Orthogonal Modelling of Positive and Negative Symptoms in Psychosis and Their Differential Treatment Responses to Antipsychotics, Ketamine, and ECT

Authors:

Ngo Cheung, FHKAM(Psychiatry)

Affiliations:

[1] Independent Researcher

Corresponding Author:

Ngo Cheung, MBBS, FHKAM(Psychiatry)

Hong Kong SAR, China

Tel: 98768323

Email: info@cheungngomedical.com

Abstract

Background: Schizophrenia is marked by persistent heterogeneity: antipsychotics reliably control positive symptoms yet leave negative symptoms largely untouched, while emerging glutamatergic and neuromodulatory treatments show inconsistent results across patients. No single computational framework has yet allowed direct, mechanism-agnostic comparison of these interventions within the same hierarchical architecture. We therefore built a fully parameter-based predictive-coding network that orthogonally implements precision dysregulation (positive symptoms) and higher-layer structural pruning, capacity loss, and effort deficits (negative symptoms), then applied antipsychotics, ketamine, and ECT as explicit parameter modifications.

Methods: A three-layer feedforward network ($2 \rightarrow [256, 256, 128] \rightarrow 4$) was trained on synthetic 2D blob classification and a harder cognitive-probe task. Psychosis was induced by shifting gains/noise, injecting aberrant weights, and pruning connections with cascading unit gating and effort reduction. Treatments were implemented solely through parameter changes (precision normalisation for antipsychotics, precision disruption plus synaptogenesis for ketamine, global reset plus cumulative structural repair for ECT) at matched "iso-dose" levels using full and targeted L1 norms. Nine experimental phases examined default-dose effects, zero-consolidation ablation, dose sweeps, chronic maintenance/withdrawal, and ten patient profiles varying in positive/negative severity ratios.

Results: Antipsychotics restored combined accuracy (97.8 %) and stress resilience with minimal structural repair and lowest side effects. ECT achieved the most comprehensive recovery (cognitive probe 79.5 %, capacity 0.912, effort 0.87) and highest durability (0.98 retention). Ketamine produced genuine structural restoration (capacity 0.786, effort 0.74) but was limited by precision disruption; monotherapy consistently underperformed across all subtypes. Iso-dose matching showed direction and targeting of change mattered more than magnitude. Patient variability revealed clear stratification: ECT optimal for positive-dominant cases, antipsychotics efficient for balanced profiles, and ketamine competitive only when combined with precision stabilisation in negative-dominant cases.

Conclusions: This single architecture unifies Bayesian and neuroplastic accounts of psychosis, demonstrates why no treatment is universally superior, and generates testable predictions for subtype-guided and adjunctive strategies. The findings argue for routine positive/negative severity stratification and suggest low-dose ketamine plus a precision-stabilising agent as a mechanistically rational approach for deficit-syndrome patients.

Introduction

Schizophrenia is widely recognised as one of the most disabling and variable psychiatric illnesses. While antipsychotic drugs can markedly reduce hallucinations and delusions, negative features such as diminished motivation, flattened affect, and slowed cognition often linger and determine long-term outcome. Because no single treatment works across the board, clinicians still rely on pragmatic trial-and-error to decide whether a patient will benefit more from dopamine antagonists, glutamatergic agents, or broader neuromodulatory options such as electroconvulsive therapy (ECT). This therapeutic uncertainty has fuelled interest in computational approaches that could replace descriptive diagnosis with mechanism-based, personalised care.

Computational psychiatry seeks to formalise mental operations as explicit algorithms—Bayesian inference, prediction-error signalling, reinforcement learning—and in doing so link brain processes to observable behaviour [1,2]. The roots go back to early connectionist models that tied dopamine dysfunction to context-processing deficits [3] and to simulations of basal-ganglia reward prediction errors [2]. Over the past decade and a half, predictive processing has become a unifying framework, supported by tools that allow entire treatment pipelines to be simulated within a single architecture.

Predictive coding now provides the dominant account of positive symptoms [4]. The brain is viewed as a hierarchical inference engine that minimises prediction errors by updating priors or by acting to confirm them [5]. Psychosis, on this view, reflects a misallocation of precision: sensory evidence is given too much weight—or priors too little—at lower levels, while overly precise higher-level beliefs become rigid and resistant to change [1,4]. NMDA-receptor hypofunction compromises the reliability

of priors, and striatal dopamine excess amplifies the salience of prediction errors, so neutral events acquire inappropriate motivational significance [6,7]. Ketamine challenge studies that mimic early psychosis support the model by reproducing both behavioural and neural signatures [8].

Most predictive-coding work, however, centres on positive symptoms. Negative symptoms are often treated as indirect consequences of the same precision imbalance or dismissed as mere cognitive deficits. Structural findings—synaptic pruning, hypofrontality, reduced dendritic spines—are seldom folded into the same computational scheme. Consequently, we still lack a unified model that can explain, in a common language, why antipsychotics target positive symptoms yet spare negative ones, why ketamine can enhance motivation but also exacerbate psychosis, and why ECT often succeeds where other treatments fail.

The present study addresses this gap. We built a three-layer feed-forward network in which pathology and treatment are expressed by direct, transparent changes to a single parameter set. Positive symptoms arise from reduced bottom-up gain, increased sensory noise, overly precise top-down priors, aberrant plasticity, and global hyper-excitability. Negative symptoms are modelled separately through higher-layer synaptic pruning, reversible and irreversible gating, capacity-dependent excitability scaling, and an effort parameter that dampens higher layers. Three interventions—antipsychotic-induced precision normalisation, ketamine-related precision disruption plus synaptogenesis, and ECT-like global reset with structural repair—are implemented by the same parameters, enabling fair "iso-dose" comparisons.

By keeping everything within one parameter space we can trace, layer by layer, why each treatment succeeds or fails for different symptom profiles. The network itself is deliberately simple, classifying synthetic stimuli and a more demanding cognitive-probe task, but its transparency allows every clinical effect—from preserved performance during psychosis to the biphasic dose–response of ketamine and the durable benefits of ECT—to be read directly from parameter trajectories. Simulations of ten virtual patients, differing only in positive/negative symptom ratios and random seed, reveal subtype-specific response patterns consistent with clinical observation: antipsychotics

excel in positive-dominant cases, ECT shows broad efficacy, and ketamine monotherapy underperforms overall yet offers structural advantages in isolation.

The sections that follow detail the architecture, the induction of pathology, treatment protocols, and the nine experimental phases, before presenting results and discussing implications for personalised, mechanism-guided psychiatry.

Methods

Network Architecture

All simulations used a three-layer feed-forward network coded in PyTorch 2.0. The model's input dimension was two, followed by hidden layers of 256, 256 and 128 units, and a four-unit soft-max output. Each hidden layer performed a linear transform with ReLU activation. Every pathophysiological or treatment manipulation was expressed through explicit nn.Parameter objects: one gain and one noise term per layer, a global excitability factor, a single effort scalar that disproportionately affects deeper layers, binary capacity masks for each hidden unit and an additive side-effect noise term. No other parameters were altered, allowing direct inspection of all changes.

During the forward pass each hidden layer applied its gain, excitability and effort scalars, masked inactive units, rescaled output by residual capacity (never below 0.5), and added Gaussian noise representing intrinsic and external stress. Side-effect noise was injected post-activation. Prediction errors were computed after the pass as the absolute, normalised difference between successive layers, replicating diagnostic error signals described in predictive-coding accounts [1].

Tasks and Datasets

Training and testing (Table 1) relied on synthetic two-dimensional four-class Gaussian "blob" data. The training set contained 10,000 samples with noise $\sigma = 0.7$. Evaluation used three independent sets: a clean set of 2,000 samples ($\sigma = 0$), a standard noisy set of 3,000 samples ($\sigma = 0.7$) and a

cognitive-probe set of 2,000 samples with cluster centres halved in radius and noise σ = 1.2. Class centres were placed at every (±3, ±3) coordinate. Networks were optimised with Adam (learning rate 0.001) and a cross-entropy loss.

Table 1. Evaluation Metrics Definitions

Metric	Description	Computation
Clean Accuracy	Accuracy on noise-free inputs	Classification accuracy on clean test set (σ=0)
Standard Accuracy	Accuracy on standard noisy inputs	Accuracy on test set (σ=0.7)
Combined Accuracy	Accuracy under combined noise + stress	Accuracy on noisy test set with added input noise (σ=1.0) and external stress (0.5)
Cognitive Probe Accuracy	Accuracy on demanding, low-separation task	Accuracy on cognitive probe set (centers scaled 0.5×, σ=1.2), ± additional stress
Stress Battery	Accuracy under graded external stress levels	Five levels (none, mild, moderate, severe, extreme) as additive noise
Hallucination Rate	Rate of confident misclassifications on pure noise	Fraction of Gaussian noise inputs (σ=4.0) exceeding healthy logit-margin threshold (75th percentile)
Flexibility	Sensitivity to input perturbations	Mean KL divergence between softmax outputs on clean vs. perturbed inputs (σ=2.0)
Prediction Error Score	Total normalized prediction errors across layers	Sum of mean absolute normalized residuals between adjacent hidden layers
Negative Score	Simple proxy for negative symptoms	(100 − clean accuracy) / 100, floored at 0
Capacity	Mean per-unit mask value across hidden layers	Average of unit mask parameters
Effort	Global motivational scalar	Direct effort parameter value
Side Effects	Mean absolute per-unit side-effect noise	Average of absolute side-effect parameters
Active Pruning Fraction	Fraction of originally pruned weights remaining suppressed	Ratio of pruned positions still <20% original magnitude
Positive Composite	Summary of precision dysregulation	Normalized average of gain imbalance, average noise, and scaled aberrant fraction
Negative Composite	Summary of structural/motivational deficit	Average of negative score, pruning fraction, capacity deficit, effort deficit

Psychosis Induction

Psychotic pathology was instantiated through parameter shifts that mirror predictive-coding explanations of positive and negative symptoms [4]. Positive-symptom severity increased lower-layer gains to 0.30, 0.60 and 2.50, raised noise to 1.20, 0.60 and 0.05, boosted global excitability to 1.50 and applied a 3 % random weight amplification biased toward deeper layers. Negative-symptom severity removed roughly 25 % of higher-layer weights (10 % set permanently to zero, the remainder reduced to 1 % of their original value), degraded capacity masks and lowered the effort scalar. Severity values scaled linearly from the healthy baseline.

Treatment Protocols

All interventions modified exactly the same parameter set in distinct ways.

Ketamine reduced lower-layer gains, raised noise, boosted excitability, reversed a proportion of pruning and restored masks and effort; side-effect noise represented dissociation. After ketamine administration, five consolidation epochs with Adam at a learning rate of 3×10^{-4} were run.

Antipsychotics gradually returned gains and noise toward healthy levels, dampened aberrant amplifications, introduced mild synaptogenesis and restored effort; lower-layer side-effect noise represented extrapyramidal symptoms, again followed by five consolidation epochs.

ECT applied brief Gaussian perturbations to weights, reset gains and noise incrementally toward baseline, triggered strong synaptogenesis (including partial reversal of irreversible pruning), fully restored capacity and effort, and normalised excitability. Post-ictal side-effect noise was added with exponential decay between sessions, and three consolidation epochs followed each treatment session.

Evaluation Metrics

Performance metrics included classification accuracy under clean, noisy and cognitive-probe conditions; hallucination rate, defined as the proportion of pure-noise inputs exceeding the 75th-percentile healthy logit margin; flexibility, measured as mean Kullback–Leibler divergence when weights were perturbed; summed prediction errors; mean capacity mask value; effort scalar; average side-effect noise; proportion of active pruning; and two composite pathology indices.

Experimental Procedure

Experiments ran in nine ordered phases with fixed random seeds (Table 2). Phase 1 established the healthy reference and hallucination thresholds. Phase 2 applied balanced psychosis at severity 1.0 for both symptom clusters. Phases 3 and 4 delivered the default treatment schedules and recorded parameter trajectories; Phase 3b repeated antipsychotic therapy without consolidation to isolate that factor. Phases 5 and 6 varied treatment intensity and matched therapies by the L1 norm of total parameter change, enabling "iso-dose" comparisons. Phase 7 calculated treatment efficiency as

improvement per unit dose. Phase 8 simulated chronic maintenance, withdrawal and relapse. Phase 9 tested ten synthetic patient profiles varying only in positive/negative ratios and random seeds, mapping symptom-specific responses. All code and configuration files are available from the authors on request.

Table 2. Experimental Phases and Activities

Phase	Name	Key Activities
1	Healthy Baseline	Train network for 25 epochs on standard blob task; calibrate logit-margin hallucination threshold from healthy model on noise inputs (75th percentile for ~25% healthy rate); record full evaluation metrics.
2	Psychosis Induction	Induce balanced psychosis (severity 1.0 positive/negative): shift gains/noise/excitability, add aberrant weights, prune ~25% connections (higher-layer bias, mix reversible/irreversible), cascade unit degradation, reduce effort; record evaluation.
3	Default-Dose Treatments	Apply ketamine (intensity 0.7), antipsychotic (dose 0.7), and ECT (8 sessions) independently to psychotic copies; each via direct parameter changes + consolidation (5 epochs for ketamine/AP, 3 per ECT session); compute L1/targeted L1 doses; evaluate post-treatment and after single relapse step.
3b	Zero-Consolidation Ablation	Repeat Phase 3 treatments with zero consolidation epochs; compare metrics (accuracy, cognitive, hallucination, etc.) to consolidated versions to isolate raw pharmacological effects.
4	Precision and Effort Trajectories	Tabulate per-layer gain/noise values and global effort across states: healthy, psychotic, post-ketamine, post-antipsychotic, post-ECT.
5	Iso-Dose Parameter Sweeps	Sweep treatment intensities independently (ketamine 0.1–0.9, antipsychotic 0.1–1.0, ECT 1–12 sessions); record full/targeted L1 doses, all metrics, and post-relapse accuracy at each point.
6	Iso-Dose Matched Comparisons	Identify overlapping L1 ranges (full and targeted); select matched targets; compare closest sweep points across treatments on key metrics (combined accuracy, cognitive, hallucination, etc.); identify best/worst per match.
7	Treatment Efficiency	For each sweep point, compute efficiency as absolute combined accuracy improvement divided by L1 (full and targeted); tabulate with supporting metrics.
8	Chronic Treatment and Withdrawal	Simulate acute treatment, 12 chronic maintenance steps (40% dose every 3 steps + homeostasis/low-rate training), and 15 withdrawal steps (heterogeneous relapse + homeostasis/side-effect decay); record metrics at baseline, acute, chronic, and withdrawal; compute durability and retained improvement.
9	Individual Variability	Simulate 10 patients with unique seeds and positive/negative severity profiles; induce patient-specific psychosis; apply all three treatments; record improvements and metrics; aggregate means/standard deviations; stratify into positive-dominant, negative-dominant, and balanced groups; identify most/least effective treatment per group.

Results

Healthy baseline and induction of psychosis

After training, the intact network performed at ceiling on most tasks: clean and standard noisy inputs were classified without error (100 % each), combined accuracy under additional noise and stress reached 98.5 %, and the demanding cognitive-probe task was solved at 79.7 %. The hallucination threshold, defined as the 75th-percentile logit margin, settled at 23.130 and yielded a baseline hallucination rate of 0.248. Flexibility—the mean KL divergence after small perturbations—was

1.031, whereas the global prediction-error score was 4.321. Structural measures showed full capacity (1.000), maximal effort (1.000) and an absence of side-effects or pruning.

Introducing balanced psychosis (positive and negative severity = 1.0) produced marked functional deterioration under load while sparing near-perfect performance on noiseless data. Clean accuracy fell only slightly to 99.6 %, but combined accuracy dropped to 82.5 % and extreme-stress accuracy to 49.8 %. Cognitive-probe accuracy declined to 68.5 % (loaded condition 52.3 %). Capacity and effort shrank to 0.678 and 0.650, respectively, and every unit earmarked for reversible pruning was now inactive (active-pruning fraction = 1.000). Hallucination rate fell to zero, flexibility collapsed to 0.095, and the prediction-error score fell to 1.958. Positive and negative pathology composites rose to 0.663 and 0.775. At parameter level, 14 780 weights were reversibly pruned, 1 641 irreversibly pruned, and 1 179 connections underwent aberrant amplification.

Default-dose treatment outcomes

At clinically representative doses the three interventions diverged (Table 3).

Ketamine increased combined accuracy by 12.2 % to 94.7 % and pushed cognitive-probe accuracy to 72.9 %. Capacity rose to 0.786 and effort to 0.740, while side-effect noise manifested at 0.035.

Antipsychotic medication gave the highest combined accuracy of the three (97.8 %, a 15.3 % gain) and restored extreme-stress performance to 98.3 %. Cognitive-probe accuracy climbed to 77.7 %. Side-effects remained minimal (0.014), but structural repair lagged behind ketamine: capacity was 0.681 and effort 0.660.

ECT achieved the broadest recovery. Combined accuracy reached 98.1 % (a 15.6 % improvement) and cognitive-probe accuracy 79.5 %, almost matching the healthy value. Capacity rebounded to 0.912 and effort to 0.874; however, side-effect noise peaked at 0.251 and the hallucination rate rose sharply to 0.732.

Table 3. Default-Dose Treatment Comparison

State	Clean Acc	Comb. Acc	Extreme Acc	Cognitive Probe	Halluc Rate	Neg Score	Capa-city	Effort	Side Effects	Full L1	Target-ed L1
Healthy	100.0%	98.5%	100.0%	79.7%	0.248	0.000	1.000	1.00	0.0000	--	--
Psychotic	99.6%	82.5%	49.8%	68.5%	0.000	0.004	0.678	0.65	0.0000	--	--
+ Ketamine	100.0%	94.7%	88.5%	72.9%	0.007	0.000	0.786	0.74	0.0350	0.01 3390	0.089 090
+ Anti-psychotic	100.0%	97.8%	98.3%	77.7%	0.015	0.000	0.681	0.66	0.0140	0.01 7050	0.005 597
+ ECT	100.0%	98.1%	100.0%	79.5%	0.732	0.000	0.912	0.87	0.2513	0.03 6525	0.191 422

Precision trajectories confirmed these qualitative differences (Table 4). Under ketamine, sensory gain fell to 0.24 with a noise level of 1.58; perceptual gain/noise moved to 0.47/0.85; belief-layer gain remained 2.50 while noise rose to 0.19; effort settled at 0.71. Antipsychotics drove these pairs toward 0.37/0.95, 0.68/0.47 and 2.06/0.04 with effort at 0.66. ECT produced the most "healthy-like" pattern, 0.57/0.74, 0.75/0.37 and 1.92/0.03 with effort 0.81.

Table 4. Precision and Effort Trajectories

Layer	Healthy (Gain/Noise)	Psychotic (Gain/Noise)	+ Ketamine (Gain/Noise)	+ Antipsychotic (Gain/Noise)	+ ECT (Gain/Noise)
Sensory	1.00 / 0.00	0.30 / 1.20	0.24 / 1.58	0.37 / 0.95	0.57 / 0.74
Perc-eptual	1.00 / 0.00	0.60 / 0.60	0.47 / 0.85	0.68 / 0.47	0.75 / 0.37
Belief	1.00 / 0.00	2.50 / 0.05	2.50 / 0.19	2.06 / 0.04	1.92 / 0.03
Effort	1.00	0.65	0.71	0.66	0.81

Removing consolidation epochs clarified how much benefit depended on post-treatment learning. Ketamine lost 29.7 % combined accuracy without consolidation (94.7 %→64.9 %), antipsychotics lost 6.9 % (97.8 %→91.0 %), and ECT was essentially unaffected (98.1 %→98.5 %).

Iso-dose comparisons and treatment efficiency

Dose–response sweeps demonstrated that no single therapy dominated across all perturbation budgets. When full L1 distance from baseline was equalised, antipsychotics were most efficient at small budgets, whereas ECT outperformed alternatives once the parameter budget increased. Ketamine often produced the most durable gains for a given perturbation. Expressed as absolute improvement per unit L1 change, antipsychotics peaked first, ketamine showed a pronounced biphasic curve, and ECT traded per-unit efficiency for larger absolute gains at high doses.

Chronic dynamics and durability

Long-term simulations revealed distinct temporal profiles (Table 5). During four simulated maintenance sessions and a subsequent withdrawal period:

Table 5. Chronic Treatment and Withdrawal Dynamics (Phase 8)

Treatment	Phase	Clean Acc	Combined Acc	Cognitive Probe	Halluc Rate	Neg Score	Cap-acity	Eff-ort	Side Effects
Ketamine	Baseline	99.7%	80.8%	67.2%	0.000	0.003	0.678	0.65	0.000
	Acute	100.0%	95.0%	72.3%	0.006	0.000	0.786	0.74	0.035
	Chronic	100.0%	96.5%	75.5%	0.110	0.000	0.880	0.84	0.038
	Withdrawal	100.0%	97.7%	78.4%	0.218	0.000	0.857	0.81	0.006
Anti-psychotic	Baseline	99.3%	81.1%	66.1%	0.000	0.007	0.678	0.65	0.000
	Acute	100.0%	98.0%	78.3%	0.014	0.000	0.681	0.66	0.014
	Chronic	100.0%	98.6%	79.3%	0.476	0.000	0.687	0.70	0.031
	Withdrawal	100.0%	98.4%	79.7%	0.570	0.000	0.686	0.70	0.023
ECT	Baseline	99.5%	80.4%	67.1%	0.000	0.005	0.678	0.65	0.000
	Acute	100.0%	98.5%	78.9%	0.760	0.000	0.912	0.87	0.251
	Chronic	100.0%	98.3%	79.7%	0.730	0.000	0.988	0.97	0.220
	Withdrawal	100.0%	99.2%	79.5%	0.706	0.000	0.986	0.97	0.063

Ketamine improved combined accuracy from 95.0 % (acute) to 97.7 % (withdrawal). Capacity climbed from 0.786 to 0.857, effort from 0.74 to 0.81, and side-effects decayed from 0.035 to 0.006.

Antipsychotics remained stable: combined accuracy hovered near 98 % in all phases, capacity stabilised around 0.686, effort near 0.70, and side-effects decreased only slightly (0.014 → 0.023).

ECT provided near-complete restitution: withdrawal accuracy was 99.2 %, capacity 0.986 and effort 0.97. Side-effects receded from 0.251 to 0.063. Durability—defined as retained benefit relative to acute gain—was 0.38 for ketamine, 0.83 for antipsychotics and 0.98 for ECT.

Patient heterogeneity

Ten virtual patients differing only in positive/negative severity ratios confirmed the group trends (Table 6). Mean combined-accuracy gains were 48.0 % ± 22.7 for ECT, 46.9 % ± 21.9 for antipsychotics and 32.4 % ± 25.1 for ketamine. Cognitive-probe accuracy averaged 79.0 % ± 0.6 with ECT, 78.1 % ± 2.0 with antipsychotics and 65.0 % ± 16.3 with ketamine.

Table 6. Individual Variability (Phase 9)

#	Seed	Pos Sev	Neg Sev	Psychotic Combined	Psychotic Cognitive	Ketamine Imp	Antipsychotic Imp	ECT Imp
1	42	1.2	0.8	64.8%	50.0%	+11.7%	+32.3%	+33.3%
2	137	0.8	1.2	51.8%	63.9%	+44.8%	+46.4%	+46.7%
3	256	1.0	1.0	77.0%	61.2%	+18.1%	+20.8%	+21.1%
4	314	1.5	0.5	18.6%	21.2%	+18.1%	+72.2%	+79.5%
5	501	0.5	1.5	23.5%	24.0%	+74.3%	+75.1%	+74.7%
6	619	1.1	1.1	55.6%	51.9%	+35.5%	+42.2%	+42.0%
7	733	0.9	0.9	94.5%	74.5%	+1.6%	+3.5%	+3.4%
8	842	1.3	0.7	43.6%	36.2%	+0.4%	+53.1%	+54.9%
9	951	0.7	1.3	31.1%	25.6%	+66.1%	+67.4%	+66.9%
10	1066	1.0	1.2	41.3%	30.4%	+53.3%	+56.3%	+57.3%

Subtype analysis showed that ECT was uniquely superior for all three positive-dominant cases and two of four negative-dominant cases, while antipsychotics edged the remaining balanced-profile patients. Ketamine ranked worst for every subtype, frequently exacerbating negative-symptom scores (mean increase 0.113).

Discussion

Results Interpretations

The simulation highlights crucial differences in how each intervention modifies the hierarchical system that underlies performance. Standard-dose antipsychotics almost completely restored combined and extreme-stress accuracy, yet left capacity and effort virtually unchanged. In clinical terms this mirrors the familiar picture on acute wards: voices quieten and delusional behaviour subsides, while apathy and blunted affect persist. From a computational angle the drug primarily retunes precision weighting; dopamine-driven aberrant salience is damped, bottom-up evidence regains influence, and the network can again test beliefs against reality [4,7]. Because the structural parameters remain largely untouched, negative symptoms endure—exactly what is seen in practice.

Electroconvulsive therapy produced the complementary pattern. Capacity climbed to 0.91 and effort to 0.87, the cognitive-probe score returned to near-healthy levels, and the network settled into a stable attractor that barely decayed after treatment stopped. Side-effects and a transient rise in hallucination rate were the trade-off. Clinically this resembles the short-lived confusion that follows ECT, yet the durability metric—0.98 retention of acute gains—suggests the procedure does more than "reset" the system; it appears to rebuild structural scaffolding that precision tuning alone cannot reach.

Ketamine occupied an uncomfortable middle ground. Synaptogenesis was genuine—pruning fell, capacity rose to 0.786, effort to 0.74—yet functional improvement was muted because precision parameters were driven further from healthy values. Without consolidation the raw pharmacological shift actually worsened performance; only subsequent retraining rendered the intervention therapeutic.

That profile—the need for learning to consolidate gains—fits the clinical impression of ketamine as a plasticity inducer that opens a window rather than supplying a finished product [8].

Iso-dose comparisons clarified these distinctions. When the overall magnitude of parameter change was held constant, antipsychotics remained the most efficient way to restore precision, ECT delivered the greatest benefit when a larger change budget was available, and ketamine consistently lagged on immediate functional indices despite superior structural repair. Quantity of perturbation therefore mattered less than which elements were perturbed—an observation that cautions against the simplistic assumption that "more plasticity is better."

Patient-level outcomes sharpened the clinical implications. Positive-dominant profiles responded robustly to antipsychotics and ECT but deteriorated on ketamine. Negative-dominant profiles improved under all agents, yet the structural gains unique to ketamine stood out when examined separately. In no subtype did ketamine monotherapy outperform the alternatives, but the data suggest that low-dose ketamine paired with a precision-stabilising agent could yield synergistic effects in patients with prominent negative symptoms—a prediction now open to empirical test.

Three overarching claims follow. First, negative symptoms require explicit structural interventions; they are not simply downstream sequelae of aberrant precision. Second, direction and targeting of change trump raw magnitude. Third, ketamine's therapeutic value is narrow and context-dependent, likely strongest as an adjunct rather than a stand-alone treatment.

These interpretations remain model-bound. A purely feed-forward architecture cannot mimic every feature of psychosis, and the behavioural tasks were necessarily simplified. Nevertheless, the internal coherence across dose schedules, chronic trajectories and heterogeneous profiles is striking. A single parameter set reproduced long-standing clinical observations and generated falsifiable predictions for stratified or combination therapy—precisely the sort of bridge computational psychiatry aims to build between mechanism and practice [9,10].

Model architecture and boundaries

The present network demonstrates how a single, transparent parameter set can reproduce a wide range of clinical phenomena. All manipulations—pathology, drug action, recovery—were expressed through explicit adjustments of gains, noises, excitability, effort, binary masks or individual weights inside one compact feed-forward classifier (input 2, hidden 256-256-128, output 4). This economy of design made it possible to blend the well-known precision-imbalance account of positive symptoms with a second, structural cascade that links pruning to capacity loss and motivational deficits. Because all interventions were coded in the same "currency" (the L1 distance each change imposed on the parameter vector), direct iso-dose comparisons became feasible: antipsychotics largely moved precision terms, ketamine exchanged precision for synaptogenesis, and ECT altered both domains simultaneously.

That same minimalism also sets clear limits. Real cortices are densely recurrent; higher areas do not simply boost incoming signals but can create perceptual content when evidence is weak. A purely feed-forward net cannot reproduce that generative loop, so when sensory noise was pushed to extremes the hallucination index paradoxically fell—an artefact of architecture, not a theoretical claim. Capturing authentic top-down hallucinations would require recurrent paths or explicit priors over sensory samples [4]. Likewise, the model includes no separate neuromodulatory channels, so the empirically documented link between dopamine and precision [7] or the glutamatergic contribution to prediction error [8] can only be mimicked indirectly through static gain changes.

Task design is another simplification. Classifying four Gaussian blobs—and a slightly harder probe—helps isolate capacity and effort, yet bears little resemblance to the socially embedded cognition disrupted in schizophrenia. A richer battery involving working memory, theory of mind or reward learning would almost certainly reveal additional dissociations that simple accuracy scores miss. Parameter values were also hand-set, not fitted. Doing so preserved a clean mapping between mechanism and symptom pattern, but any quantitative prediction (for example, the exact point where dose–response curves plateau) should be read as a hypothesis until the model is anchored to observations such as mismatch-negativity amplitudes, cortical thinning or longitudinal symptom scales.

These caveats point naturally to the next steps. Introducing recurrence and explicit feedback would test whether the same parameter shifts still produce the observed subtype patterns when real generative dynamics are allowed. Adding separate dopaminergic and glutamatergic signals would let the model capture precision changes and synaptic plasticity more faithfully. Finally, by fitting the network to individual EEG or MRI profiles the framework could move from proof-of-concept toward personalised prediction. Until then, the very sparseness of the current architecture is what makes its message easy to read: positive and negative symptoms arise from different computational faults, available treatments target distinct levels of the hierarchy, and the direction of change often matters more than the size of the dose.

Clinical implications

The modelling work shows that antipsychotics, ECT and ketamine do not perform interchangeable jobs. Each modifies a separate tier of the same hierarchical system, which helps explain why uniform prescribing strategies leave many patients unimproved (Table 7).

Antipsychotics acted as precise gain-setters. A routine dose almost normalised combined and extreme-stress accuracy and did so with the smallest side-effect term. Clinically that mirrors everyday experience: hallucinations fade, delusional actions stop, but apathy and cognitive drag stay. Within the model this looks like a tidy rebalancing of gain and noise—aberrant salience drops, sensory evidence regains weight and beliefs can again be tested against reality [4,7]. Because structural variables barely shift, negative features naturally persist. The flat dose-response seen in the parameter sweeps fits ward observations: once the precision mismatch is partly corrected, further escalation adds side-effects but little benefit.

ECT behaved more like a system rebuild. It simultaneously normalised precision and repaired structure: capacity climbed to 0.91, effort to 0.87 and cognitive-probe accuracy returned to near-healthy. Withdrawal hardly eroded those gains (0.98 retention, 104 % of the acute lift), in line with ECT's recognised durability in treatment-resistant cases. Acute confusion and an uptick in

hallucination rate also matched expectations—the model labels them as transient oversensitivity, not worsening psychosis. Taken together the findings support ECT when both positive and negative symptoms are pronounced or when antipsychotics have failed, and they give mechanistic weight to maintenance schedules that often exceed guideline recommendations.

Ketamine occupied an awkward middle ground. It triggered genuine synaptic repair—pruning fell to 0.158, capacity rose to 0.786, effort to 0.74—yet functional improvement stayed modest because precision was pushed farther from healthy values. When consolidation learning was removed, performance actually deteriorated, revealing a pure "plasticity window": pharmacology widens the window, learning has to walk through it [8]. Dose sweeps reproduced the familiar biphasic curve—best function at low intensity, best structural repair at high intensity—and explained the paradoxical post-withdrawal improvement: precision damage fades quickly, structural gains linger.

Across ten simulated patient profiles ketamine on its own was the weakest performer; positive-dominant cases even deteriorated. In negative-dominant cases structural advantages were clear but masked functionally by simultaneous precision disruption. The model therefore reframes rather than discards ketamine: capacity and effort restoration are unrivalled, yet they need parallel precision support. The obvious testable prediction is that low-dose ketamine combined with a precision-stabilising agent should beat either alone in patients with marked negative symptoms.

Subtype analyses reinforce that view. Positive-dominant profiles benefited most from antipsychotics or ECT and were harmed by ketamine. Negative-dominant profiles improved across the board but showed ketamine's unique structural recovery. Balanced profiles landed between the two extremes, with antipsychotics fractionally ahead. These patterns emerged without tailored symptom coding—only orthogonal parameter shifts—suggesting that a simple positive/negative severity ratio, complemented by an effort-sensitive probe, could guide first-line choice and enrich trials.

Three shifts in clinical thinking follow. First, negative symptoms demand structural interventions; they are not just residual positive pathology. Second, how a treatment moves the system matters more

than how much it moves it, explaining the limits of high-dose "more-is-better" strategies. Third, ketamine's usefulness is likely adjunctive and context-sensitive, not general purpose.

The discussion remains model-based and requires empirical validation, yet the internal agreement across default dosing, iso-dose matching, longitudinal trajectories and heterogeneous simulations is notable. A single parameter framework reproduced decades of clinical observation and set out falsifiable predictions for stratified and combination regimens.

Table 7. Computational mechanisms, clinical effects, and strategic implications of psychosis treatment modalities.

Treatment Modality	Computational Mechanism	Clinical & Structural Effects	Implications & Target Profile
Antipsychotics (D_2 Antagonists)	Precise Gain-Setters: Normalizes aberrant salience by restoring balanced precision weighting; reduces the weight of noisy bottom-up prediction errors without repairing structural deficits.	• Effect: Hallucinations fade and delusional actions cease; beliefs can be tested against reality. • Limitation: Negative symptoms (apathy, cognitive drag) persist as structural variables remain unchanged. • Dosing: Flat dose-response; escalation beyond precision correction adds side effects without benefit.	• Target: Positive-dominant profiles (high aberrant salience). • Strategy: Acts as a salience modulator rather than a comprehensive cure; effective for symptom suppression but insufficient for structural restoration.
Electroconvulsive Therapy (ECT)	System Rebuild / Network Reset: Simultaneously normalizes precision and repairs structure (capacity and effort); induces plasticity/neurogenesis to disrupt maladaptive attractor states.	• Effect: High efficacy for both positive and negative symptoms; restores small-world network organization. • Durability: High retention of gains (98%) post-withdrawal. • Side Effects: Acute confusion/hallucination uptick represents "transient oversensitivity" during reset.	• Target: Treatment-resistant cases; profiles with pronounced positive and negative symptoms. • Strategy: Supports maintenance schedules exceeding current guidelines; functions as a "network reset" enabling subsequent learning.
Ketamine / Glutamatergic Agents	Plasticity Window: Triggers synaptic repair (reversing pruning) via NMDA modulation; restores capacity and effort but may disrupt precision weighting (increasing noise).	• Effect: Unrivaled restoration of capacity and effort (negative symptoms). • Limitation: Weak monotherapy; can worsen positive symptoms in positive-dominant cases due to precision disruption. • Dosing: Biphasic curve (low intensity for function; high intensity for structural repair).	• Target: Negative-dominant profiles (adjunctive use only). • Strategy: Requires a "precision-stabilizing" partner (e.g., antipsychotics) to allow learning to utilize the widened plasticity window without increasing psychotic noise.

Note: Data derived from computational modeling of hierarchical systems and aberrant salience frameworks. "Structure" refers to synaptic connectivity/capacity; "Precision" refers to the weighting of prediction errors.

Solidity, novelty and potential impact

A notable strength of the present framework is its transparency. Symptoms, state transitions and treatment effects are encoded directly in the same parameter vector—gains, noises, excitability, effort, mask status and synaptic weights—so every behavioural change can be traced to an explicit numerical adjustment. The use of an L1 "dose" permits head-to-head comparisons that do not privilege any single mechanism: antipsychotics mainly rotate precision terms, ketamine exchanges precision for

synaptogenesis, and ECT moves both sets at once. Hard gating, capacity-dependent excitability and a stand-alone effort scalar give negative symptoms their own functional signature while the classic precision imbalance continues to account for positive symptoms. Revisions introduced in the second version—separating consolidation from acute drug action, dampening seizure noise in the ECT condition, adopting a logit-margin hallucination metric and normalising prediction-error scores—removed obvious confounds, and the qualitative pattern now repeats across default doses, sweep experiments, long-term trajectories and ten heterogeneous virtual patients.

The architecture is intentionally lean, and that deliberate simplicity imposes limits. Without recurrent pathways the classifier cannot create authentically top-down hallucinations; when sensory noise is extreme, bottom-up chaos swamps the belief layer and the hallucination index paradoxically falls, an artefact rather than a claim about pathophysiology. Real cortical processing will require feedback loops and explicit priors over sensory samples. Neuromodulatory dynamics are folded into static gain parameters, so dopamine–precision couplings [7] or glutamate-driven prediction-error updates [8] are represented only indirectly. Finally, the 2-D blob task captures capacity and effort effects but bears little resemblance to real-world cognition; richer batteries will be needed before quantitative claims can be taken to clinic.

Even with those caveats the model delivers something genuinely new. Previous work explained positive symptoms through predictive-coding failures [4] and tied negative symptoms to structural loss, yet no earlier study embedded both fault lines inside a single hierarchical network and then compared pharmacologically distinct treatments at the same perturbation cost. The iso-dose analysis, the durability comparison and the explicit positive/negative severity axis combine to generate predictions that have not appeared elsewhere—for example the forecast that low-dose ketamine, when paired with a precision-stabilising agent, should outperform either drug alone in deficit-dominant cases.

If future empirical work confirms even part of these forecasts, the translational implications are clear. Positive-dominant patients could start with precision-focused agents such as antipsychotics (or ECT

when rapid control is needed), whereas negative-dominant patients might benefit from adjunctive ketamine delivered in a protected precision environment. Simple bedside measures—a positive/negative ratio plus an effort-sensitive probe—could provide enough information to choose an initial regimen rationally rather than by trial and error. More broadly, the study shows how a fully parameterised model can link computational psychiatry [5] to everyday prescribing decisions, offering an auditable bridge between theory and practice.

Conclusion

The next technical steps are straightforward: add recurrence, separate dopamine and glutamate channels, and fit parameters to patient EEG or structural MRI. Conceptually, however, the essential point is already visible. Positive and negative syndromes arise from separable faults; treatments act on distinct hierarchical levels; and the direction of change matters more than its magnitude. A modest model, perhaps, but one that sets out a coherent agenda for mechanism-based personalisation.

References

[1] Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. Front Psychiatry. 2013;4:47. doi:10.3389/fpsyt.2013.00047

[2] Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. Trends Cogn Sci. 2012;16(1):72-80. doi:10.1016/j.tics.2011.11.018

[3] Cohen JD, Servan-Schreiber D. Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. Psychol Rev. 1992;99(1):45-77. doi:10.1037/0033-295X.99.1.45

[4] Sterzer P, Adams RA, Fletcher P, Frith C, Lawrie SM, Muckli L, Petrovic P, Uhlhaas P, Voss M, Corlett PR. The predictive coding account of psychosis. Biol Psychiatry. 2018;84(9):634-643. doi:10.1016/j.biopsych.2018.05.015

[5] Friston K. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 2005;360(1456):815-836. doi:10.1098/rstb.2005.1622

[6] Corlett PR, Honey GD, Krystal JH, Fletcher PC. Glutamatergic model psychoses: prediction error, learning, and inference. Neuropsychopharmacology. 2011;36(1):294-315. doi:10.1038/npp.2010.163

[7] Kapur S. Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry. 2003;160(1):13-23. doi:10.1176/appi.ajp.160.1.13

[8] Corlett PR, Honey GD, Fletcher PC. Prediction error, ketamine and psychosis: An updated model. J Psychopharmacol. 2016;30(11):1145-1155. doi:10.1177/0269881116650087

[9] Anticevic A, Murray JD, Barch DM. Bridging levels of understanding in schizophrenia through computational modeling. Clin Psychol Sci. 2015;3(3):433-459. doi:10.1177/2167702614562041

[10] Benrimoh D, Sheldon A, Sibarium E, Powers AR. Computational mechanism for the effect of psychosis community treatment: A conceptual review from neurobiology to social interaction. Front Psychiatry. 2021;12:685390. doi:10.3389/fpsyt.2021.685390