Context-dependent goal-directed control in schizophrenia spectrum disorders: Prioritizing loss avoidance over reward pursuit

Wolfgang Omlor[1]*, Giacomo Cecere[1]*, Akhil Ratan Misra[1]*, Gao-Yang Huang[2], Xiaoyan Wu[1], Victoria Edkins[1], Giulio Pergola[4], Boris Quednow[3,5], Philippe Tobler[3,6,7], Philipp Homan[1,3]

[1]Psychiatric Hospital, University of Zurich, Lenggstrasse 31, Zurich, 8032, Switzerland
[2]Swiss Federal Institute of Technology Zurich, Zurich, Switzerland
[3]Neuroscience Center Zurich, University of Zurich & Swiss Federal Institute of Technology Zurich, Zurich, Switzerland
[4]Lieber Institute for Brain Development, Johns Hopkins Medical Campus, Baltimore, Maryland; Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
[5]Experimental Pharmacopsychology and Psychological Addiction Research, Department of Adult Psychiatry and Psychotherapy, University Hospital of Psychiatry Zurich, University of Zurich, Zurich, Switzerland
[6]Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich 8006, Switzerland
[7]University Research Priority Program "Adaptive Brain Circuits in Development and Learning" (URPP AdaBD), University of Zurich, Zurich 8057, Switzerland.
*These authors contributed equally

Correspondence concerning this article should be addressed to:

Wolfgang Omlor, MD, PhD, Psychiatric Hospital, University of Zurich, Lenggstrasse 31, 8032, Zurich, Switzerland, E-Mail: omlor@bli.uzh.ch

Philipp Homan, MD, PhD, Psychiatric Hospital, University of Zurich, Lenggstrasse 31, 8032 Zurich, Switzerland. E-mail: philipp.homan@bli.uzh.ch

Abstract: While individuals with schizophrenia spectrum disorders show diminished reward pursuit, their ability to avoid negative outcomes appears relatively intact. The neural mechanisms of this dissociation and its relation to negative symptoms remain poorly understood. We hypothesized that in schizophrenia spectrum disorders, cognitive resources are preferentially directed toward avoiding losses rather than pursuing rewards, potentially limiting reward processing and contributing to negative symptoms. A two-step sequential decision task, which distinguishes between goal-directed and habitual strategies under reward and loss conditions, was completed by 50 stable individuals with schizophrenia spectrum disorders and 48 healthy controls during functional magnetic resonance imaging. Healthy participants showed more goal-directed (model-based) control in the reward setting than in the loss context. Individuals with schizophrenia spectrum disorders showed the opposite pattern, engaging model-based control more strongly in the loss context. These behavioral shifts were reflected in a prefrontal network. In anterior cingulate, orbitofrontal, and dorsolateral prefrontal regions, individuals with schizophrenia spectrum disorders showed increased activation during model-based learning in the loss condition. Within this group, loss-related activation in the anterior cingulate region was associated with anhedonia. Our findings suggest that schizophrenia spectrum disorders are associated with a reallocation of cognitive resources from reward pursuit towards loss avoidance. Enhanced loss processing in the anterior cingulate region may help explain the reduced capacity for pleasure seen in individuals with psychotic disorders.

Introduction:

How does the mind decide when it is worth investing cognitive effort to chase a reward versus avoid a loss? Everyday choices—double-checking a tax return to avoid a penalty or staying late at work to secure a bonus—force the brain to weigh potential gains against possible costs. In schizophrenia spectrum disorders (SSD), which are marked by profound disturbances in cognition and goal-directed behavior[1-3], this balance between reward pursuit and loss avoidance appears to be altered. Strikingly, the drive to pursue rewards is blunted, whereas the capacity to avoid losses remains intact or is even heightened[4,5]. This asymmetry raises the possibility that cognitive resources for goal-directed behavior are preferentially used to avoid losses at the expense of reward pursuit[6,7]. Yet this cognitive resource-reallocation hypothesis has never been tested within a computational framework.

The two-stage reinforcement-learning task, coupled with computational modeling, allows researchers to quantify the balance between two principal decision processes that place different demands on cognitive resources: model-based control is goal-directed, forward-looking, cognitively demanding, slow, and flexible, whereas model-free control is habitual, backward-looking, resource-efficient, fast and rigid[8-29]. The two-stage decision task has been used to demonstrate that compulsive conditions, including obsessive–compulsive disorder and substance dependence, are associated with reduced model-based control in reward contexts, but not in loss contexts[30,31]. Because model-based control hinges on cognitive resources[13,14] and dopaminergic signaling[32-35], both often compromised in psychotic disorders[2], individuals on the schizophrenia spectrum are expected to exhibit impaired model-based learning[3]. Indeed, in reward variants of the two-stage decision task, individuals with SSD have shown impairments in model-based control[36].

It remains unclear, however, how reward and loss contexts modulate the balance between model-based and model-free control in individuals with SSD. Resolving this knowledge gap is essential to determine whether psychosis shifts goal-directed cognitive control toward loss avoidance at the expense of reward pursuit, and whether this distortion drives reward devaluation and negative symptoms. We therefore aimed to test whether model-based control and its neural implementation are redirected from reward pursuit to loss avoidance in early psychosis, and how this reallocation relates to distinct symptom domains. Guided by evidence that SSD entail restricted cognitive capacity[1,2,37], value representation is dependent on cognition[6,7,38], and reward pursuit is attenuated relative to loss avoidance[4], we derived a set of specific expectations for behavior, neural engagement and their clinical correlates. We hypothesized that limited cognitive resources in psychotic disorders are preferentially allocated to goal-directed avoidance of negative outcomes rather than reward pursuit. We therefore predicted (i) preserved model-based control in individuals with SSD when choices

prevent losses but diminished goal-directed behavior when they procure gains. We further predicted (ii) that this goal-directed control asymmetry in SSD would be mirrored by a corresponding neural signature: stronger loss-than-reward activation in the three prefrontal hubs that instantiate model-based computations[3]—the anterior cingulate cortex (ACC) for cognitive resource allocation according to the expected value of goal-directed control[39-42], the orbitofrontal cortex (OFC) for prospective value representation[4,43,44], and the dorsolateral prefrontal cortex (dlPFC) for planning and working memory[45,46]. We further expected (iii) that the magnitude of this loss-biased recruitment would scale particularly with apathy-related negative symptoms (avolition, asociality, and anhedonia), which have been associated with a motivational asymmetry characterized by impaired reward seeking relative to preserved loss avoidance[5,47].

To test these predictions, we combined functional magnetic resonance imaging (fMRI) with the canonical two-step sequential decision task[8] (**Fig. 1a**) in 50 medicated, clinically stable individuals with SSD and 48 healthy controls. Separate reward and loss versions of the task were used during fMRI to map prefrontal activation in the ACC, DLPFC, and OFC (**Fig. 1b**) during model-based control. This design allowed us to quantify context-dependent shifts between model-based and model-free control and to characterize their neural signatures within these regions (**Fig. 1c**). Such mechanistic insights may inform interventions aimed at restoring adaptive motivational processing in SSD.

Results

Characteristic	HC	SSD	p-value
Demographics			
Age (years), mean (SD)	28.00 (5.57)	27.14 (6.42)	0.481
Sex (% female)	20 (41.7)	11 (22.0)	0.061
Ethnicity, n (% Caucasian)	38 (79.2)	40 (80.0)	
% Asian	6 (12.5)	6 (12.0)	0.995
% African	4 (8.3)	4 (8.0)	
Educational attainment, mean (SD)	2.23 (1.29)	1.66 (1.17)	0.024
Illness Characteristics			
Disease Duration (months), mean (SD)	NA	18.26 (29.63)	
Hospitalizations, mean (SD)	NA	6.96 (4.98)	
Medication Status			
Dosage, antipsychotic (CPZ eq.)	NA	277.20 (160.93)	
Clinical Ratings			
BPRS Total Score, mean (SD)	NA	39.04 (10.43)	
Suspicion	NA	1.94 (1.10)	
Hallucination	NA	1.84 (1.13)	
BNSS Total Score, mean (SD)	NA	40.44 (15.37)	
Asociality	NA	6.06 (2.65)	
Avolition	NA	6.50 (2.76)	
Anhedonia	NA	10.68 (4.63)	
Distress	NA	3.32 (1.49)	
Blunted Affect	NA	9.50 (4.55)	
Alogia	NA	4.34 (2.48)	
Cognitive Performance			
MATRICS - Attention	NA	0.97 (1.06)	
MATRICS - Working Memory	NA	1.47 (0.86)	
MATRICS - Verbal Learning	NA	1.59 (1.13)	
MATRICS - Reasoning	NA	1.44 (1.05)	
MATRICS - Visual Learning	NA	1.50 (0.90)	

Table 1. Participant characteristics. Clinical symptoms were assessed using the Brief Psychiatric Rating Scale (BPRS) for general psychopathology and the Brief Negative Symptom Scale (BNSS) for negative symptoms. Cognitive performance was evaluated with the MATRICS Consensus Cognitive Battery (MCCB), reported as standardized T-scores. Antipsychotic medication dosage is reported in chlorpromazine equivalents (CPZ eq.). Data are presented as mean (standard deviation) or count (percentage). Educational attainment is coded as follows: 0 = mandatory schooling, 1 = secondary education (high school or equivalent), 2 = vocational training, 3 = higher professional education, 4 = university degree. HC: Healthy controls. SSD: Schizophrenia spectrum disorders.

The SSD cohort (*n* = 50) and healthy controls (*n* = 48) showed no significant differences in age, sex, or ethnicity (**Table 1**). However, participants with SSD had lower educational attainment scores compared to controls (SSD: 1.66 ± 1.17; controls: 2.23 ± 1.29 years, *p* = 0.02). All participants with SSD received antipsychotic medication. Symptom severity was mild to moderate (BPRS total score: 39.04 ± 10.43). Negative symptom ratings on the BNSS showed both apathy-related symptoms (anhedonia: 10.68 ± 4.63; avolition: 6.5 ± 2.76;

asociality: 6.06 ± 2.65) and diminished-expression components (blunted affect: 9.5 ± 4.55; alogia: 4.34 ± 2.48).

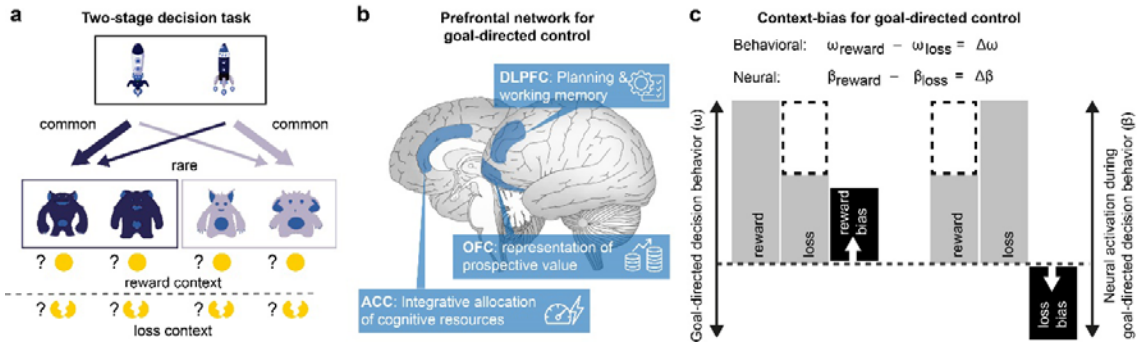

Figure 1: Experimental setting and modeling hypothesis. (a) *Two-stage decision-making task.* Adapted from Decker, et al. [48]. Participants made two consecutive choices to either earn play coins (*reward context*) or avoid losing play coins (*loss context*). Probabilistic transitions between the first and second stages (common vs. rare) allowed for the quantification of model-based and model-free control. In stage 1, each spaceship was associated predominantly (70% probability, 'common' transitions) with one of the two alien pairs in stage 2, and less frequently (30% probability, 'rare' transitions) with the other pair. **(b)** Schematic of the human pre-frontal network supporting model-based reinforcement learning. Dorsolateral pre-frontal cortex (DLPFC) supports planning and working-memory processes; orbitofrontal cortex (OFC) encodes prospective option values; anterior cingulate cortex (ACC) integrates these signals to allocate cognitive control. **(c)** Schematic showing the computation of context biases in goal-directed control. For each participant, the model-based weight ω and the task-evoked BOLD parameter β were estimated separately in *reward* and *loss* contexts (grey bars). Context bias was defined as the difference $\Delta\omega = \omega_{reward} - \omega_{loss}$ for behavior and $\Delta\beta = \beta_{reward} - \beta_{loss}$ for neural activation. Positive Δ indicates a *reward bias* (back bar with arrow pointing up), whereas negative Δ indicates a *loss bias* (back bar with arrow pointing down).

Decision behavior in reward and loss contexts: In the raw behavioral analysis (B_{raw}, see Methods), individuals with schizophrenia spectrum disorders (SSD) showed reduced model-based control compared to healthy controls (HC) in both reward (p = 0.005) and loss contexts (p = 0.014; **Supplementary Fig. 1**). Within-group comparisons revealed that HCs exhibited a numerically higher contribution of model-based control in the reward relative to the loss context, whereas individuals with SSD showed the opposite pattern—a numerically higher contribution in the loss relative to the reward context (**Fig. 2a, Supplementary Fig. 2a**). However, these context-dependent differences were not statistically significant (p = 0.28).

To obtain a more nuanced estimate of model-based versus model-free contributions, we next applied a hybrid reinforcement learning model. Consistent with the raw behavioral findings, the hybrid model confirmed reduced model-based control in individuals with SSD in the reward context ($p = 7.9 \times 10^{-13}$), but not in the loss context, when comparing individual-level ω parameters across groups (**Supplementary Fig. 2b**). The model further revealed a significant reward bias in HCs ($p = 3.6 \times 10^{-}$□) and a significant loss bias in the SSD group ($p = 2.9 \times 10^{-}$□). The difference in context-dependent biases between groups was

significant ($p = 2.8 \times 10^{-1}$□; **Fig. 2b**). We also examined whether ω-context bias was related to age, sex, illness duration, antipsychotic dose (chlorpromazine equivalents), or 5-HT_{2A} and D_2 receptor occupancy (see□Methods), but none of these associations reached significance. Across all contexts and groups, ω and B_{raw} were strongly correlated (r_s = 0.60–0.70, $p \leq 1 \times 10^{-}$□ in every case), indicating that B_{raw} reliably approximates the model-derived estimate and similarly captures the balance between model-based and model-free control. In the reward context, the posterior distribution of group mean differences confirmed that individuals with SSD had a lower model-based weight parameter (ω), as well as lower learning rates at both the first (α_1) and second stage (α_2), and reduced exploitation at the first stage (β_1) compared to HC (HC minus SSD, 95% HDIs; Δω: 0.16 – 0.68; Δα1: 0.04 – 0.86; Δα2: 0.16 – 0.65; Δβ1: 0.23 – 3.29; **Supplementary Fig. 3**). In the loss context, the posterior distribution indicated significantly lower second-stage learning rates and reduced exploitation at both stages in the SSD group, while the model-based weight parameter (ω) did not differ between groups (HC minus SSD, 95% HDIs: Δα2: 0.31 – 0.9; Δβ1: 0.61 – 3.57; Δβ2: 0.06 – 1.37; **Supplementary Fig. 4**).

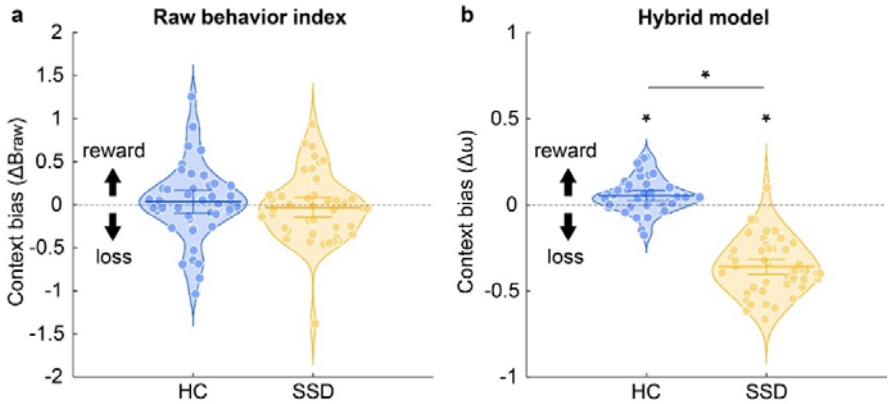

Figure 2: Analysis of decision behavior: (a) *Simplified raw behavior scores:* Regarding the relative contribution of model-based learning, healthy controls (blue) exhibited a numerically higher mean contribution of model-based control in the reward relative to the loss context, whereas individuals with schizophrenia spectrum disorders (SSD) showed a numerically higher mean contribution of model-based control in the loss relative to the reward context. When the context bias of raw behavior scores was compared between groups the difference was not significant. **(b)** *ω-scores from the modeling analysis:* Regarding the relative contribution of model-based learning, the ω-scores indicated a significant reward bias in healthy controls and a significant loss bias in individuals with SSD. In addition, the context bias of ω-values differed significantly between groups. Mean and 95% confidence intervals are shown. Asterisks indicate $p < 0.05$.

Both in the reward and loss setting, MCMC samples of all model parameters were well-mixed and converged to stationary distributions (**Supplementary Figs 5 and 6**). Convergence to stationary distributions was additionally indicated by the index of convergence (□ across parameters: 1.00 to 1.04 in the reward and loss setting). Parameter-recovery analysis confirmed that all seven hybrid-model parameters were

identifiable: across groups (HC, SSD) and contexts (reward, loss), all Spearman correlations between true and recovered values of the seven parameters were positive and significant ($p \leq 0.005$ in every case; ω: r_s = 0.7 – 0.93; for more details, see **Supplementary Figs 7 - 10**).

Context-dependent neural activations:

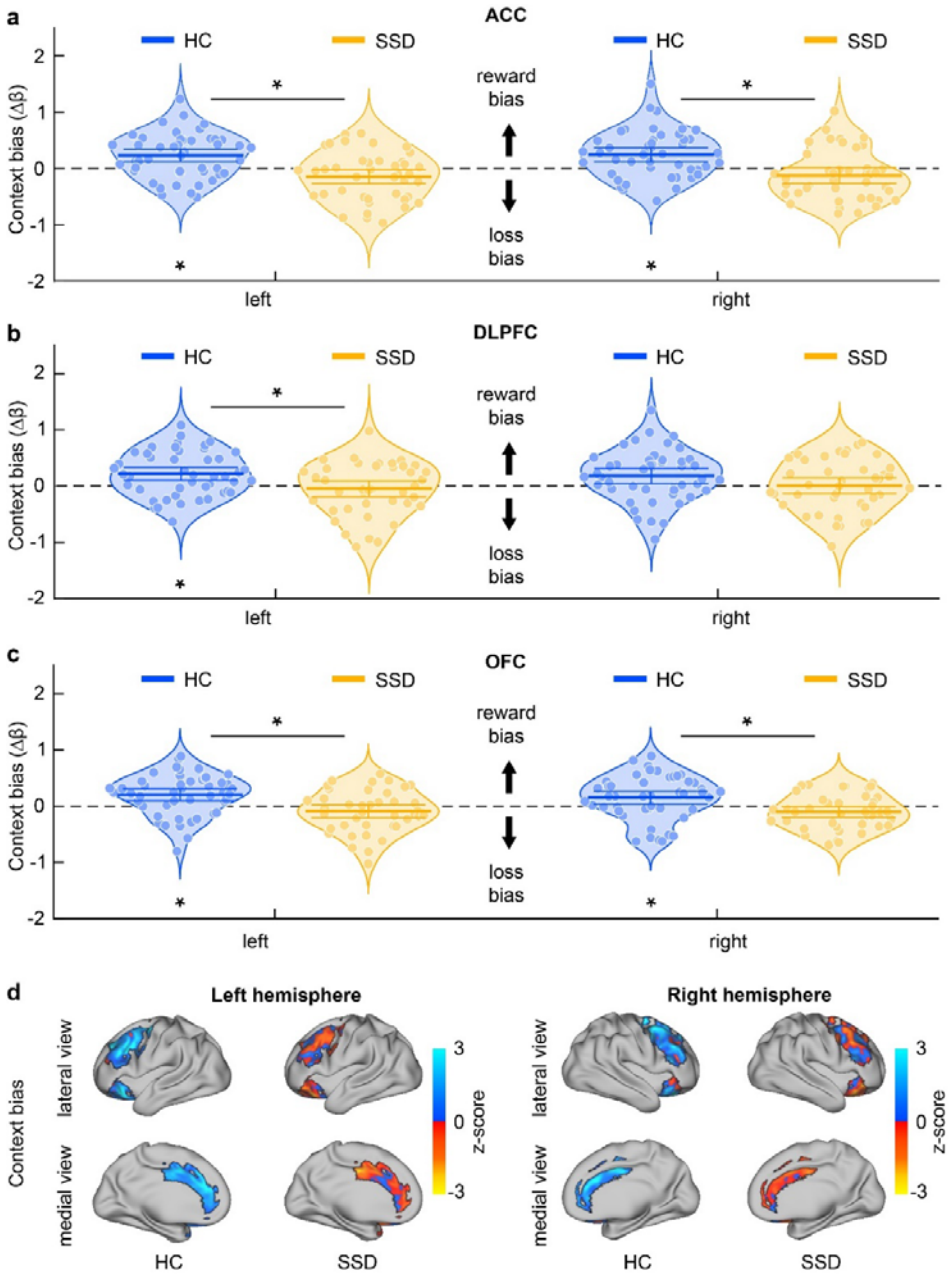

Figure 3: Context-dependent BOLD activation during model-based control: (**a-c**). The context bias of activation in the anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC) and orbitofrontal cortex (OFC) for healthy controls (HC, blue) and individuals with schizophrenia spectrum disorders (SSD, orange) is shown, plotted separately for the left and right hemispheres. Symbols indicate group means and 95 % confidence intervals. Asterisks indicate p < 0.05. (**d**) Brain maps

displaying context-biased activation within ACC, DLPFC, and OFC masks. Maps are presented unthresholded for visualization purposes.

In the ACC, healthy controls showed a significant reward bias in both hemispheres (*signrank test*; left and right: p = 0.001), while individuals with SSD exhibited a numerical loss bias. Group comparisons revealed a significant difference in context bias for both the left and right ACC (*ranksum test*; left and right: $p = 6*10^{-4}$, **Fig. 3a**). In the DLPFC, a significant reward bias was found in healthy controls in both hemispheres (*signrank test*; left: p = 0.002; right: p = 0.01). In the left hemisphere, this reward bias differed significantly from the numerical loss bias observed in individuals with SSD (ranksum test: p = 0.02, **Fig. 3b**). In the OFC, healthy controls showed a significant reward bias in both hemispheres (*signrank test*; left: p = 0.001; right: p = 0.01), while individuals with SSD demonstrated a significant shift toward a numerical loss bias (*ranksum test*; left: p = 0.001; right: p = 0.002, **Fig. 3c**). Context-biased activation within ROIs was predominantly observed in central portions of the DLPFC and OFC in both groups. Within the ACC, individuals with SSD exhibited loss-biased activation predominantly in dorsal regions, whereas reward-biased activation in controls was stronger in ventral regions (**Fig. 3d**).

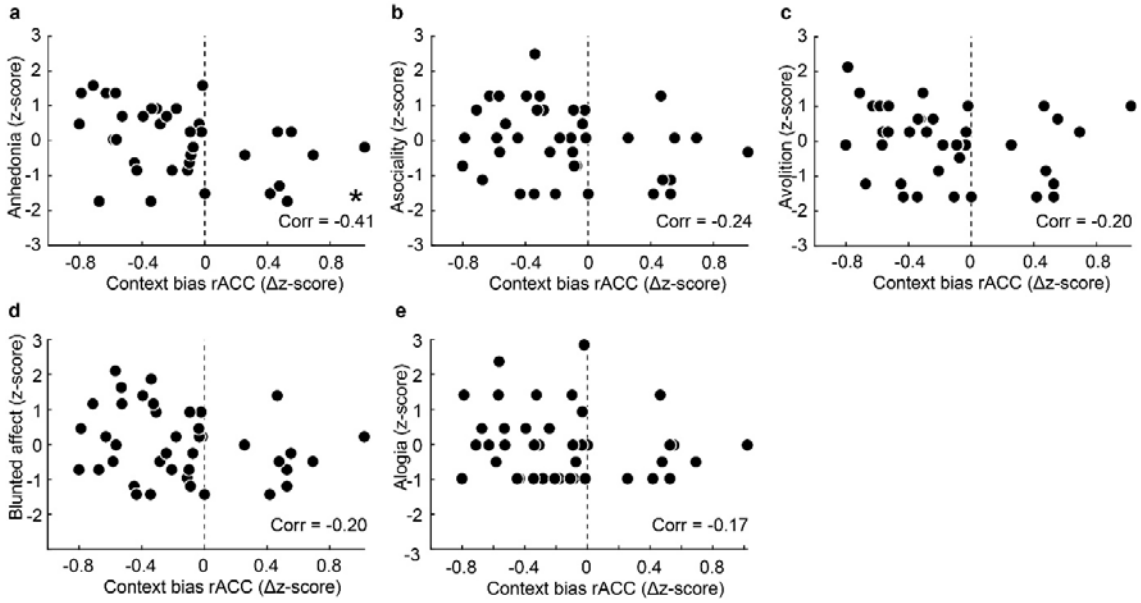

Figure 4: Relationship between right ACC context bias and negative symptom domains. (**a–c**) Scatter plots illustrating the association between the right anterior cingulate cortex (rACC) context bias and the negative symptom domains of apathy (anhedonia, asociality, and avolition). Spearman's correlation coefficients for each relationship are indicated within each plot. (**d, e**) Corresponding scatter plots for the negative symptom domains of diminished expression (blunted affect and alogia). An asterisk denotes statistical significance at $p < 0.05$.

We next examined whether context bias in any of these regions was associated with negative symptom domains. Only the context bias in the right ACC showed a significant

negative correlation with the symptom domain of anhedonia (r_s = -0.41; p = 0.048; **Fig. 4**). Thus, greater loss-biased activation in rACC was associated with greater anhedonia. As an additional control, we regressed anhedonia on right-ACC context bias while adjusting for age, sex, illness duration and antipsychotic dose (see Methods); the bias remained a significant negative predictor (β = –4.26 ± 1.51SE, p = 0.008). We also tested whether regional context bias correlated with age, sex, illness duration, antipsychotic dose (chlorpromazine equivalents), HT_{2A}- and D_2-occupancy (see Methods); none of these associations reached significance.

Apart from the a priori regions of interest (ACC, OFC, and dlPFC), our exploratory fMRI analysis identified additional areas showing reward- or loss-biased activation during model-based control (**Supplementary Fig. 11**). During model-based choices, individuals with SSD showed loss-biased activity centered in anterior and posterior cingulate as well as temporal regions, whereas reward-biased responses were observed mainly in occipito-temporal regions (**Supplementary Table 1**). By contrast, healthy controls expressed a broad reward-biased network spanning for instance cingulate and paracingulate regions, parietal, orbitofrontal and supramarginal regions, thalamus, angular gyrus as well as further medial and dorsolateral prefrontal regions, with only a small occipital-pole cluster showing the opposite loss-dominant pattern (**Supplementary Table 2**).

Discussion

The present study provides converging behavioral and neural evidence that SSD are characterized by a systematic re-allocation of goal-directed cognitive resources away from reward pursuit and toward loss avoidance. Relative to the reward-biased deployment of model-based strategies in healthy controls, individuals with SSD showed model-based control primarily when potential losses were at stake. This loss-biased strategy was mirrored at the neural level by preferential engagement of the ACC, OFC and DLPFC during model-based learning in the loss condition, and the degree of loss-biased ACC activation tracked the severity of anhedonia. Together, these results offer a mechanistic explanation for the clinical observation that individuals with schizophrenia show impaired pursuit of positive outcomes yet intact loss avoidance[4,5], while linking this asymmetry to the negative symptom domain anhedonia. It is worth mentioning that the raw behavioral analysis showed only a non-significant numerical context effect, while the hierarchical reinforcement learning model revealed a robust difference. By fitting the full choice history and jointly estimating learning-rate, perseveration and lapse noise, the hierarchical computational model isolates the underlying goal-directed signal that may be masked by excess variability in the raw

behavioral analysis. Hierarchical pooling further reduces within-group noise through shrinkage, increasing power to detect true between-group differences[49].

Canonical reinforcement-learning theory assumes that model-based control is deployed when the expected benefits of flexible planning outweigh its cognitive costs[3,8]. In healthy adults, appetitive contexts typically enhance the perceived value of planning, leading to robust model-based control in reward contexts, as observed in our control sample and in previous studies [43,45]. The loss-biased pattern in participants with SSD suggests that for these individuals the subjective cost–benefit calculus is distorted: avoiding losses is deemed more valuable than obtaining gains. One parsimonious interpretation is that patients discount prospective rewards because they anticipate diminished pleasure from realizing potential gains[50], whereas losses remain fully salient. In line with this idea, low "decision acuity"—a trait-like reduction in overall planning capacity marked by noisier, short-sighted choices[51]—could further restrict cognitive resources and channel them chiefly into preventing losses, thereby producing the loss-biased goal-directed control we observe. A complementary line of evidence shows that people with SSD are less willing to expend cognitive effort even for large rewards, pointing to a domain-general undervaluation of appetitive incentives[52,53]. These findings align with evidence for impaired value representation during reward processing relative to preserved loss avoidance in SSD[4], and our results frame this as the consequence of a strategic shift toward loss-oriented planning.

Our fMRI results reveal how the cognitive bias towards loss avoidance in individuals with SSD is implemented in the prefrontal cognitive network. In controls, reward-related increases in model-based control were accompanied by greater activation of ACC, OFC and DLPFC—regions known to orchestrate cost-dependent allocation of cognitive resources, value representation, and prospective planning according to goal-representation[3,46]. In individuals with SSD, the same circuit was preferentially engaged for model-based control when potential losses, rather than rewards, were at stake. These findings extend prior observations of blunted OFC and DLPFC responses to positive outcomes[7] by demonstrating that computations underpinning flexible decision-making are context-dependent and retuned toward loss avoidance in psychotic disorders. While loss-dominated activation in the OFC of individuals with SSD may signal a heightened valuation of avoiding negative outcomes, a similar activation pattern in the DLPFC might reflect increased recruitment of cognitive resources during the planning of loss avoidance compared to reward seeking[3,41,42,45]. The ACC emerged as a particularly pivotal hub: its loss-biased engagement was the only neural measure that covaried with a negative symptom domain. The ACC is ideally positioned to integrate motivational signals with task-control demands and to evaluate whether engaging the additional cognitive effort required for model-based learning is worthwhile[40-42]. A loss-

dominated encoding of action value within the ACC could therefore funnel limited cognitive resources into preventing negative outcomes, leaving insufficient capacity for goal-directed reward processing and manifesting clinically as anhedonia.

Our findings converge with mounting evidence that the ACC is central to psychotic disorders, with ACC dysfunction contributing to impairments in proactive control, and uniform structural features characterizing this region in schizophrenia[3,39,54]. Identifying a loss-biased planning mode in schizophrenia has potential translational relevance. First, it provides a neurocomputational phenotype that could be quantified and used to stratify patients or monitor treatment responses. Second, the association between anhedonia and ACC loss bias positions the anterior cingulate cortex as a promising target for circuit-based interventions. Non-invasive brain stimulation targeting the ACC or novel pharmacological interventions may re-balance resource allocation and thereby ameliorate motivational deficits. Our findings also extend earlier active-inference results showing elevated forgetting and reduced policy precision during action selection in early psychosis[55], and place the concomitant computational profile in a motivational framework that distinguishes loss from reward contexts.

This study has several limitations. First, it is cross-sectional and cannot establish whether loss-biased planning predates illness onset, emerges progressively or reflects a secondary adaptation to chronic disability. Longitudinal studies in high-risk and first-episode cohorts are needed to pinpoint the emergence of loss-biased cognitive processing in schizophrenia. Second, participants were medicated, and systematic modulation of cognitive processing and behavioral strategies by the complex receptor modulation of antipsychotic agents cannot be excluded. However, we found no associations between context-dependent brain activation or behavior and chlorpromazine equivalents, 5-HT_{2A} and D_2 receptor occupancy. These findings argue against the interpretation that loss-biased behavior and brain activation observed in individuals with SSD merely reflect medication effects. Third, a recent study showed that modeling parameters in reinforcement tasks can vary over time, potentially limiting their capacity for precision psychiatry[56]. Moreover, the monetary outcomes used here do not capture the rich spectrum of rewards and losses encountered in daily life, and virtual-reality tasks embedding social, vocational and other incentives could enhance ecological validity.

In summary, by integrating computational modeling with neuroimaging, we demonstrate that SSD are not marked by a uniform deficit in goal-directed cognition but by a selective redirection of cognitive resources toward loss avoidance. This behavioral shift is instantiated in a prefrontal network comprising ACC, OFC and dlPFC, and loss-biased engagement of the ACC is coupled to the negative symptom domain anhedonia. Our findings provide a

mechanistic explanation for the preservation of avoidance learning alongside impaired reward pursuit in schizophrenia and outline circuit-level targets for interventions aimed at restoring adaptive motivation.

Methods

Participants

Fifty individuals with early psychosis were recruited from the University Hospital of Psychiatry, Zurich, and 48 healthy controls were recruited through local advertisements. Psychosis diagnoses were established with a structured clinical interview and according to ICD-10 criteria. Eligible patients met criteria for schizophrenia, schizoaffective disorder, bipolar disorder with psychotic features, or acute polymorphic psychotic disorder, and were within five years of their first psychotic episode. In addition, eligible patients showed no evidence of substance-induced psychosis. All patients were clinically stable, in remission, and receiving antipsychotic medication at the time of testing. Healthy controls had no lifetime history of psychiatric or neurological disorders. The study was approved by the Cantonal Ethics Committee of Zurich, and all procedures adhered to its guidelines.

Clinical Assessments

Symptoms were assessed using several rating scales. The Brief Negative Symptom Scale (BNSS) was used to assess the different negative symptom domains of the apathy (avolition, asociality, anhedonia) and diminished expression (blunted affect, alogia) dimensions. To complement this domain-specific measure, overall psychopathology was evaluated with the Brief Psychiatric Rating Scale (BPRS), which provides a concise, clinician-rated index of positive, affective, activation and hostility symptoms in schizophrenia. Detailed cognitive functioning was characterized with the MATRICS Consensus Cognitive Battery (MCCB). All participants were required to pass a urine drug screen and a breathalyzer test.

Task Design

We used the two-stage task as described in previous research[8]. In total, participants completed 200 trials with 3 seconds allocated for each choice. The interval between first and second-stage stimuli was 1 second, followed by 1 second of feedback presentation after the second-stage choice. The sequence of events in each trial is illustrated in **Fig. 1a**. We employed the simplified version of the two-stage decision task described by Decker, et al. [48], which replaces abstract fractal cues with easily discriminable “spaceship” and “alien” stimuli. This adaptation with an embedded story line lowers perceptual and working-memory

demands, accommodating the cognitive impairments often observed in SSD, and provides a more intuitive, engaging environment for participants.

Each trial included two stages: In the first stage, participants used an MRI-compatible button box to select one of two spaceships. Choices had to be made within 2 seconds; otherwise, the trial was aborted. Once a choice was made, the selected option moved to the top of the screen, while the unselected option disappeared. In the second stage, participants were presented with one of two new states (two red or two purple aliens) each containing two additional options. The participants selected one option, and the outcome was displayed. In the reward setting, the outcome was either a monetary reward (represented by a coin) or no reward (represented by a red cross). In the loss setting, the outcome was either a monetary loss (represented by a broken coin) or no loss (represented by a green cross). Trials were separated by an inter-trial interval of randomized duration, averaging approximately one repetition time (TR). The second-stage state presented was probabilistically determined by the first-stage choice, based on the transition scheme shown in **Fig. 1a**. The position (left or right) of the options at each state was randomly varied between trials. Reward probabilities for each bottom-stage option were dynamic, fluctuating with independent Gaussian noise[8]. This dynamic structure encouraged ongoing learning. The order of reward and loss trial blocks (a block of 200 trials in each context) was randomized for each participant.

Before the fMRI task, participants completed a computerized training session to understand the task structure. In this training session, participants completed 50 practice trials. They were informed that reward probabilities would vary throughout the task, but the transition probabilities between the first and second stages would remain fixed. Participants were also told that each first-stage choice was predominantly associated with one of the two second-stage states, and the specific mapping was disclosed. Participants were informed that they would receive 50 CHF for task completion. Four participants with SSD did not engage with the task (always selected the same response option/pressed the same button during one stage)—and were therefore excluded from further analysis.

Behavior analysis

To examine learning effects without presupposing a particular reinforcement-learning model, we first asked how events on the previous trial shaped choice on the subsequent trial. We computed a stay-probability (StP) index based on the combinations of transition type (common/rare) and outcome (reward/loss). From this, we derived straightforward and intuitive “raw” metrics of model-based (MB_{raw}) and model-free (MF_{raw}) control (**Supplementary Fig. 1**). Model-based and model-free raw scores were calculated separately for reward and loss contexts as follows:

Reward context:

$$MB_{raw} = StP_{reward,\ common} - StP_{reward,\ rare} - StP_{no\ reward,\ common} + StP_{no\ reward,\ rare}$$

$$MF_{raw} = StP_{reward,\ common} + StP_{reward,\ rare} - StP_{no\ reward,\ common} - StP_{no\ reward,\ rare}$$

Loss context:

$$MB_{raw} = StP_{no\ loss,\ common} - StP_{no\ loss,\ rare} - StP_{loss,\ common} + StP_{loss,\ rare}$$

$$MF_{raw} = StP_{no\ loss,\ common} + StP_{no\ loss,\ rare} - StP_{loss,\ common} - StP_{loss,\ rare}$$

Separately for each context, we defined the balance between model-based and model-free control (B_{raw}) as the difference between the respective raw scores, providing a straightforward and intuitive metric to quantify their relative contributions:

$$B_{raw} = MB_{raw} - MF_{raw}$$

To quantify the context-dependent difference in the deployment of model-based versus model-free control, we defined the raw context bias (CB_{raw}) as the difference between the B_{raw} scores obtained in the reward and loss contexts, computed as follows:

$$CB_{raw} = B_{raw,\ reward} - B_{raw,\ loss}$$

We additionally deployed a hierarchical Bayesian reinforcement-learning model that incorporates the full choice history and individual learning rates, providing a generative, mechanistically explicit estimate of model-based versus model-free control. We fitted a hybrid reinforcement-learning model that integrates two algorithms: (i) model-based control, which computes action values through prospective (recursive) planning over the task's transition structure; and (ii) model-free control, which updates action values via reward-prediction errors, strengthening actions that are rewarded and weakening those followed by loss[8,57]. A key parameter, denoted as ω, determines the balance between these systems in shaping first-stage choices—where their predicted values diverge most. This weighting parameter is fixed across trials, with $\omega = 1$ reflecting fully model-based control, and $\omega = 0$ indicating purely model-free influence. Further model specifications and fitting procedures are detailed in the supplementary materials. Similar to ω, B_{raw} quantifies the relative contribution of model-based versus model-free learning, with higher values indicating greater reliance on model-based control. To assess the correspondence between the two metrics, we computed their Spearman correlation. To quantify the context-dependent difference in

the deployment of model-based relative to model-free control, we defined the context bias (CB) as the difference between the ω scores obtained in the reward and loss contexts (**Fig. 1c**):

$$CB = \omega_{reward} - \omega_{loss}$$

Hierarchical Bayesian parameter estimation was implemented through Stan (v. 2.18.2). Prior research has demonstrated that hierarchical parameter estimation provides more stable and reliable estimates compared to individual-level estimation[49]. Posterior distributions were approximated using the Markov Chain Monte Carlo (MCMC) algorithm. Each model was run with four chains, comprising 1000 burn-in samples and 5000 post-burn-in samples. Mixing and convergence of the chains to stationary values were assessed by computing the convergence index (□) and through visual inspection[49].

For the two-stage decision task, we utilized the *ts_par7* model that is implemented in *hBayesDM package in R* and incorporated both reward and loss contexts[49]. This model quantifies trial-by-trial decision-making dynamics, using seven parameters: α_1 (learning rate in stage 1), β_1 (inverse temperature in stage 1), α_2 (learning rate in stage 2), β_2 (inverse temperature in stage 2), pi (perseverance), ω (balance between model-based and model-free learning) and lambda (eligibility trace). To determine the optimal specification, we also fitted four- and six-parameter variants of the two-stage reinforcement-learning model and evaluated their performance using the leave-one-out information criterion (LOOIC)[49]. LOOIC favored the seven-parameter model over both the four- and six-parameter models (**Supplementary Table 3**).

fMRI Scanning protocol

Two consecutive fMRI scans were acquired. During the first (resting state), participants were instructed to relax and keep their eyes open. In the second, participants performed the modified version of the two-stage task in a “reward” and in a “loss” setting. Participants were monitored to ensure they remained awake and relaxed. They wore a respiratory belt to track breathing and were observed with a camera. Wakefulness was monitored throughout the scan, and customized head molds were used to reduce head motion. In addition, participants were provided with earplugs and headphones.

MRI was conducted on a Philips Achieva 3.0 T magnetic resonance scanner with a 32-channel SENSE head coil. During the 22-minute resting state sequence, 900 echo-planar imaging volumes were recorded (TR = 2000 ms, TE = 30 ms, matrix = 64 x 64, FOV = 240 mm, slice thickness = 3 mm, 40 continuous axial oblique slices). Each session began with the acquisition of structural images, including a sagittal localizer, a high-resolution T1-

weighted SPGR 3D spoiled gradient sequence (TR = 7.5 ms, TE = 3 ms, matrix = 256 × 256, FOV = 240 mm), and a T2 FSE sequence (TR = 6434 ms, TE = 102 ms, matrix = 256 × 256, slice thickness = 2.5 mm).

Anatomical Data Processing

The T1-weighted image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection [58], distributed with ANTs 2.5.0 [59], and used as T1-weighted reference throughout the workflow. The T1-weighted reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as the target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM), and gray-matter (GM) was performed on the brain-extracted T1-weighted image using fast [60]. Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.5.0) [59], using brain-extracted versions of both the T1-weighted reference and the T1-weighted template. The ICBM 152 Nonlinear Asymmetrical template version 2009c [61] was selected for spatial normalization and accessed with TemplateFlow.

Functional Data Processing

There was one BOLD run per subject. For pre-processing, a reference volume was first generated, using a custom methodology of *fMRIPrep*, for use in head motion correction. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated before any spatiotemporal filtering using mcflirt [62]. The BOLD reference was then co-registered to the T1-weighted reference using mri_coreg (FreeSurfer) followed by flirt with the boundary-based registration cost-function. Co-registration was configured with six degrees of freedom. Several confounding time series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS, and three region-wise global signals. FD was computed following Power, et al. [63] and Jenkinson, et al. [62]. FD and DVARS were calculated for each functional run, both using their implementations in *Nipype*[63].

The three global signals were extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow component-based noise correction[64]. Principal components were estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components were then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF + WM) were generated in anatomical space. The implementation differed from that of Behzadi, et al. [64] in that

instead of eroding the masks by 2 pixels on BOLD space, a mask of pixels that likely contain a volume fraction of GM was subtracted from the aCompCor masks. This mask was obtained by thresholding the corresponding partial volume map at 0.05, preventing extraction from voxels with a minimal fraction of GM. Finally, these masks were resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation).

Components were also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values were retained, such that the retained components' time series were sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components were dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each[65]. Frames that exceeded a threshold of 0.5mm FD or 1.5 standardized DVARS were annotated as motion outliers. Additional nuisance time series were calculated by means of principal components analysis of the signal found within a thin band (*crown*) of voxels around the edge of the brain.

All resamplings could be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using nitransforms, configured with cubic B-spline interpolation. After images were preprocessed with fMRIPrep, preprocessed images were smoothed using a Gaussian kernel with full width at half-maximum of 8mm and were high-pass filtered with a cutoff of 128s. Four participants with SSD were omitted from further fMRI analysis: two owing to excessive head motion during fMRI acquisition, and two because scanner malfunction produced incomplete imaging data.

Analysis of fMRI Data.

To investigate neural activity associated with task conditions, an event-related analysis was implemented using generalized linear modeling in FSL-FEAT. The analysis was based on the timeseries of Reward Prediction Errors (RPEs) as generated from the simulation of the model over each subject's task choices. This approach followed a framework inspired by Daw, et al. [8], concentrating on prediction errors occurring at two critical moments within each trial: the onset of second-stage stimuli and the onset of reward delivery.

For the first-level analysis, two onset regressors were included in the model (i) first-stage onset (ii) second stage and feedback onset time incorporated into the model as a single

regressor. The first onset regressor was parametrically modulated by the value of the chosen first-stage action and its partial derivative with respect to the ω-parameter from computational modeling (see also Supplementary Information). On the other hand, the second onset regressor was modulated by model-free prediction error and by the difference of model-free and model-based reward prediction error and its partial derivative with respect to the ω-parameter. Note that this difference regressor is zero at reward delivery as both algorithms converge at this point. To avoid confounding of the neural results due to activity differences between these two time points per se, within subject mean centering of the difference regressor was carried out. The analysis also included nuisance onset at feedback. Additionally, to correct for artefacts caused by physiological noise we included 6 motion parameters as estimated by fMRIPrep: a regressor for framewise displacement, a regressor to censor frames in the fMRI volumes with FD > 3mm, and an additional five noise regressors calculated using anatomical CompCor. To quantify context-dependent differences in neural activation associated with model-based versus model-free control, we defined the neural context bias (nCB) as the difference between β coefficients for model-based reward prediction errors estimated in reward versus loss contexts (**Fig. 1**):

$$nCB = \beta_{reward} - \beta_{loss}$$

We defined activation as 'loss-biased' in a given group when the significant group difference in nCB indicated a numerical reward bias in the comparison group shifting toward a numerical loss bias in the target group. Conversely, we defined activation as 'reward-biased' when a numerical loss bias in the comparison group shifted toward a numerical reward bias in the target group.

Region of interest and whole brain analysis

Guided by convergent evidence across mammalian studies, we selected the three prefrontal regions most consistently implicated in model-based control—the OFC, ACC and dlPFC—as primary regions of interest (ROIs) to examine their context-specific activation patterns [3,40,42,43,45]. Although some studies implicate additional regions such as the hippocampus in model-based control[66], a quantitative whole-brain meta-analysis in human participants found no consistent hippocampal involvement in standard two-stage decision tasks and concluded that the hippocampus plays a substantial role primarily when tasks include explicit spatial relations between decision stages[43]. We excluded striatal regions as primary ROIs because they encode a mixture of model-based and model-free signals[8,43], whereas our focus was on regions selective for model-based control. Our results also corroborate earlier evidence[8] that the ventral striatum encodes a mixture of model-based and model-free value signals (**Supplementary Fig. 12**). ROI masks were taken from the Harvard-Oxford cortical atlas to

ensure anatomical precision (see also Supplementary Information). Due to the absence of explicit labels for certain prefrontal subregions in the Harvard–Oxford atlas, the middle frontal gyrus was used as a proxy for the DLPFC [67]. Each ROI was further divided into left and right hemispheres to allow for lateralized analysis. These resulting binary masks were applied to extract the mean BOLD activity within each region for subsequent analyses. This anatomically guided approach ensured consistency across subjects and enhanced interpretability of ROI-based neural activation patterns during task performance. We additionally examined the spatial distribution of activation within each ROI. This analysis provided a finer-grained view of how each region supports model-based control in the reward relative to loss contexts. Because our ROI analysis was hypothesis-driven—and therefore might overlook regions that contribute to model-based control—we also conducted an exploratory whole-brain analysis.

Statistical analyses

Potential differences between individuals with SSD and healthy participants were assessed using two-tailed Wilcoxon rank-sum tests. Within-group differences across contexts (e.g., omega parameters in reward versus loss contexts within the schizophrenia group) were evaluated using two-tailed Wilcoxon signed-rank tests. Associations between context-dependent ROI activation bias/behavior bias and different variables (negative symptom domains, age, sex, chlorpromazine equivalents, illness duration, occupancy of 5-HT_{2A} and D_2-receptors[68]) were initially explored with Spearman's rank correlations to obtain robust estimates. Significant associations were further examined using general linear models (GLMs), adjusting for age, sex, antipsychotic dose, and illness duration. The association between true and recovered model parameters was investigated with Spearman's rank correlations. For all analyses involving multiple comparisons, p-values were corrected using FDR correction according to Benjamini-Hochberg. Statistical significance was defined as a FDR-corrected $P < 0.05$.

Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request.

Code availability

All codes used to compute and illustrate the findings of this study are available from the corresponding authors on reasonable request.

Acknowledgements

The authors thank Dr. Franziska Knolle for her critical reading of the manuscript and her valuable scientific advice. This work was supported by a NARSAD grant from the Brain & Behavior Research Foundation (28445, awarded to Philipp Homan) and by a grant from the Baugarten Stiftung via Stiftung für wissenschaftliche Forschung an der Universität Zürich (STWF-22-007, awarded to Wolfgang Omlor).

Author contributions

W.O. and P.H. designed the experiments. G.C. and W.O. performed the experiments. W.O., A.M., G.H. and G.C. analyzed the data. X.W., V.E., G.P., B.Q., P.T. contributed to the concept of the data analysis and manuscript. P.H. initiated and supervised the study. All authors approved, refined and edited the final version of the manuscript

Competing interests

Dr. P. Homan has received grants and honoraria from Boehringer Ingelheim, Janssen, Lundbeck, Mepha, Neurolite, and Novartis.

References

1 Marder, S. R. & Cannon, T. D. Schizophrenia. *N Engl J Med* **381**, 1753-1761, doi:10.1056/NEJMra1808803 (2019).

2 McCutcheon, R. A., Reis Marques, T. & Howes, O. D. Schizophrenia-An Overview. *JAMA Psychiatry* **77**, 201-210, doi:10.1001/jamapsychiatry.2019.3360 (2020).

3 Barch, D. M., Culbreth, A. J. & Sheffield, J. M. Cognitive Control in Schizophrenia: Advances in Computational Approaches. *Curr Dir Psychol Sci* **33**, 35-42, doi:10.1177/09637214231205220 (2024).

4 Gold, J. M. *et al.* Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch Gen Psychiatry* **69**, 129-138, doi:10.1001/archgenpsychiatry.2011.1269 (2012).

5 Marder, S. R. & Umbricht, D. Negative symptoms in schizophrenia: Newly emerging measurements, pathways, and treatments. *Schizophr Res* **258**, 71-77, doi:10.1016/j.schres.2023.07.010 (2023).

6 Kring, A. M. & Elis, O. Emotion deficits in people with schizophrenia. *Annu Rev Clin Psychol* **9**, 409-433, doi:10.1146/annurev-clinpsy-050212-185538 (2013).

7 Strauss, G. P., Waltz, J. A. & Gold, J. M. A review of reward processing and motivational impairment in schizophrenia. *Schizophr Bull* **40 Suppl 2**, S107-116, doi:10.1093/schbul/sbt197 (2014).

8 Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204-1215, doi:10.1016/j.neuron.2011.02.027 (2011).

9 Lee, S. W., Shimojo, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687-699, doi:10.1016/j.neuron.2013.11.028 (2014).

10 Shahar, N. *et al.* Credit assignment to state-independent task representations and its relationship with model-based decision making. *Proc Natl Acad Sci U S A* **116**, 15871-15876, doi:10.1073/pnas.1821647116 (2019).

11 Deserno, L. *et al.* Dopamine enhances model-free credit assignment through boosting of retrospective model-based inference. *Elife* **10**, doi:10.7554/eLife.67778 (2021).

12 Feher da Silva, C., Lombardi, G., Edelson, M. & Hare, T. A. Rethinking model-based and model-free influences on mental effort and striatal prediction errors. *Nat Hum Behav* **7**, 956-969, doi:10.1038/s41562-023-01573-1 (2023).

13 Otto, A. R., Gershman, S. J., Markman, A. B. & Daw, N. D. The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol Sci* **24**, 751-761, doi:10.1177/0956797612463080 (2013).

14 Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A. & Daw, N. D. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A* **110**, 20941-20946, doi:10.1073/pnas.1312011110 (2013).

15 Karthik, S. K., Pereira, K. M. K., Culbreth, A. J., Barch, D. M. & Moran, E. K. Higher motivation and pleasure scores predict more reliance on model-free

decision making. *Cogn Affect Behav Neurosci*, doi:10.3758/s13415-025-01302-3 (2025).

16 Daw, N. D. Are we of two minds? *Nat Neurosci* **21**, 1497-1499, doi:10.1038/s41593-018-0258-2 (2018).

17 Kool, W., Cushman, F. A. & Gershman, S. J. When Does Model-Based Control Pay Off? *PLoS Comput Biol* **12**, e1005090, doi:10.1371/journal.pcbi.1005090 (2016).

18 Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A. & Daw, N. D. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife* **5**, doi:10.7554/eLife.11305 (2016).

19 Seow, T. X. F. *et al.* Model-Based Planning Deficits in Compulsivity Are Linked to Faulty Neural Representations of Task Structure. *J Neurosci* **41**, 6539-6550, doi:10.1523/JNEUROSCI.0031-21.2021 (2021).

20 Vaghi, M. M. *et al.* Compulsivity is linked to reduced adolescent development of goal-directed control and frontostriatal functional connectivity. *Proc Natl Acad Sci U S A* **117**, 25911-25922, doi:10.1073/pnas.1922273117 (2020).

21 Grosskurth, E. D., Bach, D. R., Economides, M., Huys, Q. J. M. & Holper, L. No substantial change in the balance between model-free and model-based control via training on the two-step task. *PLoS Comput Biol* **15**, e1007443, doi:10.1371/journal.pcbi.1007443 (2019).

22 Sebold, M. *et al.* When Habits Are Dangerous: Alcohol Expectancies and Habitual Decision Making Predict Relapse in Alcohol Dependence. *Biol Psychiatry* **82**, 847-856, doi:10.1016/j.biopsych.2017.04.019 (2017).

23 Drummond, N. & Niv, Y. Model-based decision making and model-free learning. *Curr Biol* **30**, R860-R865, doi:10.1016/j.cub.2020.06.051 (2020).

24 Doll, B. B., Simon, D. A. & Daw, N. D. The ubiquity of model-based reinforcement learning. *Curr Opin Neurobiol* **22**, 1075-1081, doi:10.1016/j.conb.2012.08.003 (2012).

25 Collins, A. G. E. & Cockburn, J. Beyond dichotomies in reinforcement learning. *Nat Rev Neurosci* **21**, 576-586, doi:10.1038/s41583-020-0355-6 (2020).

26 Karagoz, A. B., Moran, E. K., Barch, D. M., Kool, W. & Reagh, Z. M. Evidence for shallow cognitive maps in Schizophrenia. *Cogn Affect Behav Neurosci*, doi:10.3758/s13415-025-01283-3 (2025).

27 Konovalov, A. & Krajbich, I. Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning. *Nat Commun* **7**, 12438, doi:10.1038/ncomms12438 (2016).

28 Economides, M., Kurth-Nelson, Z., Lubbert, A., Guitart-Masip, M. & Dolan, R. J. Model-Based Reasoning in Humans Becomes Automatic with Training. *PLoS Comput Biol* **11**, e1004463, doi:10.1371/journal.pcbi.1004463 (2015).

29 Knolle, F. *et al.* Investigating disorder-specific and transdiagnostic alterations in model-based and model-free decision making. *J Psychiatry Neurosci* **49(6)**, E389-E401 (2024).

30 Voon, V. *et al.* Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* **20**, 345-352, doi:10.1038/mp.2014.44 (2015).

31 Voon, V. *et al.* Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Transl Psychiatry* **5**, e670, doi:10.1038/tp.2015.165 (2015).

32 Wunderlich, K., Smittenaar, P. & Dolan, R. J. Dopamine enhances model-based over model-free choice behavior. *Neuron* **75**, 418-424, doi:10.1016/j.neuron.2012.03.042 (2012).

33 Deserno, L. *et al.* Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proc Natl Acad Sci U S A* **112**, 1595-1600, doi:10.1073/pnas.1417219112 (2015).

34 Doll, B. B., Bath, K. G., Daw, N. D. & Frank, M. J. Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *J Neurosci* **36**, 1211-1222, doi:10.1523/JNEUROSCI.1901-15.2016 (2016).

35 Sharp, M. E., Foerde, K., Daw, N. D. & Shohamy, D. Dopamine selectively remediates 'model-based' reward learning: a computational approach. *Brain* **139**, 355-364, doi:10.1093/brain/awv347 (2016).

36 Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M. & Barch, D. M. Reduced model-based decision-making in schizophrenia. *J Abnorm Psychol* **125**, 777-787, doi:10.1037/abn0000164 (2016).

37 Collins, A. G. E., Albrecht, M. A., Waltz, J. A., Gold, J. M. & Frank, M. J. Interactions Among Working Memory, Reinforcement Learning, and Effort in Value-Based Choice: A New Paradigm and Selective Deficits in Schizophrenia. *Biol Psychiatry* **82**, 431-439, doi:10.1016/j.biopsych.2017.05.017 (2017).

38 Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A. & Frank, M. J. Working memory contributions to reinforcement learning impairments in schizophrenia. *J Neurosci* **34**, 13747-13756, doi:10.1523/JNEUROSCI.0989-14.2014 (2014).

39 Omlor, W. *et al.* Estimating Multimodal Structural Brain Variability in Schizophrenia Spectrum Disorders: A Worldwide ENIGMA Study. *Am J Psychiatry* **182**, 373-388, doi:10.1176/appi.ajp.20230806 (2025).

40 Akam, T. *et al.* The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron* **109**, 149-163 e147, doi:10.1016/j.neuron.2020.10.013 (2021).

41 Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217-240, doi:10.1016/j.neuron.2013.07.007 (2013).

42 Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat Neurosci* **19**, 1286-1291, doi:10.1038/nn.4384 (2016).

43 Huang, Y., Yaple, Z. A. & Yu, R. Goal-oriented and habitual decisions: Neural signatures of model-based and model-free learning. *Neuroimage* **215**, 116834, doi:10.1016/j.neuroimage.2020.116834 (2020).

44 Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267-279, doi:10.1016/j.neuron.2013.11.005 (2014).

45 Smittenaar, P., FitzGerald, T. H., Romei, V., Wright, N. D. & Dolan, R. J. Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron* **80**, 914-919, doi:10.1016/j.neuron.2013.08.009 (2013).

46 Stein, D. J. *et al.* Obsessive-compulsive disorder. *Nat Rev Dis Primers* **5**, 52, doi:10.1038/s41572-019-0102-3 (2019).

47 Begue, I., Kaiser, S. & Kirschner, M. Pathophysiology of negative symptom dimensions of schizophrenia - Current developments and implications for

treatment. *Neurosci Biobehav Rev* **116**, 74-88, doi:10.1016/j.neubiorev.2020.06.004 (2020).

48 Decker, J. H., Otto, A. R., Daw, N. D. & Hartley, C. A. From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning. *Psychol Sci* **27**, 848-858, doi:10.1177/0956797616639301 (2016).

49 Ahn, W. Y., Haines, N. & Zhang, L. Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Comput Psychiatr* **1**, 24-57, doi:10.1162/CPSY_a_00002 (2017).

50 Kring, A. M. & Barch, D. M. The motivation and pleasure dimension of negative symptoms: neural substrates and behavioral outputs. *Eur Neuropsychopharmacol* **24**, 725-736, doi:10.1016/j.euroneuro.2013.06.007 (2014).

51 Moutoussis, M. *et al.* Decision-making ability, psychopathology, and brain connectivity. *Neuron* **109**, 2025-2040 e2027, doi:10.1016/j.neuron.2021.04.019 (2021).

52 Barch, D. M. *et al.* Dissociation of Cognitive Effort-Based Decision Making and Its Associations With Symptoms, Cognition, and Everyday Life Function Across Schizophrenia, Bipolar Disorder, and Depression. *Biol Psychiatry* **94**, 501-510, doi:10.1016/j.biopsych.2023.04.007 (2023).

53 Treadway, M. T., Peterman, J. S., Zald, D. H. & Park, S. Impaired effort allocation in patients with schizophrenia. *Schizophr Res* **161**, 382-385, doi:10.1016/j.schres.2014.11.024 (2015).

54 Brugger, S. P. & Howes, O. D. Heterogeneity and Homogeneity of Regional Brain Structure in Schizophrenia: A Meta-analysis. *JAMA Psychiatry* **74**, 1104-1111, doi:10.1001/jamapsychiatry.2017.2663 (2017).

55 Knolle, F. *et al.* Action selection in early stages of psychosis: an active inference approach. *J Psychiatry Neurosci* **48**, E78-E89, doi:10.1503/jpn.220141 (2023).

56 Vrizzi, S., Najar, A., Lemogne, C., Palminteri, S. & Lebreton, M. Behavioral, computational and self-reported measures of reward and punishment sensitivity as predictors of mental health characteristics. *Nature Mental Health* (2025).

57 Worbe, Y. *et al.* Valence-dependent influence of serotonin depletion on model-based choice strategy. *Mol Psychiatry* **21**, 624-629, doi:10.1038/mp.2015.46 (2016).

58 Tustison, N. J. *et al.* N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* **29**, 1310-1320, doi:10.1109/TMI.2010.2046908 (2010).

59 Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* **12**, 26-41, doi:10.1016/j.media.2007.06.004 (2008).

60 Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* **20**, 45-57, doi:10.1109/42.906424 (2001).

61 Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R. & Collins, D. L. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage* **47**, S102 (2009).

62 Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825-841, doi:10.1016/s1053-8119(02)91132-8 (2002).

63 Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* **84**, 320-341, doi:10.1016/j.neuroimage.2013.08.048 (2014).

64 Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* **37**, 90-101, doi:10.1016/j.neuroimage.2007.04.042 (2007).

65 Satterthwaite, T. D. *et al.* An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* **64**, 240-256, doi:10.1016/j.neuroimage.2012.08.052 (2013).

66 Miller, K. J., Botvinick, M. M. & Brody, C. D. Dorsal hippocampus contributes to model-based planning. *Nat Neurosci* **20**, 1269-1276, doi:10.1038/nn.4613 (2017).

67 Schienle, A., Kogler, W. & Wabnegger, A. A randomized trial that compared brain activity, efficacy and plausibility of open-label placebo treatment and cognitive repappraisal for reducing emotional distress. *Sci Rep* **13**, 13998, doi:10.1038/s41598-023-39806-y (2023).

68 Nyberg, S., Farde, L., Eriksson, L., Halldin, C. & Eriksson, B. 5-HT2 and D2 dopamine receptor occupancy in the living human brain. A PET study with risperidone. *Psychopharmacology (Berl)* **110**, 265-272, doi:10.1007/BF02251280 (1993).