BRAIN REPORT

N-methyl-d-aspartate receptor hypofunction causes recurrent and transient failures of perceptual inference

Veith Weilnhammer,[1,2,3,†] Marcus Rothkirch,[1,4,†] Deniz Yilmaz,[1,5,6,7,†] Merve Fritsch,[1] Lena Esther Ptasczynski,[1,5] Katrin Reichenbach,[1] Lukas Roediger,[1] Philip Corlett[8] and Philipp Sterzer[9]

[†]These authors contributed equally to this work.

Perception integrates external sensory signals with internal predictions that reflect prior knowledge about the world. Previous research suggests that this integration is governed by slow alternations between an external mode, driven by sensory signals, and an internal mode, shaped by prior knowledge.

Using a double-blind, placebo-controlled, cross-over experiment in healthy human participants, we investigated the effects of the N-methyl-D-aspartate receptor (NMDAR) antagonist S-ketamine on the balance between external and internal modes.

We found that S-ketamine causes a shift of perception towards the external mode. A case-control study revealed that individuals with paranoid schizophrenia, a disorder repeatedly associated with NMDAR hypofunction, spend more time in the external mode. This NMDAR-dependent increase in the external mode suggests that the symptoms of schizophrenia are caused by recurring dissociations of perception from prior knowledge about the world.

1 Department of Psychiatry and Neurosciences, Charité-Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin 10117, Germany
2 Berlin Institute of Health, Charité-Universitätsmedizin Berlin and Max Delbrück Center, Berlin 10178, Germany
3 Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA 94720, USA
4 Medical School Berlin, Hochschule für Gesundheit und Medizin, Berlin 14197, Germany
5 Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin 10117, Germany
6 Max Planck School of Cognition, Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany
7 Department of Psychiatry and Psychotherapy, LMU University Hospital, Munich 80336, Germany
8 Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511, USA
9 Department of Psychiatry (UPK), University of Basel, Basel 4001, Switzerland

Correspondence to: Veith Weilnhammer
Helen Wills Neuroscience Institute, University of California Berkeley
175 Li Ka Shing Center, Berkeley, CA 94720, USA
E-mail: veith.weilnhammer@gmail.com

Keywords: schizophrenia; psychosis; ketamine; predictive processing; bistable perception; visual perception

Received July 04, 2024. Revised November 06, 2024. Accepted December 03, 2024. Advance access publication January 16, 2025

Introduction

Imagine a dimly lit room at a crowded party, where unclear visual signals, indistinct sounds and complex social interactions allow for multiple—and sometimes false—interpretations. In such ambiguity, failures of perceptual inference, the ability to contextualize sensory inputs with prior knowledge about the world, can lead to profound departures from reality: faces obscured in shadow may appear distorted, random noise could be perceived as a whisper, and friendly smiles might seem derogatory.

According to the canonical predictive processing hypothesis,[1] a disruption of perceptual inference is likely to play a crucial role in schizophrenia, a severe mental disorder characterized by psychotic symptoms, such as delusions and hallucinations.[1] People with schizophrenia may fail to apply prior knowledge to the interpretation of ambiguous sensory signals, causing erratic inferences that lead to hallucinatory experiences and delusional beliefs.[1] Despite considerable progress in the computational understanding of psychosis, two key questions have remained unanswered.

The first question concerns the neural mechanisms that cause perceptual inference to fail in schizophrenia. Formal predictive processing accounts of schizophrenia foreground the role of prediction errors in updating Bayesian beliefs about the causes of sensory input.[2] Most accounts focus on a failure to predict or instantiate the precision afforded to prediction errors at various levels of the cortical hierarchy.[1] Precision refers to the confidence ascribed to prediction errors, and regulates how prior expectations are updated in response to sensory information.[2] Mathematically, precision is equivalent to the (Kalman) gain or the weighting of prediction errors in predictive processing models of perceptual inference.[3] Psychologically, the deployment of sensory precision can be understood in terms of selective attention (or sensory attenuation).[4] Physiologically, precision corresponds to the postsynaptic gain or excitability of neuronal populations that report prediction errors, commonly mediated by *N*-methyl-D-aspartate receptors[5] (NMDARs).

Beyond predictive processing theory, several lines of evidence point to NMDAR hypofunction as a key factor in the pathophysiology of psychosis.[6] NMDAR antibodies[7] and antagonists, such as ketamine,[8] mimic the symptoms of schizophrenia, which is itself associated with a reduction of NMDAR density in the prefrontal cortex.[9] In addition to their role in controlling the excitability of prediction error neurons[6] and their general function for maintaining the cortical excitation-inhibition balance,[10] NMDARs play a critical role in cortical feedback,[11] support synaptic short-term plasticity,[12] and interact with neuromodulators, such as dopamine and serotonin, via GABAergic interneurons.[13] While these NMDAR-dependent mechanisms are likely critical for perceptual inference, it is yet to be determined how NMDAR hypofunction may cause the symptoms of schizophrenia.

The second unresolved question concerns the temporal dynamics of psychotic experiences, which often unfold as short-lived events spanning from seconds to minutes, especially at early stages of schizophrenia. The transient nature of psychotic experiences challenges models that assume a constant disruption of perceptual inference.[1] A solution to this problem is suggested by the recent observation that perceptual inference is subject to spontaneous fluctuations over time.[14] Such fluctuations have been related to two opposing modes of inference, or shifts in attentional sets, during which perception is driven predominantly either by external inputs (external mode) or by internal predictions that stem from recent experiences[15] (internal mode; Fig. 1A). Although preliminary evidence indicates a tendency towards the external mode in people with schizophrenia,[16] the neural mechanisms of mode fluctuations and their potential implications for computational models of schizophrenia have remained elusive.

The objective of the current study was therefore 2-fold: (i) to test whether NMDAR hypofunction causes changes in perceptual inference that characterize schizophrenia; and (ii) to explore the effect of NMDAR hypofunction on ongoing fluctuations in perceptual inference that may explain the transient nature of psychotic experiences. We addressed these questions in a double-blind, placebo-controlled, cross-over experiment with S-ketamine in healthy participants and a case-control study that compared patients with paranoid schizophrenia to matched healthy control subjects.[17] Participants engaged in a task designed to test how internal predictions derived from previous experiences modulate the perception of sensory signals that varied in ambiguity. We found that NMDAR antagonism and schizophrenia were associated with a shift of perception towards the external mode, a minute-long state of the brain during which inference dissociates from prior knowledge. Our results suggest that NMDAR hypofunction shifts the balance between external and internal modes and may thus contribute to the symptoms of schizophrenia by causing transient and recurring failures of perceptual inference.

Materials and methods

For details on the experimental paradigm, participant recruitment and consent, inclusion/exclusion criteria, randomization and blinding, drug administration protocols, safety monitoring, data analysis, and computational modelling, please refer to the online Supplementary material, 'Methods' section.

Results

To investigate whether NMDAR hypofunction influences perceptual inference, and how NMDAR hypofunction contributes to the transient nature of psychotic experiences, we conducted a double-blind placebo-controlled cross-over experiment in 28 healthy human participants. The participants attended two experimental sessions during which they received a continuous intravenous infusion of either the NMDAR antagonist S-ketamine at a dose of 0.1 mg/kg/h or a saline placebo. In each session, the participants viewed ten 120-s blocks of an ambiguous structure-from-motion (SFM) stimulus that induced the experience of a sphere rotating around a vertical axis, and reported changes in the perceived direction of rotation (leftward versus rightward movement of the front surface) as well as their confidence in the choice (Fig. 1B and Supplementary Video 1).

The ambiguity of the display induced the phenomenon of bistable perception: even though the stimulus was physically ambiguous at each frame of the presentation, spontaneous changes in the perceived direction of rotation occurred in average intervals of 13.75 ± 3.09 s. In line with previous results,[17,18] these changes in perception occurred with a probability of $0.11 \pm 8.67 \times 10^{-3}$ at brief depth-symmetric configurations of the stimulus (Supplementary Video 1 and Supplementary Fig. 2A). We therefore divided the continuous behavioural reports into a sequence of discrete states, t. Each state was associated with a perceptual experience (y_t), confidence (c_t) and the external input (s_t).

Predictive processing conceptualizes bistable perception as an inferential process about the cause of s_t. The core idea is that previous experiences (y_{t-1}) generate internal predictions that bias the

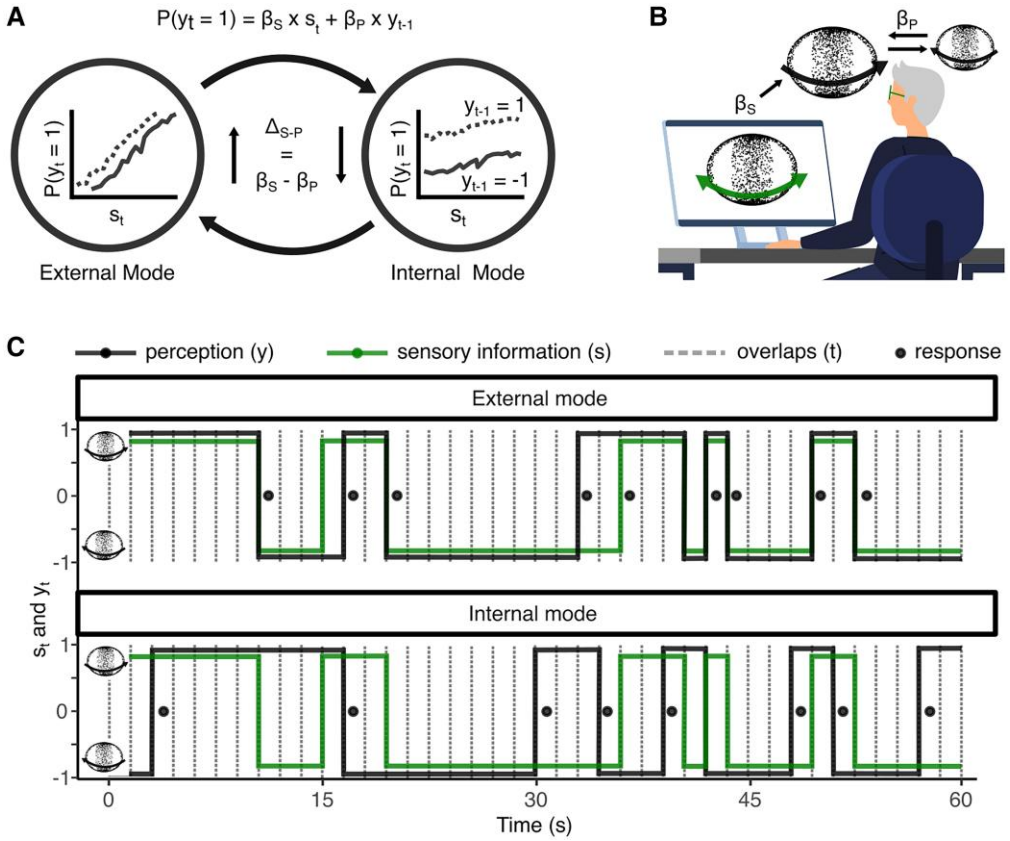

Figure 1 Paradigm. (**A**) Perception integrates ambiguous sensory signals s_t with internal predictions that reflect prior knowledge about the world. One source of prior knowledge is the temporal autocorrelation of natural environments, where the recent past often predicts the near future. The integration of external inputs and internal predictions depends on the weights assigned to incoming sensory data ($\beta_S \times s_t$) and to internal predictions derived from previous experiences ($\beta_P \times y_{t-1}$, dotted versus solid lines, simulated data), respectively. β_S determines the slope and β_P the shift of the psychometric function that links s_t and y_t. The balance $\Delta_{S-P} = \beta_S - \beta_P$ is known to alternate between two opposing modes: During the external mode (*left*), perception is largely determined by $\beta_S \times s_t$, which is reflected by a steep slope and a small shift of the psychometric curve. Conversely, during the internal mode (*right*), perception is shaped by $\beta_P \times y_{t-1}$, resulting in a shallow slope and a large shift of the psychometric curve. (**B**) We conducted a double-blind placebo-controlled experiment in 28 healthy human participants, who received a continuous infusion with either the NMDAR antagonist S-ketamine or saline. During the infusion, the participants viewed structure-from-motion (SFM) stimuli at varying levels of signal-to-ambiguity (SAR). The stimuli were compatible with two mutually exclusive subjective experiences (left versus rightward rotation of the front surface). Fully ambiguous stimuli (SAR = 0) induce the phenomenon of bistable perception, where participants perceive spontaneous changes between the two possible interpretations of the stimulus at a rate that is governed by β_P, the degree to which perception is shaped by internal predictions derived from previous experiences. For partially ambiguous stimuli (SAR > 0), perception reflects the weighted integration of internal predictions with external sensory data, which is governed by the balance $\Delta_{S-P} = \beta_S - \beta_P$. (**C**) Changes in the perceived direction of rotation of the SFM stimulus occur at brief depth-symmetric configurations of the stimulus (overlaps; Supplementary Video 1). We transformed the behavioural responses into a sequence of states t (1.5 s intervals, corresponding to the intervals between consecutive overlaps), each associated with a combination of the SAR-weighted input s_t and the perceived direction of rotation y_t. Participants reported whenever they experienced a change in conscious experience. The response time r_t was defined as the lag between the response and the last preceding overlap. We used a general linear model-hidden Markov model to quantify the weights β_S, β_P and β_B, which reflect how the reported percepts y_t were determined by the external inputs $\beta_S \times s_t$, the internal predictions $\beta_P \times y_{t-1}$ and the constant bias $\beta_B \times 1$ separately for the external mode (*top*; 60 s of example data) and the internal mode (*bottom*; 60 s of example data with identical s_t for visualization). In the external mode, perception follows the external stimulus closely (high $\Delta_{S-P} = \beta_S - \beta_P$). In the internal mode, perception is shaped more strongly by internal predictions derived from previous experiences (low $\Delta_{S-P} = \beta_S - \beta_P$).

interpretation y_t of the ambiguous stimulus[18] (Fig. 1C). In this view, inferences during bistability mirror the temporal autocorrelation of natural environments, where the recent past typically predicts the near future, much like frames captured by a video camera allow for the prediction of future frames.[19] The adaptive benefit of this predictive strategy is a stabilization of perception that prevents erratic experiences in natural environments, which are highly autocorrelated and accessible to the brain only via inherently ambiguous sensory signals.[2]

Predictive processing models of bistable perception assume that transitions between the alternative interpretations of (partially) ambiguous stimuli are driven by conflicts between the external input and stabilizing internal predictions.[17,18] To test how NMDAR antagonism alters the balance between external inputs and internal predictions, we attached a 3D signal to a fraction of the stimulus dots. The signal-to-ambiguity ratio (SAR) ranged from complete ambiguity to full disambiguation across five levels and remained constant in each block of the experiment. By changing the direction of rotation enforced by the 3D signal at random in average intervals of 10 s, we created dynamic conflicts between the SAR-weighted input s_t and the stabilizing internal prediction y_{t-1}. Due to the random changes in s_t, a shift of inference away from internal predictions and towards external sensory data, which has repeatedly been associated with NMDAR hypofunction[1]

and may be maladaptive in autocorrelated natural environments,[15] should manifest as an increase in perceptual accuracy in our experiment.

NMADR hypofunction shifts perception towards the external input and away from internal predictions

As expected, we found that y_t was driven by both s_t ($\beta = 3.01 \pm 0.06$, $z = 50.39$, $P < 0.001$) and y_{t-1} ($\beta = 2.06 \pm 0.03$, $z = 80.58$, $P < 0.001$). Importantly, S-ketamine caused perception to shift towards s_t ($\beta = 0.45 \pm 0.08$, $z = 5.6$, $P < 0.001$) (Fig. 2A and Supplementary Fig. 3), indicating a stronger weighting of external inputs over internal predictions during pharmacologically induced NMDAR hypofunction. Under the predictive processing formulation of perceptual inference, one can read the estimates for s_t and y_{t-1} as sensory and prior precision, respectively. This suggests that S-ketamine augments sensory precision by altering the interactions between pyramidal cells and fast-spiking inhibitory interneurons thought to underwrite cortical gain control or excitation-inhibition balance.[10]

Next, we performed the same analysis on data from a previous case-control study using an analogous task in patients with schizophrenia.[17] In patients with schizophrenia and controls, y_t was influenced by the SAR-weighted input s_t ($\beta = 2.77 \pm 0.11$, $z = -24.85$, $P < 0.001$) and the stabilizing prediction y_{t-1} ($\beta = 1.5 \pm 0.03$, $z = -58.2$, $P < 0.001$). Similar to S-ketamine, s_t had a larger impact on perception in patients with schizophrenia than controls ($\beta = 0.75 \pm 0.15$, $z = 4.96$, $P < 0.001$) (Fig. 2E and Supplementary Fig. 4).

Together, these results align with the canonical predictive processing theory of schizophrenia[1]: pharmacologically-induced NMDAR hypofunction and schizophrenia are associated with a shift of perceptual inference towards external inputs, and away from stabilizing internal predictions. This increase in sensory precision (relative to prior precision) is often framed as a failure of sensory attenuation, i.e. the inability to attenuate sensory precision or, psychologically, ignore unclear or irrelevant sensations.[20] In the artificial setting of our experiment, where stimuli are random, weak internal predictions under S-ketamine and in schizophrenia lead to increased perceptual accuracy. In autocorrelated natural environments, however, NMDAR hypofunction may trigger psychotic experiences by causing erratic inferences about ambiguous sensory information.

NMDARs modulate the balance between external and internal modes of perception

As a mechanism for symptoms that are transient and recurring, NMDAR-dependent changes in perceptual inference should not be constant, but fluctuate dynamically at a timescale that is compatible with the duration of individual psychotic experiences. We tested this prediction in hidden Markov models (HMM) that inferred transitions between two latent states, each linked to an independent general linear model (GLM) that predicted y_t from s_t and y_{t-1}. The β weights quantified the sensitivity to sensory information ($\beta_S \times s_t$) relative to the stabilizing effect of internal predictions provided by preceding experiences ($\beta_P \times y_{t-1}$) and allowed us to evaluate dynamic changes in the balance $\Delta_{S-P} = \beta_S - \beta_P$ between the two.

Consistent with recent findings in humans and mice,[14,15] Bayesian model comparison indicated a clear superiority of the two-state GLM-HMM over the standard one-state GLM in the S-ketamine experiment ($\delta_{BIC} = -3.65 \times 10^3$). According to the two-state GLM-HMM, perception fluctuated between an internal mode, shaped by the stabilizing internal prediction y_{t-1}, and an external mode, dominated by the SAR-weighted input s_t. External mode increased Δ_{S-P} by 2.8 ± 0.29 [$T(81) = 9.5$, $P < 0.001$] (Fig. 2B and C). Switches between modes occurred in intervals of 179.97 ± 19.39 s.

The presence of slow fluctuations between external and internal modes suggests that, instead of causing a constant increase in the sensitivity to external inputs, NMDAR hypofunction may affect perception by shifting the dynamic balance between the two modes. Indeed, S-ketamine did not alter the weights of the two-state GLM-HMM (Fig. 2C), but increased the probability of external at the expense of internal mode ($\beta = 1.01 \pm 0.03$, $z = 30.7$, $P < 0.001$) (Fig. 2D) via an effect on the stay transitions of the HMM (external-to-external and internal-to-internal; Supplementary Fig. 3D). This effect was stable over time, and present across the full range of SAR (Fig. 2D). Inter-individual differences in the effects of S-ketamine confirmed that NMDAR hypofunction raised the sensitivity to sensory information (Fig. 2A) by modulating the time participants spent in external and internal modes, respectively [$\rho = 0.41$, $T(26) = 2.3$, $P = 0.03$]. Our results therefore suggest that the failure of sensory attenuation observed under S-ketamine corresponds to an inability to disengage the external mode of perception. Through the lens of predictive processing, the external mode reflects a state of perception that is characterized by an increase in sensory precision at the expense of prior precision. Crucially, it is this balance between sensory and prior precision that determines the Kalman gain. In other words, what matters in terms of perceptual inference are the dynamic changes in relative precision over time.

Strikingly, the data from the schizophrenia-control study mirrored the effect of S-ketamine on the balance between external and internal mode: The two-state GLM-HMM outperformed the standard one-state GLM (patients: $\delta_{BIC} = -981.65$; controls: $\delta_{BIC} = -862.91$) and revealed two opposing modes [$\Delta_{S-P} = 1.44 \pm 0.33$, $T(44) = 4.33$, $P < 0.001$] (Fig. 2F) that alternated in intervals of 265.38 ± 57.76 s for patients and 230.99 ± 65.04 s for controls. Patients and controls did not differ with respect to the weights of the two-state GLM-HMM (Fig. 2G). Instead, patients with schizophrenia spent more time in external mode ($\beta = 0.52 \pm 0.03$, $z = 16.88$, $P < 0.001$) (Fig. 2H and Supplementary Fig. 4D).

External and internal modes are perceptual phenomena

Our results suggest that healthy participants under S-ketamine and schizophrenia patients spend more time in the external mode. As a dynamic mechanism for psychotic experiences, alternations between external and internal mode should have an effect at the level of perception. This means that between-mode alternations should modulate a perceptual decision variable that determines not only what is consciously experienced, but also how the contents of perception are evaluated by downstream cognition. The hypothesis that external and internal modes are perceptual phenomena needs to be contrasted against alternative scenarios in which external and internal modes are driven primarily by fluctuations in arousal, high-level cognition, or executive function. This is particularly important as behavioural reports served as the sole indicators of perceptual states in our paradigm.

To address these alternative accounts, we first performed additional tests to support our claim that external and internal mode operate at the level of perception. External and internal modes are states of a GLM-HMM that integrates the external stimulus s_t with the previous experience y_{t-1} into a perceptual decision variable

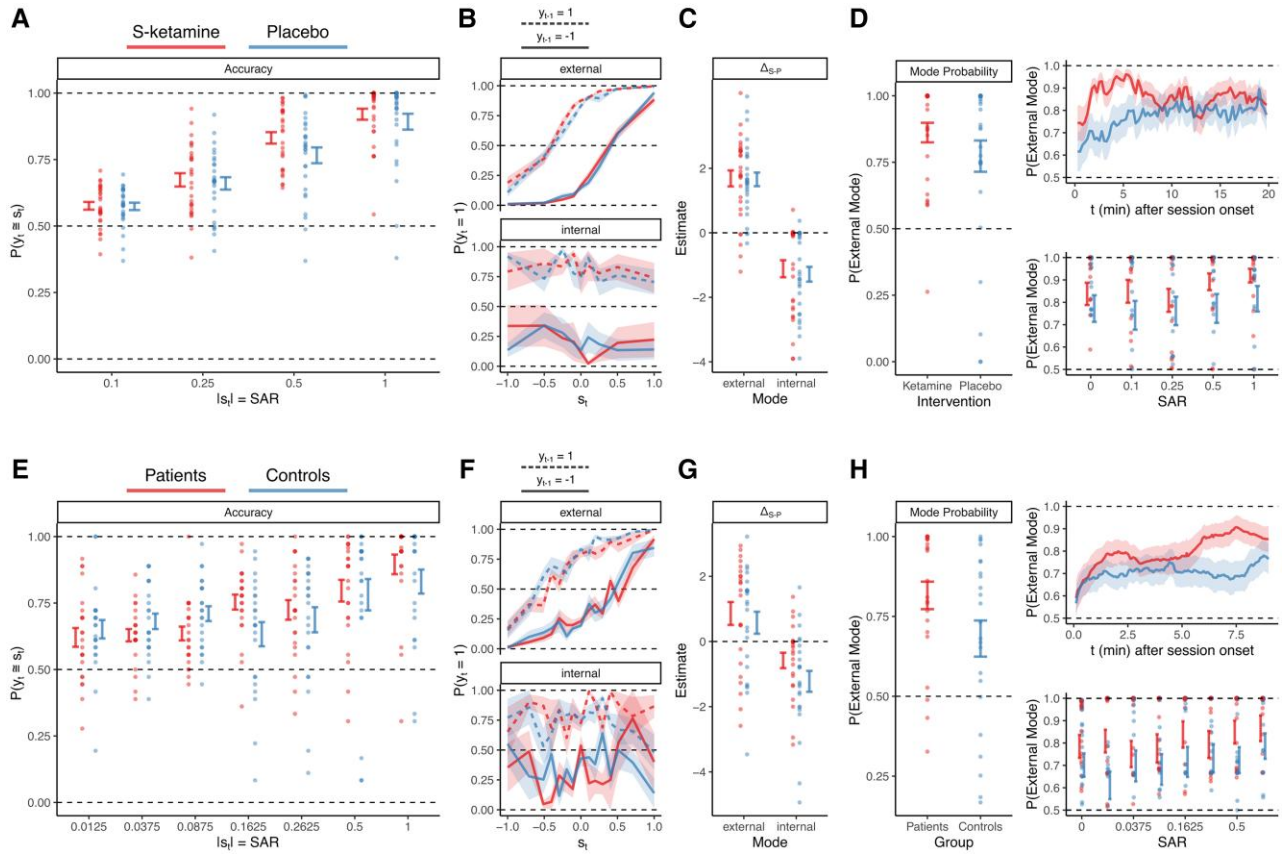

Figure 2 The balance between external and internal modes under S-ketamine and in schizophrenia. (**A**) The percepts y_t were more likely to match the stimuli s_t at higher levels of signal-to-ambiguity (SAR) ($\beta = 3.01 \pm 0.06$, $z = 50.39$, $P < 0.001$). The positive effect of SAR on $P(y_t \cong s_t)$ was more pronounced under S-ketamine relative to placebo ($\beta = 0.45 \pm 0.08$, $z = 5.6$, $P < 0.001$). (**B**) In the S-ketamine experiment, the hidden Markov models (HMM) identified two modes that differed with respect to the relative weighting of external sensory data and internal predictions: perception fluctuated between an external mode, determined by the input s_t (*top*; steep slope and small shift of the psychometric curve), and an internal mode, dominated by a stabilizing prediction that biased perception towards previous experiences y_{t-1} (*bottom*; shallow slope and large shift of the psychometric curve). Within modes, there was no significant effect of S-ketamine versus placebo on the relation of y_t with s_t and y_{t-1}. (**C**) Δ_{S-P}, the balance between the external input and the stabilizing internal predictions, was larger during external than during internal mode [$\beta = 2.8 \pm 0.29$, $T(-81) = -9.5$, $P < 0.001$]. Importantly, we found no significant effect of S-ketamine versus placebo on Δ_{S-P} within modes [$\beta = -0.03 \pm 0.29$, $T(81) = -0.1$, $P = 1$]. (**D**) S-ketamine increased the probability of external mode ($\beta = 1.01 \pm 0.03$, $z = 30.7$, $P < 0.001$) relative to placebo. The effect of S-ketamine on mode was present from the start of the session ($\beta = 1.77 \pm 0.07$, $z = 26.9$, $P < 0.001$; *top right*), with no significant effect of time ($\beta = -0.18 \pm 0.08$, $z = -2.17$, $P = 0.48$). Relative to placebo, S-ketamine increased the probability of external mode across all SARs ($\beta = 0.85 \pm 0.06$, $z = 14.14$, $P < 0.001$; *bottom right*). Higher SARs were associated with an increased probability of external mode ($\beta = 1.34 \pm 0.09$, $z = 15.01$, $P < 0.001$), in particular under S-ketamine ($\beta = 0.62 \pm 0.11$, $z = 5.52$, $P < 0.001$). Alternations between external and internal modes were found at all SARs: from full ambiguity to complete disambiguation, the probability of external mode increased by only 0.11 under S-ketamine and 0.07 under placebo. (**E**) In patients and controls, percepts y_t were more likely to match the stimuli s_t at higher levels of SAR ($\beta = 2.77 \pm 0.11$, $z = 4.96$, $P < 0.001$). Patients followed the external inputs more closely than controls ($\beta = 0.75 \pm 0.15$, $z = 4.96$, $P < 0.001$). (**F**) In analogy to the S-ketamine experiment, the HMM identified two opposing modes in schizophrenia patients and controls. The external mode increased the sensitivity towards s_t (slope of the psychometric function) and weakened the effect of the stabilizing internal prediction y_{t-1} (shift between the dotted and solid line) relative to the internal mode. Within modes, there was no effect of group on the relation of y_t with s_t and y_{t-1}. (**G**) The external mode increased Δ_{S-P}, the balance between external inputs and internal predictions, in patients and controls [$\beta = 1.44 \pm 0.33$, $T(44) = 4.33$, $P < 0.001$], with no significant effect of group [$\beta = -0.28 \pm 0.54$, $T(87.97) = -0.52$, $P = 1$]. (**H**) Relative to controls, patients spent more time in external mode ($\beta = 0.52 \pm 0.03$, $z = 16.88$, $P < 0.001$). In both group, biases towards external mode increased over time after session onset ($\beta = 2.41 \pm 0.11$, $z = 21.37$, $P < 0.001$; *top right*), with a stronger effect in patients ($\beta = -1.84 \pm 0.14$, $z = -12.97$, $P < 0.001$). Patients were more likely than controls to be in external mode across all levels of SAR ($\beta = 0.51 \pm 0.03$, $z = 14.56$, $P < 0.001$; *bottom right*). External mode increased with SAR ($\beta = 0.63 \pm 0.1$, $z = 6.47$, $P < 0.001$), with no significant difference between groups ($\beta = 0.15 \pm 0.13$, $z = 1.16$, $P = 1$). As in the S-ketamine experiment, alternations between external and internal mode were found at all SARs: from full ambiguity to complete disambiguation, the probability of external mode increased by only 0.12 in patients and 0.18 in controls.

$P(y_t = 1)$. The parameters of the GLM-HMM are optimized to predict the sequence of perceptual experiences y_t from $P(y_t = 1)$. If external and internal modes are perceptual phenomena, then the stabilization of perception should be driven by the sequence of experiences y_t, as opposed to the sequence of stimuli s_t. To test this hypothesis, we compared our experienced-based GLM-HMM, in which the stabilizing internal predictions are driven by the participants' perceptual experience at the preceding overlap, with an alternative stimulus-based GLM, in which the stabilizing internal predictions are driven by the stimulus presented at the preceding overlap. Bayesian model comparison indicated that the experienced-based GLM-HMM was better at explaining our data than a stimulus-based GLM-HMM in the S-ketamine experiment ($\delta_{BIC} = -7.4 \times 10^3$) and the case-control study (patients: $\delta_{BIC} = -981.65$; controls: $\delta_{BIC} = -862.91$).

Moreover, if external and internal modes are perceptual phenomena, then the decision variable $P(y_t = 1)$ should not only determine the contents of perception, but also metacognitive processes that depend on them. To assess this prediction, we tested whether

the posterior certainty (C_t) at which the GLM-HMM predicted the content of perception, i.e. the log probability of the experience y_t given the decision variable $P(y_t = 1)$ $\{C_t = y_t \cdot \log[P(y_t = 1)] + (1 - y_t) \cdot \log[1 - P(y_t = 1)]\}$ would correlate with the confidence reports in the S-ketamine experiment. This test is a powerful validation of our approach, as the sGLM-HMM was only fitted to binary perceptual states y_t, and not to the confidence at which they were reported. Indeed, C_t predicted the confidence reports ($\beta = 0.29 \pm 0.02$, $z = 15.4$, $P < 0.001$) without an interaction with mode ($\beta = -0.07 \pm 0.07$, $z = -1.03$, $P = 0.30$), confirming that the positive correlation between posterior certainty and confidence was present in both external and internal modes. C_t extracted from the two-state GLM-HMM was better at explaining confidence than the one-state control GLM ($\delta_{BIC} = -280.69$) and the one-state stimulus GLM ($\delta_{BIC} = -445.13$).

As a consequence, internal mode should be associated with lower metacognitive performance (i.e. the degree to which confidence correlates with accuracy), as stabilizing internal predictions have a larger effect on perception in the internal mode, and cause experiences y_t to be less constrained by the external input s_t. Indeed, accuracy was predictive of high confidence ($\beta = -1.01 \pm 0.05$, $z = 18.7$, $P < 0.001$), but to a lesser degree during the internal mode ($\beta = -0.61 \pm 0.09$, $z = -6.61$, $P < 0.001$). In line with this, metacognitive sensitivity, as measured by meta-d', was significantly lower in the internal mode [$\beta = -1.6 \pm 0.45$, $T(50) = -3.55$, $P = 0.03$]. Together, these findings support the hypothesis that external and internal modes modulate a low-level decision variable $P(y_t = 1)$ that determines the content of perception and their metacognitive evaluation.

Second, we asked whether fluctuations in global brain states can provide an alternative explanation for external and internal modes. One could assume that mode alternations could in fact reflect dynamic states of arousal, with high arousal and engaged behaviour corresponding to the external mode, and low arousal and disengaged behaviour corresponding to the internal mode. Our time-resolved assessment of internal states revealed reduced wakefulness (Q1) under S-ketamine (Supplementary Fig. 6). This observation is clearly incompatible with the hypothesis that changes in the dynamics of mode are driven by low arousal under S-ketamine, as NMDAR antagonism increased the prevalence of the external mode, improving behavioural performance in the artificial setting of our experiment. When controlling for dynamic changes in wakefulness (Q1), subjective intoxication (Q2) and nervousness (Q3), the effect of S-ketamine on mode ($P < 0.001$) and the effect of mode on Δ_{S-P} remained significant ($P < 0.001$). We observed no additional effects of or interactions with Q1–3 that could explain the observed relations between S-ketamine, mode and Δ_{S-P}. Despite its positive effect on perceptual accuracy, external mode was associated with higher levels of dissociation in the S-ketamine experiment as measured by the Clinician Administered Dissociative States Scale[21] [CADSS; $\beta = -1.05 \pm 0.54$, $T(208.05) = 1.95$, $P = 0.05$] (Supplementary Fig. 6B).

In addition to the time-resolved subjective reports on wakefulness obtained under S-ketamine and placebo (Supplementary Fig. 6), response times (r_t) can provide an indirect measure of task engagement, with longer r_t and higher r_t variability as indicators of fatigue or disengagement.[22,23] We found no significant effect of mode on r_t in either the S-ketamine experiment ($\beta = 0.02 \pm 0.03$, $z = 5.96 \times 10^3$, $P = 0.78$) or in the case-control study ($\beta = 0.03 \pm 0.04$, $z = 4.89 \times 10^3$, $P = 0.76$). r_t variability did not differ significantly between modes in the S-ketamine intervention ($V = 85$, $P = 0.47$) or in the case-control study ($W = 945$, $P = 0.59$). In both experiments, there was no main effect of time on r_t [S-ketamine intervention: $\beta = 6.11 \times 10^{-3}$, $T(6.22 \times 10^3) = 0.11$, $P = 1$; case-control study: $\beta = -0.04 \pm 0.05$, $T(5.34 \times 10^3) = -0.71$, $P = 1$]. We observed no Time × Intervention interaction [$\beta = 0.04 \pm 0.08$, $T(6.22 \times 10^3) = 0.47$, $P = 1$] nor a Time × Group interaction [$\beta = 0.06 \pm 0.07$, $T(5.35 \times 10^3) = 0.86$, $P = 1$], suggesting that interventions and groups did not differ with respect to fatigue.

Contrary to the natural dynamic of fatigue in psychophysical experiments, which increases over time, we observed no effect of time on the balance between modes in the S-ketamine experiment ($\beta = -0.18 \pm 0.08$, $z = -2.17$, $P = 0.48$; Fig. 2D). In the case-control study, external mode even became more prevalent over time ($\beta = 2.41 \pm 0.11$, $z = 21.37$, $P < 0.001$), with a stronger effect in patients ($\beta = 1.84 \pm 0.14$, $z = 12.97$, $P < 0.001$; Fig. 2H).

Furthermore, we found no evidence that external and internal modes reflect behavioural strategies that depend on task difficulty, such as using internal predictions only when the sensory information is unreliable. Individual stereodisparity thresholds were not correlated with inter-individual differences in mode (Supplementary Fig. 6). Within participants, the balance between external and internal mode was only marginally modulated by the SAR of the stimulus (Fig. 2D and H).

In sum, these findings suggest that the effect of S-ketamine on mode, and the effects of mode on the integration of external inputs with internal predictions (Δ_{S-P}), are unlikely to be mediated by dynamic changes in arousal, fatigue, task engagement or task difficulty. Rather, they indicate the NMDAR hypofunction under S-ketamine and in schizophrenia has a direct impact on perceptual processing via its effect on mode.

Discussion

Perception integrates incoming signals with internal predictions that reflect prior knowledge about the world.[2] Our results indicate that this integration is subject to dynamic changes over time, alternating between an external mode, where perception closely follows the external input, and an internal mode, where perception is shaped by internal predictions.[15,24] The internal mode enables the brain to use prior knowledge about the statistics of natural environments, such as their temporal autocorrelation, for efficient perception.[15] Intermittent episodes of external mode processing decouple perception from prior knowledge. The balance between external and internal mode may prevent circular inferences within recurrent neural networks, where predictive feedback influences early stages of sensory processing.[25] We found that healthy individuals receiving the NMDAR antagonist S-ketamine, as well as patients diagnosed with schizophrenia, are more prone to an external mode of perception. This NMDAR-dependent change in the balance between modes may expose perception to the destabilizing effects of sensory ambiguity, causing afflicted individuals to be deluded by spurious connections between unrelated events, to attribute the sensory consequences of their actions to an outside force, and to hallucinate signals in noise.[1]

External and internal mode explain dynamic failures of perceptual inference in schizophrenia

During bistable perception, previous experiences provide the predictive context in which incoming sensory data are interpreted, and lead to prolonged periods of perceptual stability despite the ambiguity of the external input.[18] Our results suggest that NMDAR hypofunction, whether due to pharmacological antagonism or as a potential endophenotype of schizophrenia, causes a

shift of bistable perception towards the external input, and away from stabilizing internal predictions that stem from previous experiences. These findings bear similarity with prior work on perceptual illusions, where prior knowledge biases perception in ways that may be adaptive in natural environments but reduce perceptual accuracy in experimental settings.[26] Weak predictions may explain why people with schizophrenia are, for example, less susceptible to the hollow-mask illusion, where knowledge about faces is thought to induce the experience of a convex face on the concave surface of a human mask; the Ebbinghaus illusion, where larger circles make a smaller central circle appear bigger; or the force-matching illusion, where humans apply less force when matching an externally applied force with their own.[26]

Our findings therefore align with the canonical predictive processing account of psychosis.[1] According to this model, NMDAR hypofunction[7] and schizophrenia[17] are associated with weak priors that cause erratic inferences in perception and cognition, ultimately leading to psychotic symptoms, such as delusions and hallucinations. At the same time, they seem at odds with the observation that psychotic experiences, and in particular false alarms that serve as an experimental proxy for hallucinations, correlate with strong priors.[27] So far, attempts to reconcile these disparate sets of findings suggest that priors may vary in strength depending on the phase of psychotic illness, with weak priors in early stages and strong priors in later stages, or depending on their position within the cognitive hierarchy, with weak priors at the perceptual level and strong priors at the cognitive level.[1] As an alternative to predictive processing, circular inference accounts of schizophrenia posit that psychotic symptoms depend on an over-counting of sensory data that are reverberated multiple times due to an imbalance of excitation and inhibition in feedforward-feedback loops of the cortical hierarchy.[28]

In line with the general principles of predictive processing, the GLM-HMM proposed here predicts the experiences y_t in a weighted integration of the external input $\beta_S \times s_t$ with internal predictions that embody the temporal autocorrelation of natural environments and are defined by the preceding experiences $\beta_P \times y_{t-1}$. The critical advance provided by the GLM-HMM is that the model allows for dynamic changes in the balance between external and internal sources of information ($\Delta_{S-P} = \beta_S - \beta_P$). In the data presented here, the GLM-HMM revealed that the general shift of perception towards the external input and away from internal predictions observed under S-ketamine and in schizophrenia is in fact driven by changes in the balance between two opposing modes of inference: an external mode, during which priors are weak, and an internal mode, during which priors are strong. The failures of perceptual inference, which are hypothesized to characterize schizophrenia,[1] may thus be transient and recurring.

To our knowledge, our results are the first to uncover a neural mechanism underlying the slow, task-related fluctuations in perceptual inference observed in both humans and mice.[14,15] In the context of schizophrenia, this extends previous predictive processing accounts by suggesting an alternative explanation for the apparent discrepancy between strong and weak priors: an imbalance between the modes may cause the brain to make erratic inferences during the external mode, when the influence of previously learned priors is weak, generating a distorted or inaccurate model of the world, which is then used maladaptively during the internal mode, when priors are strong.[24] Furthermore, the dynamic nature of between-mode transitions illustrates how constant and potentially heritable dysfunctions of the NMDAR may produce symptoms of psychosis that are recurrent and transient in nature.

Is the balance between modes a trait or a psychosis-related state of perceptual inference?

In the present data, we did not find a correlation of the balance between external and internal mode with either global psychosis proneness or the clinical severity of schizophrenia (Supplementary Fig. 6). Our study was optimized for within-participant power and not designed to detect correlations between inter-individual differences in schizophrenia-related traits and the balance between external and internal modes. One key question moving forward is whether the shift toward external mode represents a general trait-like phenomenon in schizophrenia, potentially linked to cognitive alterations that are also present to some degree under ketamine,[29] or whether external and internal modes are associated with psychosis-related, state-dependent changes in inference.

Future research could address these questions by correlating the balance between modes with both positive and negative symptoms, as well as with measures of cognitive performance, such as IQ, in larger samples. Another promising approach to distinguish between trait and state effects, which can manifest differently or even with opposite phenotypes,[30] could involve real-time symptom tracking combined with functional imaging. Such analyses could help to examine whether shifts between external and internal modes align with the on- and offset of individual psychotic experiences,[24] both at the behavioural level and in terms of their neural correlates.

Are external and internal mode perceptual or behavioural phenomena?

Previous studies have used GLM-HMMs to identify engaged and disengaged behaviour in mice tasked with discriminating the location of a visual stimulus.[14] While this terminology may suggest that GLM-HMM states reflect dynamic changes in rodent behaviour, evidence from human psychophysics indicates that external and internal modes may in fact reflect perceptual (as opposed to behavioural) states.[15,24] Specifically, when humans detect gratings in white noise, false alarms are more likely when the noise contains more power at the orientation and spatial frequency of the preceding grating, suggesting that detection relies on a predictive perceptual template.[19,24] If these detection events were purely behavioural, no correlation between false alarms and the noise power spectrum would be expected.[24] Critically, recent work demonstrates that these predictive perceptual templates are confined to the internal mode, supporting the hypothesis that the internal mode is indeed predictive and perceptual.[24] Moreover, an analysis of 66 experiments on human two-alternative forced-choice decision-making revealed a quadratic relationship of confidence with mode.[15] The observation that confidence remains high for strong biases towards both external and internal modes[15] argues against the interpretation of internal mode as disengaged behaviour.

These observations do not, however, rule out the possibility that external and internal modes have multiple and potentially independent effects on the brain, including influences on high-level cognition and response behaviour, or that they are, to some degree, dependent on global brain states. As our analyses rely on behavioural reports about changes in the content of perception, dynamic changes in response behaviour represent an additional potential confound in the identification of external and internal modes.

Future work should use trial-wise reports of perception and confidence with randomized response mappings to enable GLMs that can disentangle perception and response behaviour. No-report

functional imaging experiments, where the content of experience is decoded without overt behavioural signals, alongside pupillometry, manipulations of neuromodulators that regulate global brain states, or non-invasive brain stimulation, could help illuminate the causes and consequences of these modes across the cortical hierarchy. Mapping the neurocomputational dynamics of mode alternations will be crucial to testing whether adjusting the balance between modes can mitigate psychotic experiences and ultimately improve the lives of people living with schizophrenia.

Data availability

Further information and requests for resources should be directed to and will be fulfilled by the lead corresponding author. This study did not generate new unique reagents. All data and code associated with this study will be made available on the associated Github repository https://github.com/veithweilnhammer/modes_ketamine_scz upon publication. Key resources are listed in Supplementary Table 1.

Funding

This work was funded by the Deutsche Akademie der Naturforscher Leopoldina - Nationale Akademie der Wissenschaften (grant number: LDPS2022-16, https://www.leopoldina.org/en/leopoldina-home/), the German Research Foundation DFG (grant number: STE 1430/9-1, https://www.dfg.de), the Berlin Institute of Health Clinician Scientist Program (https://www.bihealth.org/en/translation/innovation-enabler/academy/bih-charite-clinician-scientist-program) and the German Ministry for Research and Education (ERA-NET NEURON program, grant number: 01EW2007A, https://www.neuron-eranet.eu/). The funders had no role in study design, data collection, data analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Sterzer P, Adams RA, Fletcher P, et al. The predictive coding account of psychosis. *Biol Psychiatry*. 2018;84:634-643.
2. Friston K. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 2005;360:815-836.
3. Rao RP, Ballard DH. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*. 1999;2:79-87.
4. Feldman H, Friston KJ. Attention, uncertainty, and free-energy. *Front Hum Neurosci*. 2010;4:7028.
5. Muthukumaraswamy SD, Shaw AD, Jackson LE, Hall J, Moran R, Saxena N. Evidence that subanesthetic doses of ketamine cause sustained disruptions of NMDA and AMPA-mediated frontoparietal connectivity in humans. *J Neurosci*. 2015;35:11694-11706.
6. Corlett PR, Honey GD, Krystal JH, Fletcher PC. Glutamatergic model psychoses: Prediction error, learning, and inference. *Neuropsychopharmacology*. 2011;36:294-315.
7. Stein H, Barbosa J, Rosa-Justicia M, et al. Reduced serial dependence suggests deficits in synaptic potentiation in anti-NMDAR encephalitis and schizophrenia. *Nat Commun*. 2020;11:4250.
8. Murray JD, Anticevic A, Gancsos M, et al. Linking microcircuit dysfunction to cognitive impairment: Effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cereb Cortex*. 2014;24:859-872.
9. Catts VS, Lai YL, Weickert CS, Weickert TW, Catts SV. A quantitative review of the postmortem evidence for decreased cortical N-methyl-D-aspartate receptor expression levels in schizophrenia: How can we link molecular abnormalities to mismatch negativity deficits? *Biol Psychol*. 2016;116:57-67.
10. Adams RA, Pinotsis D, Tsirlis K, et al. Computational modeling of electroencephalography and functional magnetic resonance imaging paradigms indicates a consistent loss of pyramidal cell synaptic gain in schizophrenia. *Biol Psychiatry*. 2022;91:202-215.
11. Self MW, Kooijmans RN, Supèr H, Lamme VA, Roelfsema PR. Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proc Natl Acad Sci U S A*. 2012;109:11031-11036.
12. Castro-Alamancos MA, Connors BW. Short-term synaptic enhancement and long-term potentiation in neocortex. *Proc Natl Acad Sci U S A*. 1996;93:1335-1339.
13. Nakazawa K, Jeevakumar V, Nakao K. Spatial and temporal boundaries of NMDA receptor hypofunction leading to schizophrenia. *NPJ Schizophr*. 2017;3:7.
14. Ashwood ZC, Roy NA, Stone IR, et al. Mice alternate between discrete strategies during perceptual decision-making. *Nat Neurosci*. 2022;25:201-212.
15. Weilnhammer V, Stuke H, Standvoss K, Sterzer P. Sensory processing in humans and mice fluctuates between external and internal modes. *PLOS Biol*. 2023;21:e3002410.
16. Albert S, Schmack K, Sterzer P, Schneider G. A hierarchical stochastic model for bistable perception. *PLOS Comput Biol*. 2017;13:e1005856.
17. Weilnhammer V, Röd L, Eckert A-L, Stuke H, Heinz A, Sterzer P. Psychotic experiences in schizophrenia and sensitivity to sensory evidence. *Schizophr Bull*. 2020;46:927-936.
18. Weilnhammer V, Fritsch M, Chikermane M, et al. An active role of inferior frontal cortex in conscious experience. *Curr Biol*. 2021;31:2868-2880.e8.
19. Manassi M, Whitney D. Continuity fields enhance visual perception through positive serial dependence. *Nat Rev Psychol*. 2024;3:352-366.
20. Shergill SS, Samson G, Bays PM, Frith CD, Wolpert DM. Evidence for sensory prediction deficits in schizophrenia. *Am J Psychiatry*. 2005;162:2384-2386.
21. Mertens YL, Daniels JK. The clinician-administered dissociative states scale (CADSS): Validation of the German version. *J Trauma Dissociation*. 2022;23:366-384.
22. Kucyi A, Esterman M, Riley CS, Valera EM. Spontaneous default network activity reflects behavioral variability independent of mind-wandering. *Proc Natl Acad Sci U S A*. 2016;113:13899-13904.
23. Yamashita A, Rothlein D, Kucyi A, et al. Variable rather than extreme slow reaction times distinguish brain states during sustained attention. *Sci Rep*. 2021;11:14883.
24. Weilnhammer V, Murai Y, Whitney D. Dynamic predictive templates in perception. *Curr Biol*. 2024;34:4301-4306.e2.
25. Muckli L, De Martino F, Vizioli L, et al. Contextual feedback to superficial layers of V1. *Cur Biol*. 2015;25:2690-2695.
26. Notredame C-E, Pins D, Deneve S, Jardri R. What visual illusions teach us about schizophrenia. *Front Integr Neurosci*. 2014;8:63.

27. Schmack K, Bosc M, Ott T, Sturgill JF, Kepecs A. Striatal dopamine mediates hallucination-like perception in mice. *Science*. 2021;372:eabf4740.
28. Jardri R, Duverne S, Litvinova AS, Denève S. Experimental evidence for circular inference in schizophrenia. *Nat Commun*. 2017;8:14218.
29. Morgan CJA, Curran HV; Independent Scientific Committee on Drugs. Ketamine use: A review. *Addiction* 2012;107: 27-38.
30. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. *Front Psychiatry*. 2013;4:47.