting treatment responses as well as the development of new interventions.

Predictive processing theory as a bridge to mechanism

A promising set of models that have the capacity to link higher cognition to neural mechanism lies within the predictive



(legend on next page)

processing framework.⁸⁵ Several types of predictive processing theory (PPT) exist, but they share the core idea that perception lies at the interface between memory-based predictions and environmental sensory inputs.⁸⁶ For example, responses in primary sensory cortex are thought to only partially reflect information from the periphery, as some are better explained by deviations from internal predictions.⁸⁷ These findings can be attributed to the highly interconnected nature of cortical areas; higher-order sensory cortex that responds to more complex features generates more statistically probable stimuli based on partial inputs, which then interact with the lower level inputs in primary sensory areas. Deviations from these predictions are transmitted back to higher areas to update their activity (potentially requiring neuromodulators such as DA). Based on this iterative process, PPT has striking parallels to cognitive reappraisals discussed earlier. PPT makes predictions in this domain that have been verified by experiments, such as finding dedicated cortical neurons that respond to quantities that are higher or lower than memory-based predictions.⁸⁷

Given such parallels, one would expect PPT to have important insights into the neural mechanisms of schizophrenia. Indeed, a series of studies have formulated schizophrenia pathophysiology as resulting from imprecise predictive machinery (imprecise priors in Bayesian terms) combined with overly precise sensory data (high likelihoods).^{70,88–91} This results in a higher frequency of prediction errors, which the brain may interpret as hallucinatory experiences. While intriguing, this account introduces a paradox: delusional beliefs are highly resistant to change, suggesting increased rather than decreased weighting of priors. A recent proposal resolves this by extending classical PPT accounts to the cognitive domain through a hierarchical framework.^{92–94} Specifically, reduced precision is localized to the generative process of higher sensory areas but not to the associative areas above them. Instead, these areas may have increased precision, resulting in two types of prediction errors at different levels of hierarchy (Figure 4A). The increased perceptual prediction errors would therefore explain hallucinatory phenomena as in classical PPT, but the reduced cognitive prediction errors would explain the persistence of delusional thought content that cannot be explained by classical PPT. Therefore, hierarchical PPT aligns well with cognitive models of reasoning, mapping fast and slow thinking onto lower and higher hierarchi-

cal levels, and fits with known anatomical and physiological differences between associative and higher sensory cortical areas.⁹⁵

A recent hierarchical PPT algorithmic model learned to model the process of schema formation and error-based learning.⁹⁶ Despite its algorithmic nature, this model makes predictions regarding interactions between the medial (evaluative) and lateral (executive) divisions of the PFC. This is relevant to schizophrenia pathology; one possible instantiation of network imbalance could be between the medial and lateral PFC, with the medial regions driven by overly precise mnemonic representations, whereas lateral regions show reduced activity consistent with a mid-hierarchical deficit. This potential implementation (Figure 4B), which would focus on delusional thought processing, could simultaneously relate to hippocampal lesions resulting in delusions⁹⁷ and the literature on impaired dorsolateral PFC function in schizophrenia.⁹⁸ Thus, hierarchical PPT offers a testable way to link schizophrenia circuitry to symptoms and potential treatments.

Computational models linking small circuits to networks to altered behavior in animal models relevant to schizophrenia

We described previously (in the “The role of formal computational theory in schizophrenia, linking cognitive theory to behavior” section) spatiotemporal DA dynamics across the striatum of the rodent and how it relates to agency and credit assignment.⁶⁴ Such an intricate interplay between DA dynamics at multiple time scales and striatal subregions in the service of learning about agency could have profound implications for schizophrenia. Rather than considering how a single DA signal may be aberrant, such findings imply that it may be critical to control the timing and spatial specificity of DA terminals to assign credit to the correct neural circuit. Speculatively, disruptions in such signaling could give rise to inappropriate causal attribution to the agent’s own underlying corticostriatal circuits (e.g., an outcome that should be attributed to motor activity could be erroneously attributed to cognitive activity or vice versa). In rodents, these findings align with human studies linking DA signaling to learning values associated with agency^{99–101} and with neural network models implicating the need for multidimensional DA signaling for reinforcing appropriate striatal targets.^{102,103} Interestingly, findings across

Figure 4. Hierarchical predictive processing framework and its relevance to delusions

(A) Schematic model of hierarchical predictive processing in both a healthy context (left) and in the presence of a persecutory delusion (right). Arrows depict feedback signaling of priors and feedforward signaling of prediction errors (PEs), with arrow thickness indicating their relative precision. In the healthy model, prediction errors at both perceptual and cognitive levels guide continuous updating of beliefs in a balanced manner. By contrast, in the persecutory delusion example, auditory input, such as the sound of footsteps, leads to a low-level belief of being followed, which informs the high-level belief of being tracked by the government. Here, high-level priors gain excessive precision, compensating for the decreased precision priors of low-level beliefs and thus dominating interpretation of uncertain inputs.

(B) Hypothesized medial lateral prefrontal cortex imbalance in schizophrenia. Schematic depiction of a network imbalance based on the hierarchical PPT algorithm. In this framework, the medial prefrontal cortex (mPFC) becomes dominated by memory-driven signals originating in the hippocampus (HPC), while the dorsolateral prefrontal cortex (dlPFC) shows diminished executive activity indicative of a mid-hierarchical deficit. As a result, evaluative processes in the medial area overshadow the lateral region’s executive functions, potentially giving rise to the delusional thinking observed in schizophrenia.

(C) Ultrasound uncaging in treatment of psychiatric conditions. Identified network imbalances could potentially be targeted through ultrasonic drug uncaging: a patient visiting a clinical setting for usual treatment would receive an intravenous infusion of a liposomal ultrasound-sensitive nanocarrier encapsulating the drug of interest. Then, with an ultrasound transducer system placed on their scalp noninvasively targeting the region of interest, the drug would be released from the carriers as they circulate in the blood volume of the brain region of interest. The workflow is similar to that of transcranial magnetic stimulation (TMS), with the addition of the drug product infusion.

humans,^{104,105} non-human primates¹⁰⁶ and rodents¹⁰⁷ have shown dopaminergic locus of control within frontal thalamocortical networks, distinct from the classic corticostriatal system, augmenting the notion that DA impacts the brain at multiple levels including those responsible for the maintenance and switching of working memory.^{25,50,107} Notably, any blunt pharmacological DA manipulation would not restore the endogenous dynamics directed to particular striatal subregions needed for proper credit assignment.

Frontal thalamocortical systems have gained increasing attention in the pathophysiology of schizophrenia,¹⁰⁸ while highlighting the value of animal models.^{109–111} These studies have amended the classical notion of the thalamus as a relay of peripheral inputs to the cortex, ushering in the notion of thalamic control of functional cortical connectivity.¹¹² The impetus for this change is anchored in the unique structure of the thalamus, being composed of excitatory neurons that lack local recurrent connections.^{113–117} This feature results in long-range input-output connectivity patterns within the thalamus being the primary determinant of its computational properties.^{118,119–122} Specifically, while thalamic structures associated with early sensory processing (e.g., the lateral geniculate) receive strong driving inputs with little convergence, neurons in associative structures such as the mediodorsal thalamus receive convergent inputs from multiple prefrontal regions, enabling their integration as a bona fide thalamic computation. Such integration, whose properties have been modeled as a combination of input convergence and a unique Hebbian learning rule,^{118,119} results in the generation of “summary statistics,” including task context,^{22,123} input uncertainty,¹⁰⁷ and context prediction errors.¹²³ These signals are then projected back to prefrontal regions to control their ongoing dynamics. Of note, while thalamic neurons cannot excite each other directly, they are thought to inhibit one another via the thalamic reticular nucleus.^{124–131}

Drawing from these systems neuroscience approaches,¹³² Huang et al. translated a mouse task designed to probe the role of mediodorsal thalamus in decision-making. The findings showed that mediodorsal neurons encoded cueing uncertainty based on convergent prefrontal inputs and drove prefrontal inhibition in a manner proportional to this summary statistic of signal reliability. Huang et al. found that subjects with schizophrenia showed a behavioral pattern qualitatively identical to optogenetic mediodorsal suppression in mice. What’s more, these behavioral deficits correlated with a more specific functional neuroimaging readout: the functional connectivity between the right mediodorsal thalamus and the right dorsolateral PFC. This identification of a neural readout of conflict-related executive dysfunction in schizophrenia was generalized to a larger dataset in a completely different cohort. The idea that a specific thalamic locus may be a target for intervention in schizophrenia is a tantalizing possibility. Next, we explore promising approaches for achieving this goal.

NEW METHODS FOR STUDYING AND INTERVENING AT MULTIPLE LEVELS

Although large advances in mapping circuits in preclinical models and advances in imaging methods in humans have and

will continue to play an essential role in modeling across levels, direct interventions at these mapped levels in patients have been limited by a lack of interventional approaches. Over the last decade, there has been major progress in bridging this gap, through image-guided technologies that directly interact with brain circuit function including transcranial magnetic and direct or alternating current stimulation (TMS, TDC/TAC, or tDCS or tACS) and focused ultrasound.¹³³ The transcranial electrical and magnetic stimulation techniques are widely used in clinical practice; however, they suffer from poor field penetration to the neural circuits of interest, given their relatively limited depth of penetration.¹³⁴ In contrast, focused ultrasound can precisely reach nearly any brain region of interest, including medial prefrontal and striatal circuits.¹³⁵ The following section describes progress in image-guided focused ultrasound and how its effectiveness may be improved through the parallel development of computational brain state models.

Focused ultrasound as a method for circuit-specific drug targeting

To test hypotheses of how certain neural circuits and receptor systems contribute to schizophrenia, we need more precise controls over neural activity in target brain regions, with spatial, temporal, and receptor specificity. As noted previously, pharmacology provides the mainstay of current therapy for schizophrenia and yields powerful therapeutic effects in most patients, which can be adapted to specifically target a myriad of molecular receptors. However, side effects due to action of the drug at off-target regions are the norm for most pharmacologic treatments, limiting their therapeutic potential and complicating interpretation of how that drug affects the neural circuits and processes underlying schizophrenia for that particular patient.

Ideally, pharmacological treatments would retain their receptor-specific actions while achieving greater spatiotemporal precision, limiting drug exposure outside the target region to reduce side effects. An example of this approach is work by Airan et al. who have developed a technology for on-demand action of drugs at precise brain region of interest, at targeted time periods.¹³⁶ They have created a suite of ultrasound-sensitive nanoparticulate drug carriers that can hold a drug bound until ultrasound induces release of the drug payload.^{137–139} With this system, they leverage focused ultrasound, which can target ultrasound energy to millimeter-scale regions of interest anywhere in the brain, with millisecond precision, to enable on-demand drug treatment at the right time and at the right brain target. In practice, the patient would receive an intravenous infusion of an ultrasound-sensitive drug-loaded nanocarrier while undergoing routine treatment in an ambulatory, intensive outpatient, or inpatient setting. Focused ultrasound would be applied under image guidance to the brain region(s) of interest including the key nodes of our computational framework. Importantly, this approach could be seamlessly integrated into in-person therapy session without disrupting the treatment workflow. Ultrasound would induce drug release from the nanocarrier as it traverses the blood volume of the region of interest. Then the freed drug would enter the brain, crossing the intact blood-brain barrier as it would otherwise normally do, but only at or near the sonicated site (Figure 4C). Airan et al. have validated this approach in

animal models.¹³⁷ They are now advancing toward a first-in-human application of this technology, targeting ketamine delivery to the anterior cingulate for chronic pain treatment.

Integration of focused ultrasound and other circuit-level interventions with algorithmic models

Focused ultrasound used to directly modulate circuit function or target drug delivery can be a powerful method to validate algorithmic models and map their neural substrates. Let's take the idea of paranoid delusions being the product of abnormal hierarchical PPT as suggested earlier. With focused ultrasound and its depth of precise focusing throughout the brain, we can selectively target different nodes of the neural hierarchy to test such hypotheses. By modulating excitatory-inhibitory balance in higher sensory or lower associative areas, it may be possible to enhance cognitive prediction errors sent to higher associative centers. Similarly, reducing activity in regions generating highly precise assumptions (e.g., hippocampus and medial PFC) could also prove beneficial. Crucially, these interventions should not rely solely on symptom reduction but should incorporate cognitive tasks that clarify when and why an approach fails. Therefore, having tasks that formalize slow and fast thinking and/or bias against disconfirmatory evidence would be a welcome approach.¹⁴⁰

THE CHALLENGE OF DELIVERING NEW NEUROSCIENCE INNOVATION TO THE CLINIC

For regulatory approval and insurance coverage by third-party payers, interventions must demonstrate clear benefit in clinical trials. This is complicated by the large magnitude of contextual effects seen in most trials.^{141,142} If, as we posit, there are multiple circuits and pathways, at different scales, likely contributing to both the pathophysiology of disorder and mechanism of actions of treatments, we must also accept that contextual effects including the placebo effect and sequential effects of failed drug interventions can dramatically impact future treatment outcomes. We now know that several components of the placebo effect, expectations, conditioning, and therapeutic alliances, are mediated by, and themselves mediate, circuits involved with many of the core pathophysiological features of mental illness.¹⁴³ Directly related to this, changes in belief updating and prediction error are thought to mediate part of placebo effect as well as play a central role in the pathophysiology of disorders like schizophrenia.^{144–146} It should therefore not be surprising to see large placebo effect sizes with clinical trial designs that focus on treatments targeting similar cognitive processes. When adequately masked, it could be very difficult to distinguish the effects of the drug or device from the behavioral or cognitive interventions alone.

A key challenge in demonstrating the efficacy of novel treatments is isolating the specific mechanisms driving clinical benefits. This issue is exemplified by 3,4-Methylenedioxymethamphetamine (MDMA)-assisted therapy for post-traumatic stress disorder (PTSD), where regulatory agencies struggle to determine whether the benefits stem from the drug's plasticity-enhancing effects, the therapeutic alliance with the treaters, patient expectancy, or, we must be willing to consider, all or

combinations of these factors.^{147,148} Asking such questions go beyond simply proving whether the treatment has superior efficacy over a placebo control. To combine therapeutics targeting various scalar levels of pathology, future clinical trialists must devise ways to disentangle these contributing factors, and regulatory bodies must develop new paradigms for paths to approval. It is also likely that the field will have to come together to better define the data needed to convince third-party payers. Furthermore, considerations must be given to ways of easily introducing these new modalities into existing healthcare delivery systems in a cost-effective manner. We have recently seen several examples of novel treatments such as IN esketamine for treatment-resistant depression that despite strong evidence of efficacy has limited use due to problems with patient access.¹⁴⁹

CONCLUSIONS AND FUTURE OUTLOOK

Joan: Doctor, I can't go back on the clozapine.

Doctor: Clozapine has helped you before, and it's the best option we have for you right now. Can you share your concerns with me?

Joan: It's too much for me. I do think it helps with voices, but I'm not me anymore.

Doctor: Can you expand on that Joan?

Joan: Sure—I totally see that it dulls the voices, and that relaxes me ... somewhat. The issue is that I don't feel the decisions I'm making are really mine anymore. Say I wanted to go out for a walk or go to the store to get ice cream. I normally know how to think about that choice, but not when I'm on clozapine. Do you see what I'm saying Doc?

Doctor: I understand what you're saying, Joan, and think it is important we try to address this, but to be honest, we aren't able to offer a clear description of how your decision-making may be affected with clozapine. We're working hard on this and are making some progress on how decisions are affected in schizophrenia, but we have a long way to go. In the meantime, let's work together on what medications we can use to give you some relief from the voices, but also keep you feeling like yourself and in control.

Contrasting the aforementioned conversation with that of Robert is meant to illustrate the heterogeneity within schizophrenia. There are individuals capable of eloquently describing what they experience, yet research lacks the tools and frameworks to fully capture these nuances. This underscores the need for integrated models that account for the full breadth of lived experiences in ways that facilitate individualized interventions and mitigate issues of "illusory generalizability" where group-based findings fail to reflect individual experiences. By linking computational, neural, and behavioral data, we can develop models that allow for more accurate predictions, applicable to unique manifestations of disease states at the individual level.¹⁵⁰ Moreover, computational models offer the potential to bridge symptomatology and underlying neural mechanisms, providing a clearer understanding of how specific brain circuit disruptions translate into observable behaviors and symptoms.

As we refine these models, they can assist in guiding the selection of personalized interventions—whether pharmacological, neuromodulatory, or behavioral—by identifying strategies that would be effective to individual’s unique cognitive and neural profile. This personalized approach would permit real-time treatment plan adjustments, resulting in better management of schizophrenia’s evolving course.

In this piece, we have provided an integrative perspective that centers algorithmic models of behavior as a critical node between brain science and mental health, using schizophrenia as a case example. This node provides the necessary language to translate symptoms into computational units and portray them in a manner that can be independently targeted with different treatment modalities. While astute clinicians already navigate psychiatric complexities in a similar way, this approach offers a structured framework that accommodates emerging advances in neural measurements and interventions.

We envision a future where integrative, algorithm-driven models revolutionize mental health care, particularly in treating complex conditions such as schizophrenia. The next 20 years hold immense potential for advances in treatment, but realizing this potential requires a commitment to implementing frameworks that capture the wide spectrum of patient experiences. Only through this concerted effort from the larger clinical and research community will we be able to address the needs of individuals like Robert and Joan in profound and transformative ways, ultimately delivering the life-changing solutions that they, and countless others, deserve.

ACKNOWLEDGMENTS

We thank the James S. McDonnell Foundation for hosting the workshop, which inspired the ideas outlined in this perspective. The patient stories depict “composite patients” that are based on real patient encounters.

DECLARATION OF INTERESTS

D.O. has served on advisory board for Rapport Therapeutics.

REFERENCES

1. Braslow, J.T., and Marder, S.R. (2019). History of Psychopharmacology. *Annu. Rev. Clin. Psychol.* *15*, 25–50. <https://doi.org/10.1146/annurev-clinpsy-050718-095514>.
2. Valenstein, E.S. (1986). *Great and Desperate Cures: The Rise and Decline of Psychosurgery and Other Radical Treatments for Mental Illness (Basic Books)*.
3. George, M.S. (2010). Transcranial magnetic stimulation for the treatment of depression. *Expert Rev. Neurother.* *10*, 1761–1772. <https://doi.org/10.1586/ern.10.95>.
4. Lozano, A.M., Lipsman, N., Bergman, H., Brown, P., Chabardes, S., Chang, J.W., Matthews, K., McIntyre, C.C., Schlaepfer, T.E., Schulder, M., et al. (2019). Deep brain stimulation: current challenges and future directions. *Nat. Rev. Neurol.* *15*, 148–160. <https://doi.org/10.1038/s41582-018-0128-2>.
5. Mayberg, H.S., Lozano, A.M., Voon, V., McNeely, H.E., Seminowicz, D., Hamani, C., Schwab, J.M., and Kennedy, S.H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron* *45*, 651–660. <https://doi.org/10.1016/j.neuron.2005.02.014>.
6. Williams, L.M. (2016). Precision psychiatry: a neural circuit taxonomy for depression and anxiety. *Lancet Psychiatry* *3*, 472–480. [https://doi.org/10.1016/S2215-0366\(15\)00579-9](https://doi.org/10.1016/S2215-0366(15)00579-9).
7. Fornito, A., Zalesky, A., and Breakspear, M. (2015). The connectomics of brain disorders. *Nat. Rev. Neurosci.* *16*, 159–172. <https://doi.org/10.1038/nrn3901>.
8. Wang, J.B., Aryal, M., Zhong, Q., Vyas, D.B., and Airan, R.D. (2018). Noninvasive ultrasonic drug uncaging maps whole-brain functional networks. *Neuron* *100*, 728–738.e7. <https://doi.org/10.1016/j.neuron.2018.10.042>.
9. Kaul, I., Sawchak, S., Correll, C.U., Kakar, R., Breier, A., Zhu, H., Miller, A.C., Paul, S.M., and 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. *Lancet* *403*, 160–170. [https://doi.org/10.1016/S0140-6736\(23\)02190-6](https://doi.org/10.1016/S0140-6736(23)02190-6).
10. Deisseroth, K. (2011). Optogenetics. *Nat. Methods* *8*, 26–29. <https://doi.org/10.1038/nmeth.f.324>.
11. Luo, L., Callaway, E.M., and Svoboda, K. (2008). Genetic dissection of neural circuits. *Neuron* *57*, 634–660. <https://doi.org/10.1016/j.neuron.2008.01.002>.
12. Ripke, S., Neale, B.M., Corvin, A., Walters, J.T.R., Farh, K.-H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A., Huang, H., et al. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* *511*, 421–427. <https://doi.org/10.1038/nature13595>.
13. Collins, F.S., and Varmus, H. (2015). A New Initiative on Precision Medicine. *New Engl. J. Med.* *372*, 793–795. <https://doi.org/10.1056/NEJMp1500523>.
14. Insel, T.R. (2010). Rethinking schizophrenia. *Nature* *468*, 187–193. <https://doi.org/10.1038/nature09552>.
15. Buck, S.A., Quincy Erickson-Oberg, M., Logan, R.W., and Freyberg, Z. (2022). Relevance of interactions between dopamine and glutamate neurotransmission in schizophrenia. *Mol. Psychiatry* *27*, 3583–3591. <https://doi.org/10.1038/s41380-022-01649-w>.
16. Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* *160*, 13–23. <https://doi.org/10.1176/appi.ajp.160.1.13>.
17. Buzzoni, M., Tesio, L., and Stuart, M.T. (2022). Holism and Reductionism in the Illness/Disease Debate. In *From Electrons to Elephants and Elections: Exploring the Role of Content and Context*, S. Wuppuluri and I. Stewart, eds. (Springer International Publishing), pp. 743–778. https://doi.org/10.1007/978-3-030-92192-7_39.
18. Hyman, S.E. (2012). Revolution stalled. *Sci. Transl. Med.* *4*, 155cm11. <https://doi.org/10.1126/scitranslmed.3003142>.
19. Hasin, Y., Seldin, M., and Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biol.* *18*, 83. <https://doi.org/10.1186/s13059-017-1215-1>.
20. Ma, W.J., Beck, J.M., Latham, P.E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.* *9*, 1432–1438. <https://doi.org/10.1038/nn1790>.
21. Scott, D.N., Mukherjee, A., Nassar, M.R., and Halassa, M.M. (2024). Thalamocortical architectures for flexible cognition and efficient learning. *Trends Cogn. Sci.* *28*, 739–756. <https://doi.org/10.1016/j.tics.2024.05.006>.
22. Rikhye, R.V., Gilra, A., and Halassa, M.M. (2018). Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nat. Neurosci.* *21*, 1753–1763. <https://doi.org/10.1038/s41593-018-0269-z>.
23. O’Doherty, J.P., Dayan, P., Friston, K., Critchley, H., and Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* *38*, 329–337. [https://doi.org/10.1016/s0896-6273\(03\)00169-7](https://doi.org/10.1016/s0896-6273(03)00169-7).

24. Körding, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247. <https://doi.org/10.1038/nature02169>.
25. Hawkins, G.E., Cavanagh, J.F., Brown, S.D., and Steyvers, M. (2024). Cognitive Models as a Tool to Link Decision Behavior with EEG Signals. In *An Introduction to Model-Based Cognitive Neuroscience*, B.U. Forstmann and B.M. Turner, eds. (Springer International Publishing), pp. 241–263. https://doi.org/10.1007/978-3-031-45271-0_10.
26. Mollakazemiha, M., and Bahrami Samani, E. (2024). A Joint Cognitive Latent Variable Model for Binary Decision-making Tasks and Reaction Time Outcomes. *Ann. Data Sci.* 12, 499–516. <https://doi.org/10.1007/s40745-024-00519-2>.
27. Su, C., Xu, Z., Pathak, J., and Wang, F. (2020). Deep learning in mental health outcome research: a scoping review. *Transl. Psychiatry* 10, 116. <https://doi.org/10.1038/s41398-020-0780-3>.
28. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936. <https://doi.org/10.1152/jn.2001.86.4.1916.1>.
29. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78–84. <https://doi.org/10.1038/nature12742>.
30. Wills, T.J., Lever, C., Cacucci, F., Burgess, N., and O’Keefe, J. (2005). Attractor dynamics in the hippocampal representation of the local environment. *Science* 308, 873–876. <https://doi.org/10.1126/science.1108905>.
31. Jain, V. (2023). How AI could lead to a better understanding of the brain. *Nature* 623, 247–250. <https://doi.org/10.1038/d41586-023-03426-3>.
32. Amunts, K., Knoll, A.C., Lippert, T., Pennartz, C.M.A., Rylvlin, P., Destexhe, A., Jirsa, V.K., D’Angelo, E., and Bjaalie, J.G. (2019). The Human Brain Project—Synergy between neuroscience, computing, informatics, and brain-inspired technologies. *PLoS Biol.* 17, e3000344. <https://doi.org/10.1371/journal.pbio.3000344>.
33. Turner, J.A., and Turner, M.D. (2021). Re-conceptualizing domains in neuroscience, hopes and utopias aside. *Nat. Neurosci.* 24, 1643–1644. <https://doi.org/10.1038/s41593-021-00946-x>.
34. Hitchcock, P.F., Fried, E.I., and Frank, M.J. (2022). Computational Psychiatry Needs Time and Context. *Annu. Rev. Psychol.* 73, 243–270. <https://doi.org/10.1146/annurev-psych-021621-124910>.
35. Okubo, R., Okada, M., and Motomura, E. (2024). Dysfunction of the NMDA Receptor in the Pathophysiology of Schizophrenia and/or the Pathomechanisms of Treatment-Resistant Schizophrenia. *Biomolecules* 14, 1128. <https://doi.org/10.3390/biom14091128>.
36. Robbins, T.W., and Cools, R. (2014). Cognitive deficits in Parkinson’s disease: a cognitive neuroscience perspective. *Mov. Disord.* 29, 597–607. <https://doi.org/10.1002/mds.25853>.
37. Marras, C., and Chaudhuri, K.R. (2016). Nonmotor features of Parkinson’s disease subtypes. *Mov. Disord.* 31, 1095–1102. <https://doi.org/10.1002/mds.26510>.
38. Zhang, Q., Aldridge, G.M., Narayanan, N.S., Anderson, S.W., and Uc, E.Y. (2020). Approach to Cognitive Impairment in Parkinson’s Disease. *Neurotherapeutics* 17, 1495–1510. <https://doi.org/10.1007/s13311-020-00963-x>.
39. Wiecki, T.V., and Frank, M.J. (2010). Neurocomputational models of motor and cognitive deficits in Parkinson’s disease. *Prog. Brain Res.* 183, 275–297. [https://doi.org/10.1016/S0079-6123\(10\)83014-6](https://doi.org/10.1016/S0079-6123(10)83014-6).
40. Jaskir, A., and Frank, M.J. (2023). On the normative advantages of dopamine and striatal opponency for learning and choice. *Elife* 12, e85107. <https://doi.org/10.7554/eLife.85107>.
41. Beeler, J.A., Frank, M.J., McDaid, J., Alexander, E., Turkson, S., Bernardes Sarria, M.S., McGehee, D.S., and Zhuang, X. (2012). A Role for Dopamine-Mediated Learning in the Pathophysiology and Treatment of Parkinson’s Disease. *Cell Rep.* 2, 1747–1761. <https://doi.org/10.1016/j.celrep.2012.11.014>.
42. Cheung, T.H.C., Ding, Y., Zhuang, X., and Kang, U.J. (2023). Learning critically drives parkinsonian motor deficits through imbalanced striatal pathway recruitment. *Proc. Natl. Acad. Sci. USA* 120, e2213093120. <https://doi.org/10.1073/pnas.2213093120>.
43. Westbrook, A., van den Bosch, R., Määttä, J.I., Hofmans, L., Papadopoulos, D., Cools, R., and Frank, M.J. (2020). Dopamine Promotes Cognitive Effort by Biasing the Benefits Versus Costs of Cognitive Work. *Science* 367, 1362–1366. <https://doi.org/10.1126/science.aaz5891>.
44. Frank, M.J., Seeberger, L.C., and O’Reilly, R.C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940–1943. <https://doi.org/10.1126/science.1102941>.
45. Huys, Q.J.M., Daw, N.D., and Dayan, P. (2015). Depression: A Decision-Theoretic Analysis. *Annu. Rev. Neurosci.* 38, 1–23. <https://doi.org/10.1146/annurev-neuro-071714-033928>.
46. Wiecki, T.V., Poland, J., and Frank, M.J. (2015). Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry. *Clin. Psychol. Sci.* 3, 378–399. <https://doi.org/10.1177/2167702614565359>.
47. Geana, A., Barch, D.M., Gold, J.M., Carter, C.S., MacDonald, A.W., 3rd, Ragland, J.D., Silverstein, S.M., and Frank, M.J. (2022). Using Computational Modeling to Capture Schizophrenia-Specific Reinforcement Learning Differences and Their Implications on Patient Classification. *Biol. Psychiatry. Cogn. Neurosci. Neuroimaging* 7, 1035–1046. <https://doi.org/10.1016/j.bpsc.2021.03.017>.
48. Pedersen, M.L., Ironside, M., Amemori, K.I., McGrath, C.L., Kang, M.S., Graybiel, A.M., Pizzagalli, D.A., and Frank, M.J. (2021). Computational phenotyping of brain-behavior dynamics underlying approach-avoidance conflict in major depressive disorder. *PLoS Comput. Biol.* 17, e1008955. <https://doi.org/10.1371/journal.pcbi.1008955>.
49. Maia, T.V., and Frank, M.J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* 14, 154–162. <https://doi.org/10.1038/nn.2723>.
50. Schmitt, L.L., Wimmer, R.D., Nakajima, M., Happ, M., Mofakham, S., and Halassa, M.M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. *Nature* 545, 219–223. <https://doi.org/10.1038/nature22073>.
51. Parr, T., Rikhye, R.V., Halassa, M.M., and Friston, K.J. (2020). Prefrontal Computation as Active Inference. *Cereb. Cortex* 30, 682–695. <https://doi.org/10.1093/cercor/bhz118>.
52. Lewis, D.A., Curley, A.A., Glausier, J.R., and Volk, D.W. (2012). Cortical Parvalbumin Interneurons and Cognitive Dysfunction in Schizophrenia. *Trends Neurosci.* 35, 57–67. <https://doi.org/10.1016/j.tins.2011.10.004>.
53. Toda, M., and Abi-Dargham, A. (2007). Dopamine hypothesis of schizophrenia: making sense of it all. *Curr. Psychiatry Rep.* 9, 329–336. <https://doi.org/10.1007/s11920-007-0041-7>.
54. Howes, O.D., and Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III—The Final Common Pathway. *Schizophr. Bull.* 35, 549–562. <https://doi.org/10.1093/schbul/sbp006>.
55. Maia, T.V., and Frank, M.J. (2017). An Integrative Perspective on the Role of Dopamine in Schizophrenia. *Biol. Psychiatry* 81, 52–66. <https://doi.org/10.1016/j.biopsych.2016.05.021>.
56. Collins, A.G.E., Brown, J.K., Gold, J.M., Waltz, J.A., and Frank, M.J. (2014). Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *J. Neurosci.* 34, 13747–13756. <https://doi.org/10.1523/JNEUROSCI.0989-14.2014>.
57. Gold, J.M., Waltz, J.A., Matveeva, T.M., Kasanova, Z., Strauss, G.P., Herbener, E.S., Collins, A.G.E., and Frank, M.J. (2012). Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch. Gen. Psychiatry* 69, 129–138. <https://doi.org/10.1001/archgenpsychiatry.2011.1269>.
58. Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian

- learning. *J. Neurosci.* 16, 1936–1947. <https://doi.org/10.1523/JNEUROSCI.16-05-01936.1996>.
59. Janssen, L.K., Sescousse, G., Hashemi, M.M., Timmer, M.H.M., ter Huurne, N.P., Geurts, D.E.M., and Cools, R. (2015). Abnormal modulation of reward versus punishment learning by a dopamine D2-receptor antagonist in pathological gamblers. *Psychopharmacology* 232, 3345–3353. <https://doi.org/10.1007/s00213-015-3986-y>.
 60. Lutas, A., Kucukdereli, H., Alturkistani, O., Carty, C., Sugden, A.U., Fernando, K., Diaz, V., Flores-Maldonado, V., and Andermann, M.L. (2019). State-specific gating of salient cues by midbrain dopaminergic input to basal amygdala. *Nat. Neurosci.* 22, 1820–1833. <https://doi.org/10.1038/s41593-019-0506-0>.
 61. Langdon, A.J., Sharpe, M.J., Schoenbaum, G., and Niv, Y. (2018). Model-based predictions for dopamine. *Curr. Opin. Neurobiol.* 49, 1–7. <https://doi.org/10.1016/j.conb.2017.10.006>.
 62. Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W., and Witten, I.B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* 570, 509–513. <https://doi.org/10.1038/s41586-019-1261-9>.
 63. Tsutsui-Kimura, I., Matsumoto, H., Akiti, K., Yamada, M.M., Uchida, N., and Watabe-Uchida, M. (2020). Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *Elife* 9, e62390. <https://doi.org/10.7554/eLife.62390>.
 64. Hamid, A.A., Frank, M.J., and Moore, C.I. (2021). Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. *Cell* 184, 2733–2749.e16. <https://doi.org/10.1016/j.cell.2021.03.046>.
 65. Liu, C., Goel, P., and Kaeser, P.S. (2021). Spatial and temporal scales of dopamine transmission. *Nat. Rev. Neurosci.* 22, 345–358. <https://doi.org/10.1038/s41583-021-00455-7>.
 66. Garety, P.A., Bebbington, P., Fowler, D., Freeman, D., and Kuipers, E. (2007). Implications for neurobiological research of cognitive models of psychosis: a theoretical paper. *Psychol. Med.* 37, 1377–1391. <https://doi.org/10.1017/S003329170700013X>.
 67. Freeman, D., Garety, P.A., Kuipers, E., Fowler, D., and Bebbington, P.E. (2002). A cognitive model of persecutory delusions. *Br. J. Clin. Psychol.* 41, 331–347. <https://doi.org/10.1348/014466502760387461>.
 68. Garety, P.A., Kuipers, E., Fowler, D., Freeman, D., and Bebbington, P.E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychol. Med.* 31, 189–195. <https://doi.org/10.1017/s0033291701003312>.
 69. Morrison, A.P. (2001). The interpretation of intrusions in psychosis: An integrative cognitive approach to hallucinations and delusions. *Behav. Cogn. Psychother.* 29, 257–276. <https://doi.org/10.1017/S1352465801003010>.
 70. Corlett, P.R., Horga, G., Fletcher, P.C., Alderson-Day, B., Schmack, K., and Powers, A.R. (2019). Hallucinations and Strong Priors. *Trends Cogn. Sci.* 23, 114–127. <https://doi.org/10.1016/j.tics.2018.12.001>.
 71. Garety, P.A., and Freeman, D. (2013). The past and future of delusions research: from the inexplicable to the treatable. *Br. J. Psychiatry* 203, 327–333. <https://doi.org/10.1192/bjp.bp.113.126953>.
 72. Dudley, R., Taylor, P., Wickham, S., and Hutton, P. (2016). Psychosis, Delusions and the “Jumping to Conclusions” Reasoning Bias: A Systematic Review and Meta-analysis. *Schizophr. Bull.* 42, 652–665. <https://doi.org/10.1093/schbul/sbv150>.
 73. McLean, B.F., Mattiske, J.K., and Balzan, R.P. (2017). Association of the Jumping to Conclusions and Evidence Integration Biases With Delusions in Psychosis: A Detailed Meta-analysis. *Schizophr. Bull.* 43, 344–354. <https://doi.org/10.1093/schbul/sbw056>.
 74. So, S.H.W., Siu, N.Y.F., Wong, H.L., Chan, W., and Garety, P.A. (2016). Jumping to conclusions’ data-gathering bias in psychosis and other psychiatric disorders — Two meta-analyses of comparisons between patients and healthy individuals. *Clin. Psychol. Rev.* 46, 151–167. <https://doi.org/10.1016/j.cpr.2016.05.001>.
 75. Ward, T., and Garety, P.A. (2019). Fast and slow thinking in distressing delusions: A review of the literature and implications for targeted therapy. *Schizophr. Res.* 203, 80–87. <https://doi.org/10.1016/j.schres.2017.08.045>.
 76. Tversky, A., and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185, 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>.
 77. Kahneman, D. (2013). *Thinking, Fast and Slow* 1st Pbk (Farrar, Straus and Giroux).
 78. Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *Am. Psychol.* 49, 709–724. <https://doi.org/10.1037/0003-066x.49.8.709>.
 79. Evans, J.S.B.T., and Over, D.E. (1996). *Rationality and Reasoning* (Psychology Press).
 80. Evans, J.S.B.T., and Stanovich, K.E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* 8, 223–241. <https://doi.org/10.1177/1745691612460685>.
 81. Stanovich, K.E., and West, R.F. (2000). Individual differences in reasoning: implications for the rationality debate? *Behav. Brain Sci.* 23, 645–726. <https://doi.org/10.1017/s0140525x00003435>.
 82. Garety, P., Ward, T., Emsley, R., Greenwood, K., Freeman, D., Fowler, D., Kuipers, E., Bebbington, P., Rus-Calafell, M., McGourty, A., et al. (2021). Effects of SlowMo, a Blended Digital Therapy Targeting Reasoning, on Paranoia Among People With Psychosis: A Randomized Clinical Trial. *JAMA Psychiatry* 78, 714–725. <https://doi.org/10.1001/jamapsychiatry.2021.0326>.
 83. Woodward, T.S., Buchy, L., Moritz, S., and Liotti, M. (2007). A Bias Against Disconfirmatory Evidence Is Associated With Delusion Prone-ness in a Nonclinical Sample. *Schizophr. Bull.* 33, 1023–1028. <https://doi.org/10.1093/schbul/sbm013>.
 84. Frank, M.J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J. Cogn. Neurosci.* 17, 51–72. <https://doi.org/10.1162/089929052880093>.
 85. Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. <https://doi.org/10.1038/4580>.
 86. Keller, G.B., Bonhoeffer, T., and Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron* 74, 809–815. <https://doi.org/10.1016/j.neuron.2012.03.040>.
 87. Keller, G.B., and Msrhc-Flogel, T.D. (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron* 100, 424–435. <https://doi.org/10.1016/j.neuron.2018.10.003>.
 88. Fletcher, P.C., and Frith, C.D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. <https://doi.org/10.1038/nrn2536>.
 89. Sterzer, P., Adams, R.A., Fletcher, P., Frith, C., Lawrie, S.M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., and Corlett, P.R. (2018). The Predictive Coding Account of Psychosis. *Biol. Psychiatry* 84, 634–643. <https://doi.org/10.1016/j.biopsych.2018.05.015>.
 90. Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. <https://doi.org/10.1038/nrn2787>.
 91. Powers, A.R., Mathys, C., and Corlett, P.R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* 357, 596–600. <https://doi.org/10.1126/science.aan3458>.
 92. Petrovic, P., and Sterzer, P. (2023). Resolving the Delusion Paradox. *Schizophr. Bull.* 49, 1425–1436. <https://doi.org/10.1093/schbul/sbad084>.
 93. Friston, K., Brown, H.R., Siemerkus, J., and Stephan, K.E. (2016). The dysconnection hypothesis (2016). *Schizophr. Res.* 176, 83–94. <https://doi.org/10.1016/j.schres.2016.07.014>.
 94. Adams, R.A., Huys, Q.J.M., and Roiser, J.P. (2016). Computational Psychiatry: towards a mathematically informed understanding of mental

- illness. *J. Neurol. Neurosurg. Psychiatry* 87, 53–63. <https://doi.org/10.1136/jnnp-2015-310737>.
95. Palomero-Gallagher, N., Eickhoff, S.B., Hoffstaedter, F., Schleicher, A., Mohlberg, H., Vogt, B.A., Amunts, K., and Zilles, K. (2015). Functional organization of human subgenual cortical areas: Relationship between architectonical segregation and connectional heterogeneity. *Neuroimage* 115, 177–190. <https://doi.org/10.1016/j.neuroimage.2015.04.053>.
 96. Alexander, W.H., and Womelsdorf, T. (2021). Interactions of Medial and Lateral Prefrontal Cortex in Hierarchical Predictive Coding. *Front. Comput. Neurosci.* 15, 605271. <https://doi.org/10.3389/fncom.2021.605271.1>.
 97. Pines, A.R., Frandsen, S.B., Drew, W., Meyer, G.M., Howard, C., Palm, S.T., Schaper, F.L.W.V.J., Lin, C., Butenko, K., Ferguson, M.A., et al. (2024). Mapping Lesions that Cause Psychosis to a Human Brain Circuit and Proposed Stimulation Target. Preprint at bioRxiv. <https://doi.org/10.1101/2024.04.28.591471.1>.
 98. Lewis, D.A., and Gonzalez-Burgos, G. (2006). Pathophysiologically based treatment interventions in schizophrenia. *Nat. Med.* 12, 1016–1022. <https://doi.org/10.1038/nm1478>.
 99. Cockburn, J., Collins, A.G.E., and Frank, M.J. (2014). A reinforcement learning mechanism responsible for the valuation of free choice. *Neuron* 83, 551–557. <https://doi.org/10.1016/j.neuron.2014.06.035>.
 100. Fraser, K.M., Pribut, H.J., Janak, P.H., and Keiflin, R. (2023). From Prediction to Action: Dissociable Roles of Ventral Tegmental Area and Substantia Nigra Dopamine Neurons in Instrumental Reinforcement. *J. Neurosci.* 43, 3895–3908. <https://doi.org/10.1523/JNEUROSCI.0028-23.2023>.
 101. Bech, P., Crochet, S., Dard, R., Ghaderi, P., Liu, Y., Malekzadeh, M., Petersen, C.C.H., Pulin, M., Renard, A., and Sourmpis, C. (2023). Striatal Dopamine Signals and Reward Learning. *Function* 4, zqad056. <https://doi.org/10.1093/function/zqad056>.
 102. O'Reilly, R.C., and Frank, M.J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* 18, 283–328. <https://doi.org/10.1162/089976606775093909>.
 103. Frank, M.J., and Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* 22, 509–526. <https://doi.org/10.1093/cercor/bhr114>.
 104. Alelú-Paz, R., and Giménez-Amaya, J.M. (2008). The mediodorsal thalamic nucleus and schizophrenia. *J. Psychiatry Neurosci.* 33, 489–498.
 105. Kessler, R.M., Woodward, N.D., Riccardi, P., Li, R., Ansari, M.S., Anderson, S., Dawant, B., Zald, D., and Meltzer, H.Y. (2009). Dopamine D2 receptor levels in striatum, thalamus, substantia nigra, limbic regions, and cortex in schizophrenic subjects. *Biol. Psychiatry* 65, 1024–1031. <https://doi.org/10.1016/j.biopsych.2008.12.029>.
 106. Sánchez-González, M.A., García-Cabezas, M.A., Rico, B., and Cavada, C. (2005). The primate thalamus is a key target for brain dopamine. *J. Neurosci.* 25, 6076–6083. <https://doi.org/10.1523/JNEUROSCI.0968-05.2005>.
 107. Mukherjee, A., Lam, N.H., Wimmer, R.D., and Halassa, M.M. (2021). Thalamic circuits for independent control of prefrontal signal and noise. *Nature* 600, 100–104. <https://doi.org/10.1038/s41586-021-04056-3>.
 108. Anticevic, A., and Halassa, M.M. (2023). The thalamus in psychosis spectrum disorder. *Front. Neurosci.* 17, 1163600. <https://doi.org/10.3389/fnins.2023.1163600>.
 109. Mukherjee, A., and Halassa, M.M. (2024). The Associative Thalamus: A Switchboard for Cortical Operations and a Promising Target for Schizophrenia. *Neuroscientist* 30, 132–147. <https://doi.org/10.1177/10738584221112861>.
 110. Halassa, M.M. (2020). Genetic variability of memory performance is explained by differences in the brain's thalamus. *Nature* 587, 549–550. <https://doi.org/10.1038/d41586-020-03195-3>.
 111. Halassa, M.M. (2018). Fronto-thalamic Architectures for Cognitive Algorithms. *Neuron* 98, 237–239. <https://doi.org/10.1016/j.neuron.2018.04.006>.
 112. Nakajima, M., and Halassa, M.M. (2017). Thalamic control of functional cortical connectivity. *Curr. Opin. Neurobiol.* 44, 127–131. <https://doi.org/10.1016/j.conb.2017.04.001>.
 113. Rikhye, R.V., Wimmer, R.D., and Halassa, M.M. (2018). Toward an Integrative Theory of Thalamic Function. *Annu. Rev. Neurosci.* 41, 163–183. <https://doi.org/10.1146/annurev-neuro-080317-062144>.
 114. Halassa, M.M., and Kastner, S. (2017). Thalamic functions in distributed cognitive control. *Nat. Neurosci.* 20, 1669–1679. <https://doi.org/10.1038/s41593-017-0020-1>.
 115. Halassa, M.M., and Sherman, S.M. (2019). Thalamicocortical Circuit Motifs: A General Framework. *Neuron* 103, 762–770. <https://doi.org/10.1016/j.neuron.2019.06.005>.
 116. (2022). *The Thalamus*. In *The Thalamus*, M.M. Halassa, ed. (Cambridge University Press), p. iii.
 117. Wolff, M., and Halassa, M.M. (2024). The mediodorsal thalamus in executive control. *Neuron* 112, 893–908. <https://doi.org/10.1016/j.neuron.2024.01.002>.
 118. Hummos, A., Wang, B.A., Drammis, S., Halassa, M.M., and Pleger, B. (2022). Thalamic regulation of frontal interactions in human cognitive flexibility. *PLoS Comput. Biol.* 18, e1010500. <https://doi.org/10.1371/journal.pcbi.1010500>.
 119. Zheng, W.-L., Wu, Z., Hummos, A., Yang, G.R., and Halassa, M.M. (2024). Rapid context inference in a thalamocortical model using recurrent neural networks. *Nat. Commun.* 15, 8275. <https://doi.org/10.1038/s41467-024-52289-3>.
 120. Zhang, X., Mukherjee, A., Halassa, M.M., and Chen, Z.S. (2025). Mediodorsal thalamus regulates task uncertainty to enable cognitive flexibility. *Nat. Commun.* 16, 2640. <https://doi.org/10.1038/s41467-025-58011-1>.
 121. Wang, M.B., and Halassa, M.M. (2022). Thalamicocortical contribution to flexible learning in neural systems. *Netw. Neurosci.* 6, 980–997. https://doi.org/10.1162/netn_a_00235.
 122. Roy, D.S., Zhang, Y., Halassa, M.M., and Feng, G. (2022). Thalamic subnetworks as units of function. *Nat. Neurosci.* 25, 140–153. <https://doi.org/10.1038/s41593-021-00996-1>.
 123. Lam, N.H., Mukherjee, A., Wimmer, R.D., Nassar, M.R., Chen, Z., and Halassa, M.M. (2025). Prefrontal transthalamic processing of uncertainty drives flexible switching. *Nature* 637, 127–136.
 124. Halassa, M.M., Siegle, J.H., Ritt, J.T., Ting, J.T., Feng, G., and Moore, C.I. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. *Nat. Neurosci.* 14, 1118–1120. <https://doi.org/10.1038/nn.2880>.
 125. Halassa, M.M., and Acsády, L. (2016). Thalamic Inhibition: Diverse Sources, Diverse Scales. *Trends Neurosci.* 39, 680–693. <https://doi.org/10.1016/j.tins.2016.08.001>.
 126. Halassa, M.M., Chen, Z., Wimmer, R.D., Brunetti, P.M., Zhao, S., Zikopoulos, B., Wang, F., Brown, E.N., and Wilson, M.A. (2014). State-dependent architecture of thalamic reticular subnetworks. *Cell* 158, 808–821. <https://doi.org/10.1016/j.cell.2014.06.025>.
 127. Wimmer, R.D., Schmitt, L.I., Davidson, T.J., Nakajima, M., Deisseroth, K., and Halassa, M.M. (2015). Thalamic control of sensory selection in divided attention. *Nature* 526, 705–709. <https://doi.org/10.1038/nature15398>.
 128. Nakajima, M., Schmitt, L.I., and Halassa, M.M. (2019). Prefrontal Cortex Regulates Sensory Filtering through a Basal Ganglia-to-Thalamus Pathway. *Neuron* 103, 445–458.e10. <https://doi.org/10.1016/j.neuron.2019.05.026>.
 129. Nakajima, M., Schmitt, L.I., Feng, G., and Halassa, M.M. (2019). Combinatorial Targeting of Distributed Forebrain Networks Reverses Noise Hypersensitivity in a Model of Autism Spectrum Disorder. *Neuron* 104, 488–500.e11. <https://doi.org/10.1016/j.neuron.2019.09.040>.

130. Krol, A., Wimmer, R.D., Halassa, M.M., and Feng, G. (2018). Thalamic Reticular Dysfunction as a Circuit Endophenotype in Neurodevelopmental Disorders. *Neuron* 98, 282–295. <https://doi.org/10.1016/j.neuron.2018.03.021>.
131. Chen, Z., Wimmer, R.D., Wilson, M.A., and Halassa, M.M. (2015). Thalamic Circuit Mechanisms Link Sensory Processing in Sleep and Attention. *Front. Neural Circ.* 9, 83. <https://doi.org/10.3389/fncir.2015.00083>.
132. Huang, A.S., Wimmer, R., Woodward, N.D., Lam, N.H., Suresh, S., Wang, B.A., and Halassa, M. (2024). A Prefrontal thalamocortical readout for conflict-related executive dysfunction in schizophrenia. *Cell Rep. Med.* 5, 101802. <https://doi.org/10.1016/j.xcrm.2024.101802>.
133. Badran, B.W., and Peng, X. (2024). Transcranial focused ultrasound (tFUS): a promising noninvasive deep brain stimulation approach for pain. *Neuropsychopharmacology* 49, 351–352. <https://doi.org/10.1038/s41386-023-01699-w>.
134. Alekseichuk, I., Mantell, K., Shirinpour, S., and Opitez, A. (2019). Comparative modeling of transcranial magnetic and electric stimulation in mouse, monkey, and human. *Neuroimage* 194, 136–148. <https://doi.org/10.1016/j.neuroimage.2019.03.044>.
135. O'Reilly, M.A. (2024). Exploiting the mechanical effects of ultrasound for noninvasive therapy. *Science* 385, eadp7206. <https://doi.org/10.1126/science.adp7206>.
136. Airan, R. (2017). Neuromodulation with nanoparticles. *Science* 357, 465. <https://doi.org/10.1126/science.aao1200>.
137. Airan, R.D., Meyer, R.A., Ellens, N.P.K., Rhodes, K.R., Farahani, K., Pomper, M.G., Kadam, S.D., and Green, J.J. (2017). Noninvasive Targeted Transcranial Neuromodulation via Focused Ultrasound Gated Drug Release from Nanoemulsions. *Nano Lett.* 17, 652–659. <https://doi.org/10.1021/acs.nanolett.6b03517>.
138. Purohit, M.P., Roy, K.S., Xiang, Y., Yu, B.J., Azadian, M.M., Muwanga, G., Hart, A.R., Taoube, A.K., Lopez, D.G., and Airan, R.D. (2023). Acoustomechanically activatable liposomes for ultrasonic drug uncaging. Preprint at bioRxiv. <https://doi.org/10.1101/2023.10.23.563690>.
139. Zhong, Q., Yoon, B.C., Aryal, M., Wang, J.B., Illovitsh, T., Baikoghli, M.A., Hosseini-Nassab, N., Karthik, A., Cheng, R.H., Ferrara, K.W., and Airan, R.D. (2019). Polymeric perfluorocarbon nanoemulsions are ultrasound-activated wireless drug infusion catheters. *Biomaterials* 206, 73–86. <https://doi.org/10.1016/j.biomaterials.2019.03.021>.
140. Lee, M.H., Smyser, C.D., and Shimony, J.S. (2013). Resting-State fMRI: A Review of Methods and Clinical Applications. *Am. J. Neuroradiol.* 34, 1866–1872. <https://www.ajnr.org/content/34/10/1866>.
141. Huneke, N.T.M., Amin, J., Baldwin, D.S., Bellato, A., Brandt, V., Chamberlain, S.R., Correll, C.U., Eudave, L., Garner, M., Gosling, C.J., et al. (2024). Placebo effects in randomized trials of pharmacological and neurostimulation interventions for mental disorders: An umbrella review. *Mol. Psychiatry* 29, 3915–3925. <https://doi.org/10.1038/s41380-024-02638-x>.
142. Stone, A.A., Schwartz, J.E., Broderick, J.E., and Deaton, A. (2010). A snapshot of the age distribution of psychological well-being in the United States. *Proc. Natl. Acad. Sci. USA* 107, 9985–9990. <https://doi.org/10.1073/pnas.1003744107>.
143. Enck, P., and Zipfel, S. (2019). Placebo Effects in Psychotherapy: A Framework. *Front. Psychiatry* 10, 456. <https://doi.org/10.3389/fpsy.2019.00456>.
144. Ghane, M., and Sparer, T. (2019). An Integrative and Mechanistic Model of Impaired Belief Updating in Schizophrenia. *J. Neurosci.* 39, 5630–5633. <https://doi.org/10.1523/JNEUROSCI.0002-19.2019>.
145. Ongaro, G., and Kaptchuk, T.J. (2019). Symptom perception, placebo effects, and the Bayesian brain. *Pain* 160, 1–4. <https://doi.org/10.1097/j.pain.0000000000001367>.
146. Wager, T.D., and Atlas, L.Y. (2015). The neuroscience of placebo effects: connecting context, learning and health. *Nat. Rev. Neurosci.* 16, 403–418. <https://doi.org/10.1038/nrn3976>.
147. Mitchell, J.M., Ot'alora, G.M., van der Kolk, B., Shannon, S., Bogenschutz, M., Gelfand, Y., Paleos, C., Nicholas, C.R., Quevedo, S., Balliett, B., et al. (2023). MDMA-assisted therapy for moderate to severe PTSD: a randomized, placebo-controlled phase 3 trial. *Nat. Med.* 29, 2473–2480. <https://doi.org/10.1038/s41591-023-02565-4>.
148. Nichols, D.E., Johnson, M.W., and Nichols, C.D. (2017). Psychedelics as Medicines: An Emerging New Paradigm. *Clin. Pharmacol. Ther.* 101, 209–219. <https://doi.org/10.1002/cpt.557>.
149. Daly, E.J., Trivedi, M.H., Janik, A., Li, H., Zhang, Y., Li, X., Lane, R., Lim, P., Duca, A.R., Hough, D., et al. (2019). Efficacy of Esketamine Nasal Spray Plus Oral Antidepressant Treatment for Relapse Prevention in Patients With Treatment-Resistant Depression: A Randomized Clinical Trial. *JAMA Psychiatry* 76, 893–903. <https://doi.org/10.1001/jamapsychiatry.2019.1189>.
150. Chekroud, A.M., Hawrilenko, M., Loho, H., Bondar, J., Gueorguieva, R., Hasan, A., Kambeitz, J., Corlett, P.R., Koutsouleris, N., Krumholz, H.M., et al. (2024). Illusory generalizability of clinical prediction models. *Science* 383, 164–167. <https://doi.org/10.1126/science.adg8538>.