Schizophrenia Bulletin Open
https://doi.org/10.1093/schizbullopen/sgac074

Belief Updating in Subclinical and Clinical Delusions

Sophie Fromm[*,1,2,3,4], Teresa Katthagen[1], Lorenz Deserno[5,6,7], Andreas Heinz[1,2,3], Jakob Kaminski[1], and Florian Schlagenhauf[*,1,2,3]

[1]Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health CCM, Department of Psychiatry and Neuroscience | CCM, NeuroCure Clinical Research Center, Berlin, Germany; [2]Charité – Universitätsmedizin Berlin, Einstein Center for Neurosciences Berlin, Berlin, Germany; [3]Bernstein Center for Computational Neuroscience, Berlin, Germany; [4]Department of Psychology, Humboldt-Universität zu Berlin, Germany; [5]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Würzburg, Würzburg, Germany; [6]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [7]Department of Psychiatry and Psychotherapy, Technische Universität, Dresden, Germany

[*] To whom correspondence should be addressed; Department of Psychiatry and Neurosciences Campus Charité Mitte Bonhoefferweg 3, 10117 Berlin, Germany; tel: +49 30 450 517289, e-mail: Sophie.fromm@charite.de

***Background and Hypothesis*: Current frameworks propose that delusions result from aberrant belief updating due to altered prediction error (PE) signaling and misestimation of environmental volatility. We aimed to investigate whether behavioral and neural signatures of belief updating are specifically related to the presence of delusions or generally associated with manifest schizophrenia.*Methods*: Our cross-sectional design includes human participants (*n*[female/male] = 66[25/41]), stratified into four groups: healthy participants with minimal (*n* = 22) or strong delusional-like ideation (*n* = 18), and participants with diagnosed schizophrenia with minimal (*n* = 13) or strong delusions (*n* = 13), resulting in a 2 × 2 design, which allows to test for the effects of delusion and diagnosis. Participants performed a reversal learning task with stable and volatile task contingencies during fMRI scanning. We formalized learning with a hierarchical Gaussian filter model and conducted model-based fMRI analysis regarding beliefs of outcome uncertainty and volatility, precision-weighted PEs of the outcome- and the volatility-belief.*Results*: Patients with schizophrenia as compared to healthy controls showed lower accuracy and heightened choice switching, while delusional ideation did not affect these measures. Participants with delusions showed increased precision-weighted PE-related neural activation in fronto-striatal regions. People with diagnosed schizophrenia overestimated environmental volatility and showed an attenuated neural representation of volatility in the anterior insula, medial frontal and angular gyrus.*Conclusions*: Delusional beliefs are associated with altered striatal PE-signals. Juxtaposing, the potentially unsettling belief that the environment is constantly changing and weaker neural encoding of this subjective volatility seems to be associated with manifest schizophrenia, but not with the presence of delusional ideation.**

Key words: schizophrenia/computational psychiatry /reward learning/prediction error/volatility/fMRI

Introduction

One characteristic of schizophrenia is the presence of delusions, defined as fixed beliefs that are not amenable to change despite conflicting evidence.[1] Delusional ideations are not only experienced by people with manifest psychiatric disorders, but also to varying degree distributed in the nonclinical population.[2–4] Altered belief updating and processing of environmental uncertainty was related to delusions in patients with schizophrenia (PSZ),[5–7] and recently to paranoid and fixed societal and political beliefs in the general population.[8] We aim to further elucidate to what extent altered belief updating as a proposed mechanisms of delusional formation can be considered symptom specific or is rather related to the syndrome of schizophrenia in general.

In the context of schizophrenia, the predictive-coding framework advanced the mechanistic understanding of delusions.[6,9,10] Within this framework, belief formation is seen as a hierarchical inference process where abstract and complex (higher-level) beliefs are continuously updated based on incoming (lower-level) sensory evidence.[6,11,12] According to this theory, an agent integrates

new evidence into the current belief to obtain an adequate representation of the environment and better predict future states. Prediction errors (PE) inform the updating process and signal how much the current belief deviates from the observed evidence. Critically in a Bayesian framework, the update magnitude depends on the certainty, also termed precision of prior belief and sensory evidence. A precision-weighted PE is thus scaled by this precision ratio, so that highly precise PEs lead to larger belief updates. One formalization of this theory is a hierarchical Gaussian filter model that not only includes the level of sensory evidence (1st) and belief (2nd), but also a 3rd level representing an estimate of environmental volatility.[13,14] This model considers that effective belief updating not only requires adequate precision-weighting but also balancing the belief stability-flexibility trade-off according to environmental dynamics.

In the brain, precision-weighted PEs are proposed to be encoded by dopamine.[15] Meta-analytic evidence indicates elevated presynaptic striatal dopamine synthesis and release capacity in PSZ[16,17] and people at a clinical high risk.[18] Aberrant dopaminergic neurotransmission may erroneously elicit highly precise PEs, so that generally irrelevant cues gain subjective significance. Aberrant PEs may signal that constant belief updating is required, and that the environment is highly volatile. Delusions could then result from an attempt to make systematic sense of these experiences and in a compensatory manner turn to be rigid.[19–21] Also behaviorally, a cognitive bias towards generally irrelevant information was observed in PSZ[22,23] and positively associated with positive psychotic-like experiences in non-clinical community samples.[24] This suggests a specific role of aberrant PE-signaling as a mechanism involved in delusion formation, along with an overestimation of environmental volatility. Yet in the long run, compensatory mechanisms may contribute to the characterizing rigidity of delusions.

Computational psychiatry accounts refine this picture by showing aberrant PE-signaling in PSZ,[23,25,26] people at clinical high risk,[14] and healthy people with delusional/paranoid beliefs.[8,27] In this context, the role of neural precision-weighting was highlighted in individuals spanning the psychosis spectrum by Powers et al.,[28] who investigated perceptual belief updating in people with and without psychotic illness, each group comprising people who do and who do not experience auditory hallucinations.[29] The study reports stronger sensory prior beliefs in people with hallucinations, whereas people with psychotic illness failed to represent environmental volatility. Till date, computational accounts dissociating within one framework to what extent altered belief updating relates to the presence of delusional beliefs, regardless of being part of manifest schizophrenia, are missing.[30]

Leveraging computational modeling and model-based fMRI, we investigated how PE-signaling and the representation of environmental volatility relates to delusion-like ideation and diagnostic status. To this end, we applied a 2 × 2 orthogonal design similar to Powers et al.[29] with four groups: Participants with and without diagnosed schizophrenia were delineated based on whether they report minimal or strong delusion-like ideation, resulting in healthy controls with high and minimal delusion-like ideations (HC D+ and HC + -) and schizophrenia patients with high and minimal delusions (PSZ D+ and PS + D-). Notably, when referring to delusions in HC, we mean delusion-like beliefs that are not necessarily clinically relevant or functionally impairing, but will refer to the term "delusions" for brevity. This design allowed us to test if behavioral and neural belief updating alterations indeed relate to the presence of delusions and/or to the diagnostic status of schizophrenia. We re-analyzed data from participants performing a probabilistic reversal learning task during fMRI-scanning[31–33] and adopted an analysis approach from Cole et al.[14] who used a hierarchical Gaussian filter model[13,34] to formalize belief updating in a similar task paradigm.

Methods and Materials

The analysis strategy was preregistered before data inspection and analysis (https://aspredicted.org/8j4u7.pdf).

Participant Characteristics

All individuals provided written informed consent prior to participation and received monetary compensation upon completion. The study was approved by the ethics committee of the Charité Universitätsmedizin and was conducted in accordance with the Declaration of Helsinki. Our sample consists of healthy control subjects (HC) and people with diagnosed schizophrenia (PSZ), each group comprising participants with strong (D+) and minimal delusions (D-) (figure 1) Recruitment of PSZ took place on the wards as well as in the outpatient unit of the Department of Psychiatry and Psychotherapy Charité – Universitätsmedizin Berlin, Campus Mitte. We determined that patients met ICD-10 criteria for schizophrenia and had no other psychiatric Axis I disorder (except substance use disorder), and that HC did not fulfill criteria for any past or present Axis 1 disorder. HC were further screened for severe medical or neurological conditions. Additional inclusion criteria were age over 18 years, MRI-compatibility, and fluency in German language. Initially, our sample included 86 participants. Since subsequent analyses are based on the computational model, we excluded 20 subjects where the

computational model did not fit better then chance (criterion: exp(log-likelihood/ntrials) > 0.5; details in the supplementary materials). All further analyses were conducted in a subset of 66 subjects (sample description table 1). Details on model selection and assessment of fit, as well as characteristics of the initial

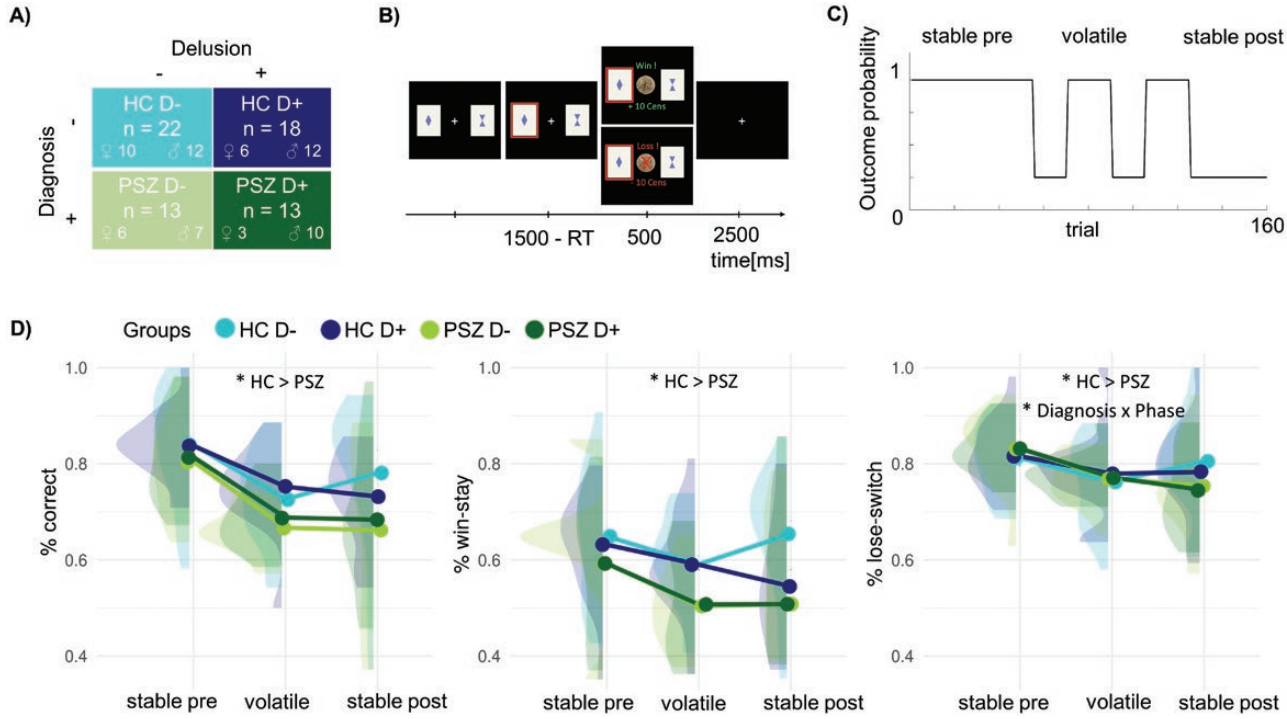

Fig. 1. (A) Sample classification, (B) trial sequence, (C) task contingencies and (D) raw behavioral indices compared across three task phases and between healthy controls (HC) and patients with schizophrenia (PSZ), each group with minimal (D−) or with strong delusions (D+).

Table 1. Demographic, cognitive, and clinical characteristics with test-statistics from two-way ANOVAs factorized by diagnosis and delusion.

	Healthy Controls		Patients with Schizophrenia		ANOVA			
	HC D−	HC D+	PSZ D−	PSZ D+	Diagnosis-effect		Delusion-effect	
N	22	18	13	13				
	Mean ± SD	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*F*	*p*	*F*	*p*
Age (years)	25.6 ± 4.84	24.8 ± 5.36	35.2 ± 7.19	32.5 ± 7.88	**31.12**	**<.001**	1.33	.253
Word memory	33.09 ± 3.48	34.39 ± 3.79	32.00 ± 4.92	30.83[a] ± 5.36	.01	.93	2.48	.120
Word list	9.05 ± 1.46	9.56 ± 0.78	9.15 ± 0.99	9.42[a] ± 0.90	**4.41**	**.04**	0.06	.800
Digit Span	7.50 ± 2.26	8.89 ± 2.35	6.62 ± 2.02	6.92[a] ± 1.98	**6.33**	**.015**	2.78	.101
Trail Making Test-A	22.62 ± 6.67	24.39 ± 7.76	39.46 ± 16.42	36.83[a] ± 15.79	**25.25**	**<.001***	< 0.001	.990
Trail Making Test-B	52.48 ± 20.94	48.72 ± 12.34	82.00 ± 51.40	77.08[a] ± 31.41	**14.16**	**<.001***	0.31	.577
Digit Symbol Substitution Test	84.32 ± 16.25	80.17 ± 19.31	66.46 ± 16.91	65.17[a] ± 13.28	**16.13**	**<.001***	0.79	.378
IQ-composite	.308 ± 0.458	.459 ± 0.428	-0.287 ± 0.804	-0.228 ± 0.544	**20.44**	**<.001***	0.55	.462
Peters' Delusion Inventory	1.86 ± 2.05	12.22 ± 2.07	7.35 ± 4.54	9.20[c] ± 3.22	2.83	.097	**71.51**	**< 0.001**
PANSS-total	/	/	72.00 ± 21.31	100.15 ± 20.59	/	/	**11.7**	**.002**
PANSS-positive dimension	/	/	8.38 ± 1.71	18.23 ± 3.59	/	/	**79.9**	**< 0.001**
PANSS-disorganized dimension	/	/	7.46 ± 3.6	11.00 ± 3.0	/	/	**7.42**	**.012**
PANSS-negative dimension	/	/	18.2 ± 7.12	19.8 ± 6.78	/	/	**0.33**	**.57**
Chlorpromazine Equivalens (mg)	/	/	312 ± 189	338 ± 259	/	/	0.50	.770
$N_{unmedicated}$			1	4				

PANSS = Positive and Negative Syndrome Scale; *p** = significant after Bonferroni-correction for multiple testing. [a]One, [b]two or [c]three participants with missing values are not included in the summary statistics. Since we observed no interactions, only main effects are reported.

sample are reported in the supplementary materials (table S2-S4). As specified in the supplement table S1, data from a subset of 75 participants was published before.[31–33]

Clinical and Cognitive Measures

We assessed delusion-like ideation in HC with the 21-item version of Peters' Delusion Inventory (PDI).[35,36] HC were re-contacted from an online sample of 1059 people who completed the PDI online. They were stratified into D+ if they had a PDI total score in the upper quartile of the total population (PDI total > 9) as described in a previous study.[37] Participants drawn from the lower end of the PDI distribution (PDI total < 7) were included in the D− group. Critically, a high total score does not imply clinical relevance. Therefore, delusions in HC, as defined by PDI, refer to subclinical delusion-like ideations. Since the PDI is not intended and validated specifically for PSZ, the assessment of delusion severity in PSZ was based on clinician-rating (P1 delusion-item of the PANSS interview).[38] Patients with a P1-score < 3 were grouped as D− (none or minimal delusions) and P1 > 4 as D + (at least moderate severity). Patients with a P1-score of 3 or 4 were excluded from the current analysis. For clinical description, we report PANSS dimensions for positive, disorganized[39] and negative symptom dimension[40] according to Wallwork et al. (2012) and Galderisi et al. (2021) and compared those between PSZ D+ and PST D−. To compare self-reported delusion-like ideation across all four groups, the total PDI-score was analyzed using a two-way ANOVA with the factors diagnosis- and delusion-status. Cognitive functioning was assessed in terms of verbal intelligence (vocabulary) and memory (word list), working memory (digit span), attention (Trail Making Test-A), executive functioning (Trail Making Test-B), and cognitive speed (digit symbol substitution test) in all subjects. We computed a 2 × 2 MANOVAs with the factors diagnosis and delusion to compare the four groups on all cognitive indices, with *p*-values adjusted for the number of cognitive indices tested ($p < 0.008$).

Behavioral Task

Participants performed 160 trials of a probabilistic reversal-learning task with stable and volatile task phases[31,32] during fMRI acquisition (figure 1). Participants chose between two card stimuli that were presented in random ordering on the right and left side on the screen. Choices were reinforced with a win or loss of 10 Eurocents. One card was initially rewarded more often (80% of the trials), whereas the other card was rewarded less frequently (20%). In the first 55 trials, the reward contingency remained stable, followed by a volatile task phase, where the anti-correlated contingencies changed every 15–20 trials. The last block consisted of 35 stable trials with a reversal of the more rewarding card from the first phase.

Computational Modeling

Models were fitted to individual trial-by-trial choice data using the same model space as Cole et al.[14] with the perceptual models: i) Rescorla Wagner learning model, ii) Two-level Hierarchical Gaussian Filter (HGF2), iii) Three-level HGF (HGF3) iv) Three-level "mean-reverting" HGF (HGF3-rev) with a subject-specific equilibrium parameter (*m3*). The *m3* parameter represents an average volatility estimate and attracts trial-wise estimates like an individual set point. Choice probabilities were computed using a softmax (logistic sigmoid) function. For each perceptual HGF model, we applied two response models, either with i) fixed decision noise for all trials or ii) free decision noise that varies trial-by-trial as a function of the estimated 3rd level volatility of the stimulus-outcome probabilities. Following Bayesian model selection, we based further analysis on the best-fitting model and included only subjects with a model fit better than chance. A model was classified to fit the subject behavior better than chance when the geometric mean likelihood per trial (given by exp(log-likelihood/ntrials)) exceeded 0.5, corresponding to $p < 0.05$ in a binomial test. Priors were set according to the behavioral modelling code in the repository from Cole et al.[14] (https://gitlab.ethz.ch/dandreea/apup). To ensure comparability of modelling results, we did not alter these priors to fit our data. Details regarding the computational modelling, model and parameter comparisons are described in the supplementary materials.

Functional Magnetic Resonance Imaging Analysis

After preprocessing (supplementary methods), we conducted an event-related analysis using the general linear model approach implemented in SPM12 similar to Cole et al.[14] Due to the short decision period, we could not set up a separate base regressor for the decision period. Instead, we modeled one base regressor that spanned the entire trial from cue onset to feedback offset.[31,41] We included four parametric modulators from the best-fitting HGF-model: μ_3 and σ_2 updated at feedback onset, whereas ε_2 and ε_3 modeled only the feedback period (non-orthogonalized). Additionally, six movement regressors and their temporal derivatives were included as regressors of no interest, plus additional regressors to flag scan-to-scan movement > 1 mm. On the group-level, we applied a flexible factorial design implemented in SPM12[42] to compute a (2 × 2 × 5) ANOVA with the factors diagnosis (HC vs. PSZ), delusion (D + vs D−) and regressor, which comprised the first-level contrast images related to the computational parameters μ_3, σ_2, ε_2, and ε_3, plus age and gender as covariates of no interest.

Appropriate contrasts examined the average activation across all groups (task man effects) for each computational parametric modulator, the main effects for delusion [(HC D+ & PSZ D+) vs. (HC D− & PSZ D−)] and diagnosis [(HC D+ & HC D−) vs. (PSZ D+ & PSZ D−)] and the delusion by diagnosis interaction. Significance level was set to $p < 0.05$ FWE-corrected at the whole brain level or for the respective main effects to correct orthogonal group comparisons. Similar to Cole et al., small volume correction was applied for all regressors using a midbrain mask[41] and for ε_2 and ε_3 using an ACC mask (derived from the WFU PickAtlas[43]).

Results

We observed significant effect of diagnosis, with PSZ (PSZ D + and PSZ D−) being significantly older than HC (HC D + and HC D−) and performing worse on cognitive tasks of attention, executive functioning, and cognitive speed (table 1). No significant effect of delusion was present. The four groups did not differ in gender ($X^2(3, 66) = 2.28$, $p = 0.518$, figure 1). As reported in table 1, patients with delusions (PSZ D+) scored higher on the PANSS positive and disorganized dimension as well as on the total score, compared to PSZ D−. PDI was significantly higher in participants with strong delusions (HC D + and PSZ D+) compared to participants with minimal delusions (HC D− and PSZ D−), and showed a diagnosis by delusion interaction ($F(1,59) = 30.84$, $p < 0.001$), due to significant differences between all subgroups, but D + and D− in PSZ ($t(59) = -1.48$, $p = 0.456$) and between PSZ and HC in D + groups ($t(59) = 2.63$, $p = 0.052$).

Computational Model Comparison

Across all subjects, within our model space the 3-level "mean reverting" HGF (HGF3-rev) with dynamic decision noise best fitted task behavior (PXP~100, figure S1 and table S5). We computed a 2 × 2 ANOVA, with factors Diagnosis and Delusion to compare model fit of the HGF3-rev model between groups. There were no differences regarding the model log-likelihood between D + and D− ($F = 1.24$, $p = .271$). However, in HC the computational model fit significantly better than in PSZ ($F = 8.85$, $p = .004$). Computational modelling in four subjects did not converge and individual model fit of HGF3-rev in sixteen subjects did not exceed chance levels, resulting in a sample of n = 66 (for sample size of subgroups, see table 1 and figure 1), which will be used in all reported analyses. Following computation of individual model fit, we identified outliers as specified in the supplementary materials and thus labelled in total 12 trials of three subjects (2 subjects in SZ D + and 1 subject in SZ D-) as missing.

Raw Behavioral Results

HC chose the correct card more often than PSZ and showed less choice-switching after correct choices (table 2, figure 1). We observed no effect of delusions on accuracy and choice-switching. Correct choices differed significantly across phases ($F(1.55,95.85) = 22.72$, $p < .001$; Greenhouse-Geisser corrected for non-sphericity) due to more correct choices in the stable-pre compared to the volatile ($t(124) = 5.99$, $p < .001$) and stable-post phase ($t(124) = 5.68$, $p < .001$). Win-stay ($F(1.74,107.89) = 8.1$, $p < .001$) and lose-switch behavior ($F(1.60, 99.08) = 13.82$, $p < .001$) also differed across phases with a similar pattern. We observed a significant interaction of diagnosis by phase on lose-switch behavior ($F(1.6, 99.08) = 3.96$, $p = .031$). Lose-switch decreased from the stable-pre to the volatile phase in HC ($t(124) = 3.04$, $p = .03$) and PSZ ($t(124) = 3.46$, $p = .009$), but differed between

Table 2. Raw behavioral and modeling parameters with descriptive and test-statistics from two-way ANOVAs factorized by diagnosis and delusion.

	Healthy Controls		Patients with Schizophrenia		2 × 2 ANOVA			
	HC D-	HC D+	PSZ D-	PSZ D+	Diagnosis-effect		Delusion-effect	
	Mean ± SD	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*F*	*p*	*F*	*P*
% correct	79 ± 8	78 ± 4	75 ± 7	72 ± 8	**8.8**	**.004**	0.4	.513
%win stay	92 ± 8	92 ± 8	80 ± 15	82 ± 8	**7.3**	**.009**	0.6	.433
%lose switch	36 ± 17	43 ± 17	36 ± 11	40 ± 12	0.3	.575	0.0	.970
*m*3	1.60 ± 0.39	1.50 ± 0.35	1.83 ± 0.72	1.83 ± 0.45	**7.63**	**.008**	0.10	.521
$\mu_3^{(k=0)}$	1.11 ± 0.58	1.31 ± 0.62	1.38 ± 0.78	1.17 ± 0.71	0.04	.845	0.01	.943
ω	-2.94 ± 0.63	-3.04 ± 1.02	-2.98 ± 0.66	-3.57 ± 2.06	< 0.01	.953	0.34	.560
κ	1.08 ± 0.19	1.04 ± 0.09	1.11 ± 0.28	1.02 ± 0.21	0.17	.680	1.01	.319
ϑ	0.27 ± 0.04	0.25 ± 0.06	0.26 ± 0.04	0.24 ± 0.04	2.15	.148	1.19	.279
β	28.90 ± 10.00	25.90 ± 8.12	25.40 ± 8.98	21.20 ± 7.35	3.24	.077	1.83	.182

m3 = equilibrium parameter; μ3(k = 0) = initial volatility estimate; κ = coupling of 2nd and 3rd level belief; ω = tonic component (log-volatility) of the outcome belief; ϑ = meta-volatility; β = inverse decision temperature of response model.

stable-pre- and stable-post phase only in PSZ ($t(124) = 4.65$, $p < .001$), but not HC ($t(124) = 1.4$, $p = .72$).

Group Differences in Computational Model Parameters and Trajectories

We examined group differences on MAP estimates from the best fitting model (HGF3-rev), which included six free parameters ($m3$, $\mu_3^{(k=0)}$, κ, ω, ϑ, β, table 2 and figure S2A). A MANOVA suggested a significant effect of schizophrenia diagnosis ($F(6,57) = 4.02$, $p = .002$; Wilks' Λ=.703) and of present delusion ($F(6,57)$=2.4, $p = .039$; Wilks' $\Lambda = 0.798$). Robust post hoc tests did not confirm any significant delusion effect but revealed higher equilibrium parameter ($m3$) in PSZ compared to HC. An increase in this parameter suggests that volatility estimates in the patient group converged to higher levels. To verify that this effect was not driven by worse model fit in PSZ or group difference in cognitive capacity, the MANOVA was repeated with LME and cognitive capacity as covariates. The diagnosis effect on m3 persisted when controlling for LME ($F(6,56) = 21.9$, $p < .001$, Wilks' $\Lambda = 0.299$) and for cognitive capacity ($F(6,54) = 3.68$, $p = .004$; Wilks' $\Lambda = 0.71$), computed as a composite score of all cognitive tests (supplementary results p.10-11). Higher $m3$ was associated with less correct responses ($r = -0.42$, $p < .001$), more switching between options overall ($r = 0.55$, $p < .001$) as well as after rewarded choices ($r = 0.45$, $p < .001$) and after unrewarded choices ($r = 0.32$, $p = .01$) (figure S3). Regarding the trial-by-trial computational trajectories, we examined how the median estimated volatility ($\mu_3^{(k)}$), absolute precision weighted outcome-related PE ($\varepsilon_2^{(k)}$), absolute precision-weighted volatility-related PE ($\varepsilon_3^{(k)}$) and belief uncertainty ($\sigma_2^{(k)}$) differed between groups (delusion and diagnosis) and task phases (stable-pre, volatile, stable-post). Across groups, all trajectories increased from the stable-pre to the volatile task phase. Pooled across phases, precision-weighted volatility $\mu_3^{(k)}$ was higher in PSZ ($F(1,62) = 8.54$, $p = .005$). Precision-weighted outcome-related prediction errors $\varepsilon_2^{(k)}$ showed an interaction of phase and delusion ($F(1.59,98.67) = 3.39$, $p = .048$). Post-hoc tests indicated that $\varepsilon_2^{(k)}$ increased from the stable-pre to the volatile phase only in D- ($t(124) = -4.71$, $p < .001$), but not D+ ($t(124) = -1.7$, $p = .515$). These findings suggests that participants with delusions did not finetune the belief updating signal according to changes of reward probabilities in the environment, which would allow more flexible belief updating (figure S2B).

fMRI Results

Across groups, we observed modulation of brain activity by μ_3, ε_2 and ε_3, but not by σ_2 (figure 2 and table S8). The volatility belief μ_3 covaried with BOLD-response in a network comprising fronto-parietal areas with middle frontal and angular gyrus; with anterior insula, superior frontal and supplementary motor cortex. HC showed stronger μ_3-related BOLD-response in superior frontal cortex ($t = 5.52$; x = 4 y = 30 z = 40, $p_{\text{FWE whole-brain corrected}} = 0.001$) and right angular gyrus ($t = 5.3$; x = 50 y = -48 z = 54, $p_{\text{FWE whole-brain corrected}} = 0.008$). Furthermore, HC displayed stronger activation in the right anterior insula and a cluster in the anterior superior frontal gyrus compared to PSZ, when correcting with the task main effect. We observed no μ_3-differences between D + and D−.

The precision-weighted outcome-related prediction error ε_2 covaried bilaterally with activity in a widely distributed network, including ventral striatum (nucleus accumbens), medial frontal and temporal gyrus, and middle and posterior cingulate (figure 2, table S8). D + compared to D− showed stronger representation of ε_2 in the striatum, including caudate and putamen, as well as anterior cingulate, inferior frontal gyrus and posterior insula. We observed no significant effect of diagnosis corrected at the whole-brain level nor corrected for the ε_2 main effect. Stronger ε_2-related activation was found in PSZ compared to HC in the right midbrain using small volume correction.

The prediction error of the volatility-belief ε_3 activated a bilateral network including the accumbens, precuneus, superior parietal and middle frontal gyrus (figure 2, table S8). Participants with delusional ideation showed stronger ε_3-related activity in the accumbens and anterior cingulate. We did not observe significant group differences between HC and PSZ on ε_3. Regarding group differences in BOLD-activity related to each regressor (table 3), we observed no interaction of diagnosis and delusion.

Discussion

We investigated how behavioral and neural alterations of belief updating relate to present delusions versus diagnosed schizophrenia to eventually extract delusion-specific inferential alterations. The present study has several strengths, including the orthogonal design of schizophrenia diagnosis and delusions, the preregistered replication- and multimethod-approach using computational modelling and neuroimaging techniques. Among participants with strong delusions, neural activation associated with precision-weighted prediction errors (PEs) (ε2, ε3) was heightened in fronto-striatal regions compared to participants with minimal delusions, irrespective of a clinical diagnosis. Diagnosed schizophrenia instead, was related to overestimation of environmental volatility and decreased neural representation of the volatility trajectories in fronto-parietal brain regions.

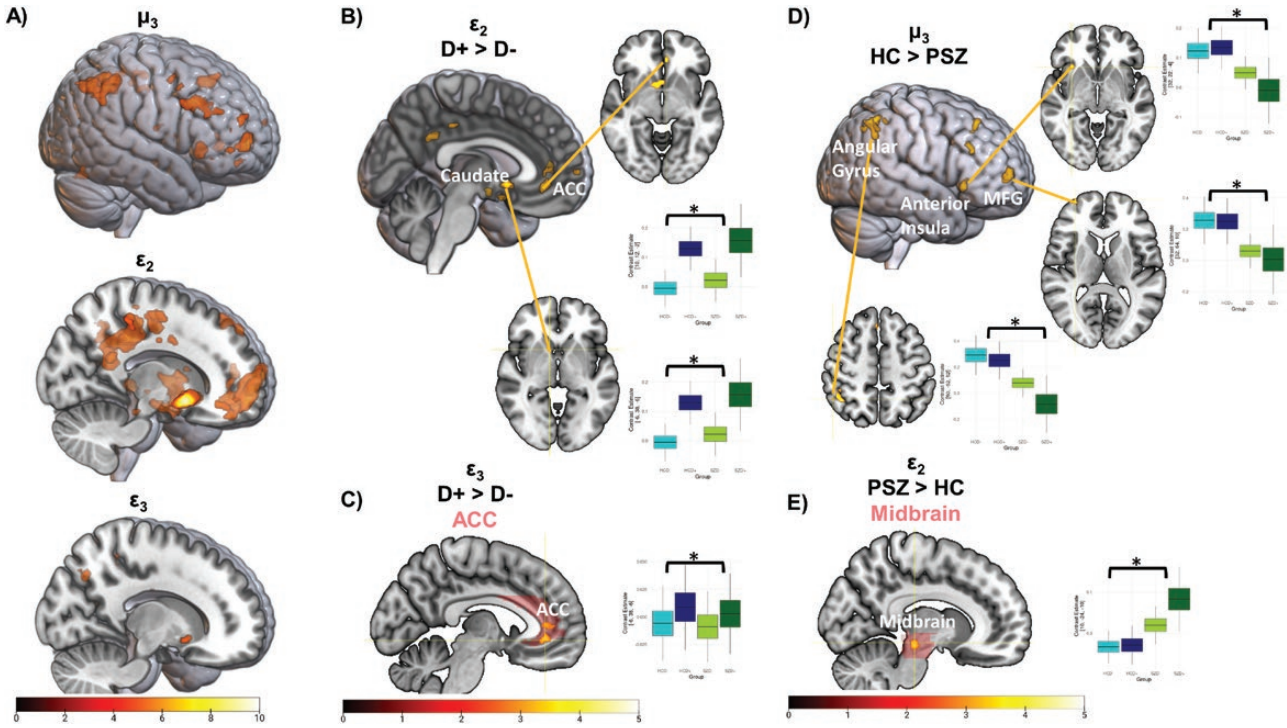

Fig. 2. Neural representation of the volatility estimate (*μ3)*, precision-weighted prediction error of the outcome (ε2) and volatility belief (ε3). A) Effects combined over all groups (main effects at pFWE < .05 whole brain corrected). B-C) Stronger representation of ε2 and ε3 in people with strong (D+) versus minimal delusions (D-), corrected for the task main effect (B) or an anatomical mask (red shading) of the anterior cingulate cortex (ACC) in (C). D-E) Stronger representation of μ3 and ε2 in patients with diagnosed schizophrenia versus healthy controls, corrected for the task main effect (D) an anatomical mask (red shading) of the midbrain (E). Plots show the contrast estimates at the peak voxel in the clusters pointed out by the lines.

Table 3. Group-differences of brain activation related to computational trajectories μ_3, ε_2 and ε_3 at $p < 0.05$ FWE corrected for the respective main effect masks (cluster extend reported at $p < 0.001$ uncorrected).

ANOVA (4 × 5 × 66) Feedback	Groups	Brain Region	Cluster size (k)	Side	P_{FWE} peak voxel	T- peak voxel	MNI peak-coordinate
μ_3							
	HC > PSZ	Angular Gyrus	139	R	.001	5.03	50, -52, 52
		Middle frontal gyrus	20	R	.002	4.76	48, 30, 28
		Superior frontal gyrus	33	R, L	.013	4.24	2, 28, 46
		Middle and superior frontal gyrus	65	R	.031	4.01	32, 64, 10
		Anterior Insula	29	R	.035	3.97	32, 22, -4
ε_2							
	D > noD	Caudate	75	R	.006	4.67	10, 12, -2
		Putamen	63	R	.021	4.36	26, 0, -12
		Putamen, Hippocampus	12	L	.041	4.18	-30, -12, -10
		Anterior cingulate	36	L, R	.017	4.41	-6, 38, -6
		Posterior Insula	1	R	.007	4.63	40, -4, 0
		Inferior frontal gyrus	16	L	.036	4.21	-52, 30, 12
$SVC_{midbrain}$ ε_3	PSZ > HC	Midbrain	35	R	.008	4.41	10, -24, -10
	D > noD	Accumbens	8	R	.003	3.88	8, 10, -8
SVC_{ACC}	D > noD	Anterior cingulate	88	L, R	.01	4.41	-6, 38, -6

μ_3 = volatility belief; ε2 = the precision-weighted prediction error of the outcome-belief, ε3 = precision-weighted prediction error of the volatility-belief; $SVC_{midbrain =}$ corrected for the anatomical midbrain mask; SVC_{ACC} = corrected for the anatomical anterior cingulate cortex mask; R = right hemisphere; L = left hemisphere.

Altered Prediction Error Signaling in Individuals with Delusions

In line with its role as reward and belief updating signal in the mesocortical dopamine system[41,44,45] and similar to prior fMRI studies,[46,47] precision-weighted outcome-related PEs covaried with neural activity predominantly in the fronto-striatal network including midbrain, caudate and medial frontal cortex, cingulate and insular regions. People with clinical delusions or nonclinical delusion-like ideation showed stronger activation in the striatum and cortical areas such as anterior cingulate, posterior insula and inferior frontal gyrus, suggesting that alterations in PE-signaling may thus be specifically related to delusions, irrespective of a schizophrenia diagnosis. This is in line with Cole et al.[14] who report heightened neural activity of precision-weighted outcome-related PEs in people at clinical high risk for schizophrenia, although group differences were located in anterior insula, frontal and parietal cortex. Also, neural activation related to higher-level PEs that update volatility estimates were more strongly represented in the accumbens and anterior cingulate in individuals with strong versus minimal delusions. In prior studies, the anterior cingulate was involved in value-based risk-predictions and error processing[48,49] and stronger activated in uncertain environments,[50] suggesting that our participants with delusions may have processed uncertainty differently.

This aligns with our finding that on the behavioral level, participants with delusions failed to increase precision-weighted PEs in the volatile task phase. Although we did not observe impaired choice-accuracy in the reversal-learning task, the profile of attenuated outcome-PEs in the volatile phase might be regarded as less favorable, since adaptive behavior in volatile environments requires a high precision of lower-level PEs to foster flexible belief updating.[51–55] This may resemble belief rigidity, which is characteristic to delusional thinking and seems to step alight particularly in volatile environments. Across human history, social crises, which are often accompanied by feelings of uncertainty or loss of control, have stimulated conspiracy theories.[56] This further corroborates the role of uncertainty in the formation of delusional thinking in the general population, without being necessarily related to manifest schizophrenia. We can rule out that these alterations are due to broader cognitive deficits, since groups with strong and minimal delusion-like belief did not differ in overall cognitive performance, which is supported by research suggesting that delusional severity is uncorrelated with cognitive performance.[7]

Increased Estimates of Environmental Volatility in Patients with Schizophrenia

In line with prior research,[25,31,57,58] PSZ responded with lower accuracy and more aberrant choice-switching. Our computational model suggests that in PSZ average estimations of environmental volatility converged to significantly higher levels, as reported previously in PSZ,[31] people at a clinical high risk[14] and HC with high paranoia.[8,27] In our sample, PSZ showed attenuated volatility-related neural activation in the anterior insula, angular and medial frontal gyri. Although the modeling results clearly converge with previous analysis in a partly overlapping sample,[31] we previously observed stronger volatility-related BOLD-response in the dlPFC of PSZ, potentially due to different sample characteristics. Interestingly, also Cole et al. observed stronger volatility coding on the neural level together with heightened volatility beliefs behaviorally.[14] How lower neural coding relates to heightened volatility beliefs remains to be established, but it cannot be easily explained by differences in model fit as model selection was similar for all groups and participants with poor model fit were excluded. We found no effect of "delusion" or "delusion by diagnosis" interaction on volatility estimates, suggesting that associated behavioral and neural alterations may be a feature associated with schizophrenia and not specific to present delusional beliefs. This may hint toward a lack of specificity in the relationship between positive symptom severity and altered volatility estimation. Although multiple theoretical accounts propose altered volatility estimation as mechanism of psychotic symptoms in PSZ,[5,12] empirical evidence for a direct association between increased volatility estimation and psychotic symptom severity is sparse. Instead, group comparisons may be confounded, just like our study, by lower cognitive capacity which may at least partially account for observed group differences.

Besides altered volatility estimation on the behavioral and neural level, PSZ showed increased midbrain-activation in response to precision-weighted outcome-related PEs compared to HC. This is in contrast to several previous studies investigating unweighted reward PEs, which found reduced PE-related activation of patients with schizophrenia compared to controls in midbrain[59] but also other areas including striatum[25] or prefrontal regions.[60–62] The heterogeneity of findings may be due to different task paradigms, different severity levels and diagnoses within the psychosis spectrum.[25,63,64] Furthermore, medication status and substance use contribute to dopaminergic activity and likely influence PE-signaling.[65,66]

Taken together, our results partly mirror studies investigating perceptual belief updating. Whereas Schmack et al.[67] observed less perceptual belief-stability among delusional subjects, Powers and colleagues' report increased perceptual priors at lower levels of the processing hierarchy (i.e., more stability) in people with acoustic hallucinations and reduced adaptation toward increasing environmental volatility in people with psychotic illness. While we did not investigate perceptual alterations in our study, our computational model also suggested attenuated belief updating (lower-level PEs) in the volatile task phase in people with delusions and in PSZ altered higher-level environmental

volatility estimates. Since hallucinations and delusions often occur together, a "similar failure mode" at different levels (perceptual level for hallucinations and abstract level for delusions) would be plausible and could involve overweighing of priors at different levels in the processing hierarchy.[30] Yet, more research is needed to elucidate how different disease stages and manifestation of delusions and/or hallucinations may impact misjudgment of environmental volatility; and whether this results from or in altered precision-weighted prediction errors.

Limitations

Several limitations must be addressed. First, although we tried to adjust our analysis accordingly, we were unable to match HC and PSZ for age and cognitive capacity and measures of task performance correlated with cognitive capacity. Therefore, we cannot rule out that differences between these groups may be biased by these confounds, although introducing cognitive capacity as covariate did not alter the results. Second, due to our cross-sectional design, we lack follow-up information if delusional-like ideations in HC progressed towards pathology and if so, whether this preceded or followed an overestimation of environmental volatility. Another limitation of the present study is that we had to exclude subjects with bad model fit, which resulted in exclusion of patients with significantly worse task performance (supplement table S4). Moreover, our computational model may not optimally track when volatility estimates decrease in the second stable phase.[68] Although the behavioral effects could be recovered in synthetic choice data that was simulated with the estimated parameters from the HGF3-rev model (supplementary results), other models allowing for a more flexible adjustment to environmental change-points should be utilized in future studies and compared with the current approach.[55,69] Traditional fit indices (AIC and BIC) favored more parsimonious models (supplementary materials). While the HGF3-rev model provided best fit according to LME, the discrepancy of AIC, BIC and LME may raise doubt about the most parsimonious description of the behavioral data. Yet, we decided to stick to our model selection metrics, since we wanted to replicate the analysis of Cole et al.[14] as pre-registered, in order to allow comparability of results between our and their clinical high risk sample. Another limitation is our small sample size within sub-groups potentially preventing the detection of significant diagnosis by delusion interaction effects. As explained in the Methods section, the clinical sample was split according to expert ratings of the PANSS delusion item, instead of self-reported PDI. This may represent a limitation, as subgroups of schizophrenia patients with minimal versus strong delusions according to expert rating did not significantly differ on self-rating scores. Last, we would like to note that we did not take into account form and content of delusions-like beliefs by use of the PDI as criterion in the healthy subgroup, which may be critical for assessing delusions in the context of schizophrenia.[70]

Conclusion

Altogether, our results suggest that people with delusions show altered prediction error signals in various brain areas, whereas a diagnosis of schizophrenia was associated with altered volatility estimates. Possibly, a failure to integrate more complex information, like environmental volatility may be characteristic to pathological schizophrenia or accompany cognitive deficits. This study supports the great potential computational psychiatry holds to investigate differential mechanisms underlying psychotic symptoms and may eventually pave the way towards precision medicine. Particularly, in heterogeneous psychiatric syndromes, like psychosis spectrum disorders, greater understanding of differential mechanisms that delineate specific patient-subgroups may enable more efficient outcome prediction and treatment.

Acknowledgments

All authors report no potential conflicts of interest, or biomedical financial interest. We gratefully thank all participants for their contribution. The full analysis code will be made available upon request.

Funding

This work was supported by research grants of the German Research Foundation awarded to Florian Schlagenhauf (SCHL 1969/1-1/1-2/3-1), by the Heisenberg program from the German Research Foundation (SCHL 1969/5-1). Sophie Fromm was supported by the Einstein Center for Neurosciences Berlin and by the Bernstein Center for Computational Neurosciences Berlin. Jakob Kaminski was supported by the Charité Clinician-Scientist Program of the Berlin Institute of Health.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*.; 2013. doi:10.1176/appi.books.9780890425596.744053
2. Hanssen M, Bak M, Bijl R, Vollebergh W, van Os J. The incidence and outcome of subclinical psychotic experiences in the general population. *Br J Clin Psychol.* 2005;44(2):181–191. doi:10.1348/014466505X29611.
3. Rössler W, Riecher-Rössler A, Angst J, *et al.* Psychotic experiences in the general population: A twenty-year prospective community study. *Schizophr Res.* 2007;92(1-3):1–14. doi:10.1016/J.SCHRES.2007.01.002.
4. McGrath JJ, Saha S, Al-Hamzawi A, *et al.* Psychotic experiences in the general population: a cross-national analysis based on 31 261 respondents from 18 countries. *JAMA Psychiatry* 2015;72(7):697–705. doi:10.1001/jamapsychiatry.2015.0575.
5. Adams RA, Napier G, Roiser JP, Mathys C, Gilleen J. Attractor-like dynamics in belief updating in schizophrenia.

J Neurosci. 2018;38(44):9471–9485. doi:10.1523/JNEUROSCI.3163-17.2018.

6. Heinz A, Murray GK, Schlagenhauf F, Sterzer P, Grace AA, Waltz JA. Towards a unifying cognitive, neurophysiological, and computational neuroscience account of schizophrenia. *Schizophr Bull.* 2019;45(5):1092–1100. doi:10.1093/schbul/sby154.
7. Baker SC, Konova AB, Daw ND, Horga G. A distinct inferential mechanism for delusions in schizophrenia. *Brain* 2019;142(6):1797–1812. doi:10.1093/brain/awz051.
8. Suthaharan P, Reed E, Leptourgos P, *et al.* Paranoia and belief updating during a crisis. *Res Sq* Published online July 27, 2021;5(9):1–13. doi:10.21203/rs.3.rs-145987/v1.
9. Adams RA, Huys QJM, Roiser JP. Computational Psychiatry: Towards a mathematically informed understanding of mental illness. *J Neurol Neurosurg Psychiatry.* 2016;87(1):53–63. doi:10.1136/jnnp-2015-310737.
10. Huys QJM, Maia T V., Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci*. 2016;19(3):404–413. doi:10.1038/nn.4238
11. 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.
12. Sterzer P, Voss M, Schlagenhauf F, Heinz A. Decision-making in schizophrenia: A predictive-coding perspective. *Neuroimage* 2019;190:133–143. doi:10.1016/j.neuroimage.2018.05.074.
13. Mathys C, Daunizeau J, Friston KJ, Stephan KE. A Bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci.* 2011;5(MAY):9. doi:10.3389/fnhum.2011.00039.
14. Cole DM, Diaconescu AO, Pfeiffer UJ, *et al.* Atypical processing of uncertainty in individuals at risk for psychosis. *NeuroImage Clin*. 2020;26(December 2019):102239. doi:10.1016/j.nicl.2020.102239
15. Diederen KMJ, Fletcher PC. Dopamine, Prediction Error and Beyond. *Neuroscientist.* 2021;27(1):30–46. doi:10.1177/1073858420907591.
16. Fusar-Poli P, Meyer-Lindenberg A. Striatal presynaptic dopamine in schizophrenia, part II: Meta-analysis of [18F/11C]-DOPA PET studies. *Schizophr Bull.* 2013;39(1):33–42. doi:10.1093/schbul/sbr180.
17. Brugger SP, Angelescu I, Abi-Dargham A, Mizrahi R, Shahrezaei V, Howes OD. Heterogeneity of striatal dopamine function in schizophrenia: meta-analysis of variance. *Biol Psychiatry.* 2020;87(3):215–224. doi:10.1016/j.biopsych.2019.07.008.
18. Howes OD, Hird EJ, Adams RA, Corlett PR, McGuire P. Aberrant salience, information processing, and dopaminergic signaling in people at clinical high risk for psychosis. *Biol Psychiatry.* 2020;88(4):304–314. doi:10.1016/j.biopsych.2020.03.012.
19. 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.
20. Howes OD, Kapur S. The Dopamine Hypothesis of Schizophrenia: Version III--The Final Common Pathway. *Schizophr Bull.* 2009;35(3):549–562. doi:10.1093/schbul/sbp006.
21. Winton-Brown TT, Fusar-Poli P, Ungless MA, Howes OD. Dopaminergic basis of salience dysregulation in psychosis. *Trends Neurosci.* 2014;37(2):85–94. doi:10.1016/j.tins.2013.11.003.
22. Stuke H, Stuke H, Weilnhammer VA, Schmack K. Psychotic experiences and overhasty inferences are related to maladaptive learning. *PLoS Comput Biol.* 2017;13(1):e10053931–e10053920. doi:10.1371/journal.pcbi.1005328.
23. Katthagen T, Mathys C, Deserno L, *et al.* Modeling subjective relevance in schizophrenia and its relation to aberrant salience. Daunizeau J, ed. *PLoS Comput Biol.* 2018;14(8):e1006319. doi:10.1371/journal.pcbi.1006319.
24. Livet A, Navarri X, Potvin S, Conrod P. Cognitive biases in individuals with psychotic-like experiences: A systematic review and a meta-analysis. *Schizophr Res.* 2020;222:10–22. doi:10.1016/j.schres.2020.06.016.
25. Schlagenhauf F, Huys QJM, Deserno L, *et al.* Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage* 2014;89:171–180. doi:10.1016/j.neuroimage.2013.11.034.
26. Kaplan CM, Saha D, Molina JL, *et al.* Estimating changing contexts in schizophrenia. *Brain* 2016;139(7):2082–2095. doi:10.1093/brain/aww095.
27. Reed EJ, Uddenberg S, Suthaharan P, *et al.* Paranoia as a deficit in non-social belief updating. *Elife* 2020;9:1–55. doi:10.7554/eLife.56345.
28. Haarsma J, Fletcher PC, Griffin JD, *et al.* Precision weighting of cortical unsigned prediction error signals benefits learning, is mediated by dopamine, and is impaired in psychosis. *Mol Psychiatry.* 2021;26(9):5320–5333. doi:10.1038/s41380-020-0803-8.
29. Powers AR, Mathys C, Corlett PR. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science (80-)* 2017;357(6351):596–600. doi:10.1126/science.aan3458.
30. Ashinoff BK, Singletary NM, Baker SC, Horga G. Rethinking delusions: A selective review of delusion research through a computational lens. *Schizophr Res.* 2021;3:23–41. doi:10.1016/j.schres.2021.01.023.
31. Deserno L, Boehme R, Mathys C, *et al.* Volatility estimates increase choice switching and relate to prefrontal activity in schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2020;5(2):173–183. doi:10.1016/j.bpsc.2019.10.007.
32. Boehme R, Deserno L, Gleich T, *et al.* Aberrant salience is related to reduced reinforcement learning signals and elevated dopamine synthesis capacity in healthy adults. *J Neurosci.* 2015;35(28):10103–10111. doi:10.1523/JNEUROSCI.0805-15.2015.
33. Katthagen T, Kaminski J, Heinz A, Buchert R, Schlagenhauf F. Striatal dopamine and reward prediction error signaling in unmedicated schizophrenia patients. *Schizophr Bull.* 2020;46(6):1535–1546. doi:10.1093/schbul/sbaa055.
34. Mathys CD, Lomakina EI, Daunizeau J, *et al.* Uncertainty in perception and the Hierarchical Gaussian filter. *Front Hum Neurosci*. 2014;8(November):825. doi:10.3389/fnhum.2014.00825
35. Peters ER, Joseph SA, Garety PA. Measurement of delusional ideation in the normal population: Introducing the PDI (Peters *et al*. Delusions Inventory). *Schizophr Bull.* 1999;25(3):553–576. doi:10.1093/oxfordjournals.schbul.a033401.
36. Peters E, Joseph S, Day S, Garety P. Measuring delusional ideation: the 21-item Peters *et al*. Delusions Inventory (PDI). *Schizophr Bull.* 2004;30(4):1005–1022. doi:10.1093/OXFORDJOURNALS.SCHBUL.A007116.
37. Fukuda Y, Katthagen T, Deserno L, *et al.* Reduced parietofrontal effective connectivity during a working-memory task in people with high delusional ideation. *J Psychiatry Neurosci.* 2019;44(3):195–204. doi:10.1503/jpn.180043.
38. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13(2):261–276. doi:10.1093/schbul/13.2.261.

39. Wallwork RS, Fortgang R, Hashimoto R, Weinberger DR, Dickinson D. Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale for schizophrenia. *Schizophr Res.* 2012;137(1-3):246–250. doi:10.1016/j.schres.2012.01.031.
40. Galderisi S, Mucci A, Dollfus S, *et al.* EPA guidance on assessment of negative symptoms in schizophrenia. *Eur Psychiatry.* 2021;64(1):e1–e23. doi:10.1192/j.eurpsy.2021.11.
41. Iglesias S, Mathys C, Brodersen KH, *et al.* Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 2013;80(2):519–530. doi:10.1016/j.neuron.2013.09.009.
42. The FIL Methods Group. SPM12 Manual. Published online 2020. doi:10.1111/j.1365-294X.2006.02813.x
43. Maldjian J, Laurienti P, Kraft R, Neuroimage JB-, 2003. *undefined. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets*. Elsevier. https://www.sciencedirect.com/science/article/pii/S1053811903001691. Accessed March 28, 2022.
44. Diederen KMJ, Ziauddeen H, Vestergaard MD, Spencer T, Schultz W, Fletcher PC. Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *J Neurosci.* 2017;37(7):1708–1720. doi:10.1523/JNEUROSCI.1979-16.2016.
45. Nour MM, Dahoun T, Schwartenbeck P, *et al.* Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. *Proc Natl Acad Sci USA.* 2018;115(43):E10167–E10176. doi:10.1073/pnas.1809298115.
46. Diaconescu AO, Mathys C, Weber LAE, Kasper L, Mauer J, Stephan KE. Hierarchical prediction errors in midbrain and septum during social learning. *Soc Cogn Affect Neurosci* 2017;12(4):618–634. doi:10.1093/SCAN/NSW171.
47. Corlett PR, Mollick JA, Kober H. Meta-analysis of human prediction error for incentives, perception, cognition, and action. *Neuropsychopharmacol* 2022. Published online January 11, 2022:1–11. doi:10.1038/s41386-021-01264-3
48. d'Acremont M, Lu ZL, Li X, Van der Linden M, Bechara A. Neural correlates of risk prediction error during reinforcement learning in humans. *Neuroimage* 2009;47(4):1929–1939. doi:10.1016/J.NEUROIMAGE.2009.04.096.
49. Vassena E, Deraeve J, Alexander WH. Surprise, value and control in anterior cingulate cortex during speeded decision-making. *Nat Hum Behav.* 2020;4(4):412–422. doi:10.1038/s41562-019-0801-5.
50. Silvetti M, Seurinck R, Verguts T. Value and prediction error estimation account for volatility effects in ACC: A model-based fMRI study. *Cortex.* 2013;49(6):1627–1635. doi:10.1016/J.CORTEX.2012.05.008.
51. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. *Nat Neurosci.* 2007;10(9):1214–1221. doi:10.1038/nn1954.
52. Wilson RC, Nassar MR, Gold JI. A mixture of delta-rules approximation to bayesian inference in change-point problems. *PLoS Comput Biol.* 2013;9(7):1003150. doi:10.1371/journal.pcbi.1003150.
53. Moens V, Zénon A. Learning and forgetting using reinforced Bayesian change detection. *PLoS Comput Biol.* 2019;15(4):e1006713. doi:10.1371/journal.pcbi.1006713.
54. Nassar MR, Troiani V. The stability flexibility tradeoff and the dark side of detail. *Cogn Affect Behav Neurosci.* 2020;21(November):2020.01.03.894014. doi:10.3758/s13415-020-00848-8
55. Piray P, Daw ND. A simple model for learning in volatile environments. Soltani A, ed. *PLoS Comput Biol.* 2020;16(7):e1007963. doi:10.1101/701466.
56. van Prooijen JW, Douglas KM. Conspiracy theories as part of history: The role of societal crisis situations. *Mem Stud* 2017;10(3):323–333. doi:10.1177/1750698017701615.
57. Waltz JA, Kasanova Z, Ross TJ, Salmeron BJ, Mcmahon RP. The Roles of Reward, Default, and Executive Control Networks in Set-Shifting Impairments in Schizophrenia. *PLoS One.* 2013;8(2):57257. doi:10.1371/journal.pone.0057257.
58. Culbreth AJ, Westbrook A, Daw ND, Botvinick M, Barch DM. Reduced model-based decision-making in schizophrenia. *J Abnorm Psychol.* 2016;125(6):777–787. doi:10.1037/abn0000164.
59. Murray GK, Corlett PR, Clark L, *et al.* Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry.* 2008;13(3):267–276. doi:10.1038/SJ.MP.4002058.
60. Corlett PR, Murray GK, Honey GD, *et al.* Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain* 2007;130(9):2387–2400. doi:10.1093/brain/awm173.
61. Schlagenhauf F, Sterzer P, Schmack K, *et al.* Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. *Biol Psychiatry.* 2009;65(12):1032–1039. doi:10.1016/J.BIOPSYCH.2008.12.016.
62. Waltz JA, Schweitzer JB, Ross TJ, *et al.* Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology.* 2010;35(12):2427–2439. doi:10.1038/NPP.2010.126.
63. Reinen JM, Van Snellenberg JX, Horga G, Abi-Dargham A, Daw ND, Shohamy D. Motivational context modulates prediction error response in schizophrenia. *Schizophr Bull.* 2016;42(6):1467–1475. doi:10.1093/schbul/sbw045.
64. Ermakova AO, Knolle F, Justicia A, *et al.* Abnormal reward prediction-error signalling in antipsychotic naive individuals with first-episode psychosis or clinical risk for psychosis. *Neuropsychopharmacology.* 2018;43(8):1691–1699. doi:10.1038/s41386-018-0056-2.
65. Farde L, Wiesel F-A, Halldin C, Sedvall G. Central D2-dopamine receptor occupancy in schizophrenic patients treated with antipsychotic drugs. *Arch Gen Psychiatry.* 1988;45(1):71. doi:10.1001/archpsyc.1988.01800250087012.
66. Heinz A, Knable MB, Coppola R, *et al.* Psychomotor slowing, negative symptoms and dopamine receptor availability--an IBZM SPECT study in neuroleptic-treated and drug-free schizophrenic patients. *Schizophr Res.* 1998;31(1):19–26. doi:10.1016/S0920-9964(98)00003-6.
67. Schmack K, de Castro AGC, Rothkirch M, *et al.* Delusions and the role of beliefs in perceptual inference. *J Neurosci.* 2013;33(34):13701–13712. doi:10.1523/JNEUROSCI.1778-13.2013.
68. Fletcher PC. Predictive coding and hallucinations: a question of balance: Comment on Powers, Mathys, and Corlett (2017) "Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors.". *Cogn Neuropsychiatry.* 2017;22(6):453–460. doi:10.1080/13546805.2017.1391083.
69. Nassar MR, Waltz JA, Albrecht MA, Gold JM, Frank MJ. All or nothing belief updating in patients with schizophrenia reduces precision and flexibility of beliefs. *Brain* 2021;144(3):1–17. doi:10.1093/brain/awaa453.
70. Heinz A, Voss M, Lawrie SM, *et al.* Shall we really say goodbye to first rank symptoms? *Eur Psychiatry.* 2016;37:8–13. doi:10.1016/j.eurpsy.2016.04.010.