

All or nothing belief updating in patients with schizophrenia reduces precision and flexibility of beliefs

Matthew R. Nassar,^{1,2} James A. Waltz,³ Matthew A. Albrecht,⁴ James M. Gold³ and Michael J. Frank^{1,5}

Schizophrenia is characterized by abnormal perceptions and beliefs, but the computational mechanisms through which these abnormalities emerge remain unclear. One prominent hypothesis asserts that such abnormalities result from overly precise representations of prior knowledge, which in turn lead beliefs to become insensitive to feedback. In contrast, another prominent hypothesis asserts that such abnormalities result from a tendency to interpret prediction errors as indicating meaningful change, leading to the assignment of aberrant salience to noisy or misleading information. Here we examine behaviour of patients and control subjects in a behavioural paradigm capable of adjudicating between these competing hypotheses and characterizing belief updates directly on individual trials. We show that patients are more prone to completely ignoring new information and perseverating on previous responses, but when they do update, tend to do so completely. This updating strategy limits the integration of information over time, reducing both the flexibility and precision of beliefs and provides a potential explanation for how patients could simultaneously show over-sensitivity and under-sensitivity to feedback in different paradigms.

- 1 Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence RI 02912-1821, USA
- 2 Department of Neuroscience, Brown University, Providence RI 02912-1821, USA
- 3 Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, MD, USA
- 4 School of Public Health, Curtin University, Perth, Western Australia, Australia
- 5 Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence RI 02912-1821, USA

Correspondence to: Matthew Nassar
164 Angell Street
Providence RI, 02906, USA
E-mail: matthew_nassar@brown.edu

Keywords: schizophrenia; belief updating; learning; computational psychiatry

Introduction

Schizophrenia is a mental illness characterized by diverse symptomatology. Many patients with schizophrenia experience positive symptoms, such as delusions or hallucinations, as well as negative and cognitive ones. Delusions common to schizophrenia often involve persisting false beliefs. However, the mechanism through which these false beliefs

arise is unclear and the computational and biological factors giving rise to them are actively debated.

One influential theory speculates that delusions emerge as a result of aberrant salience assigned to information through dopaminergic overactivation (Kapur, 2003; Howes and Kapur, 2009). Recently, this theory has been interpreted in a computational framework originally applied to learning in dynamic environments in which salient information is

regarded as suggesting a likely environmental change that requires rapid updating of beliefs (Nassar *et al.*, 2010). Within this framework, aberrant salience can be thought of as resulting from an abnormally high estimate of the rate at which the environment is changing, or hazard rate, thereby leading to more rapid belief updating in changing environments (Stephan *et al.*, 2016). Consistent with this interpretation of aberrant salience theory, patients with schizophrenia updated beliefs more rapidly than control subjects, were best described by models that overestimated the hazard rate of the environment, and displayed task-related connectivity between prefrontal and midbrain regions that predicted the severity of delusions (Kaplan *et al.*, 2016). Nonetheless, previous work leaves open a number of questions regarding the exact computations that are altered in schizophrenia.

One such question is how this high hazard rate characterization, which implies increased sensitivity to prediction errors under many conditions (Nassar *et al.*, 2010), can be reconciled with the slower rates of learning and perseverative responding characteristic of patient behaviour in many other tasks (Goldberg *et al.*, 1987; Laws, 1999; Leeson *et al.*, 2009; Reddy *et al.*, 2016; Baker *et al.*, 2019). Indeed, computational formulations of the latter observations suggest that they arise because of overly precise prior representations that reduce the speed with which beliefs are updated according to unpredicted events (Corlett *et al.*, 2019; Horga and Abi-Dargham, 2019). The coexistence of over-updating and under-updating phenotypes is perplexing from a computational standpoint, as these extreme behaviours are thought to occupy opposite ends of a spectrum of belief updating policies that ranges from those emphasizing stability (slow learning) versus those emphasizing flexibility (rapid learning) (Behrens *et al.*, 2007; Nassar *et al.*, 2010).

A possible resolution to this apparent discrepancy is that patient learning is not stationary but is instead sensitive to statistical context. In this view, perseveration and promiscuous updating may be both observed, depending on the statistical context and whether, normatively, extreme events should drive more or less learning within it. Here we examined how patients with schizophrenia update beliefs in different statistical environments in order to better characterize the computations affected by the illness. We utilized a task that probed beliefs directly on each trial (Nassar *et al.*, 2010) and could measure both over- and under-learning in separate statistical contexts that favour either more or less learning from surprising events (Nassar *et al.*, 2019). Specifically, our framework dissociated salience from learning by including surprising events that should either be used for updating or ignored (d'Acromont and Bossaerts, 2016; Nassar *et al.*, 2019), allowing us to define the computational differences between patients and controls.

We found that while average rates of learning did not differ systematically between schizophrenia patients and controls, patients showed a pronounced reduction in a specific category of moderate belief updates. Patients relied instead on a combination of total belief updates (as might be

predicted by the aberrant salience account) and non-updates (as might be predicted by an overly strong prior). This led patients to form beliefs that were both less flexible after change-points and less precise during periods of stability than those formed by control subjects. Patient behaviour could be described by an extension of the normative updating model in which belief updates are omitted as a probabilistic function of their expected magnitude. Parameter estimates from this model could predict patient status on an individual basis. Together, these results provide a unified account of the seemingly contradictory observations that schizophrenia patients over-interpret noisy information, but also underuse feedback for learning. In particular, our results suggest that both of these behaviours emerge from a single deficiency in the sort of moderate belief updates that facilitate integration of information across multiple observations.

Materials and methods

Participants

To determine the effect of psychotic illness on directed and random exploration, 108 subjects with a diagnosis of schizophrenia or schizoaffective disorder (referred to, collectively, as PSZ) and 33 healthy age-matched community volunteers performed our behavioural task at the Maryland Psychiatric Research Center (MPRC), University of Maryland School of Medicine. All participants gave informed consent, and the research was approved by the Institutional Review Board at the University of Maryland School of Medicine.

Clinical and cognitive measures

Patients were clinically and pharmacologically stable (no change in drug or dose for at least 4 weeks) outpatients from the MPRC or other nearby clinics. Almost all PSZ patients were being treated with antipsychotic medications (Supplementary Table 1). The presence of a schizophrenia spectrum disorder in patients, as well as the absence of a current Axis I disorder (including drug dependence) and lifetime diagnosis of a psychotic disorder in healthy volunteers, was verified by screening with the Structured Clinical Interview for DSM-IV (First *et al.*, 1997). The absence of a neurological disorder, cognitively impairing medical disorder (e.g. chronic, untreated hypertension or diabetes), and psychosis in first-degree relatives was verified by self-report. PSZ patients were further assessed with the Scale for the Assessment of Negative Symptoms (SANS) (Andreasen, 1984), and the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 1962).

Patients with schizophrenia and healthy volunteers were tested using a cognitive battery including the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), the Wechsler Test of Adult Reading (The Psychological Corporation, 2001) and the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (Green *et al.*, 2004). There were significant differences between patients and community control subjects on all measures of cognition (Table 1).

Table 1 Demographic, cognitive, and clinical measures in samples of patients and controls included in final analyses

Measure	Patients (n = 94) Mean (SD)	Controls (n = 31) Mean (SD)	Inferential statistic
Demographic			
Age	37.1 (10.0)	37.2 (10.3)	$t = -0.05$
Gender	33 F, 61 M	11 F, 20 M	$\chi^2 = 0.001$
Race	45 C, 38 AA, 4 AS, 6 M/O	17 C, 12 AA, 0 AS, 2 M/O	$\chi^2 = 1.554$
Subject education	13.2 (2.2)	15.0 (2.1)	$t = -3.94^{***}$
Parental education	14.3 (3.0)	14.0 (2.5)	$t = 0.52$
Cognitive			
WASI Estimated IQ (four subtests)	94.4 (14.4)	111.4 (14.3)	$t = -5.71^{***}$
WRAT-Reading Scaled Score	96.7 (15.0)	109.4 (15.4)	$t = -4.05^{***}$
WTAR Scaled Score	99.0 (18.1)	110.7 (14.4)	$t = -3.31^{**}$
MATRICES Composite Score	33.2 (12.7)	51.4 (11.3)	$t = -7.08^{***}$
MATRICES Domain Scores			
Processing Speed	38.4 (12.8)	53.1 (12.2)	$t = -5.60^{***}$
Attention/Vigilance	40.1 (11.3)	52.9 (11.4)	$t = -5.46^{***}$
Working Memory	39.9 (10.8)	52.5 (11.5)	$t = -5.52^{***}$
Verbal Learning	37.8 (8.1)	50.7 (9.0)	$t = -7.52^{***}$
Visuospatial Learning	36.4 (12.0)	45.0 (10.8)	$t = -3.52^{**}$
Reasoning/Problem Solving	43.6 (10.3)	49.6 (10.0)	$t = -2.80^{**}$
Social Cognition	41.9 (12.2)	54.2 (8.0)	$t = -5.24^{***}$
Clinical			
BPRS Mean Item Score—All Items	1.7 (0.4)		
BPRS Mean Item Score—Psychosis	2.2 (1.2)		
BPRS Mean Item Score—Depression	1.9 (0.9)		
SANS Mean Item Score—All Items	1.5 (0.6)		
SANS Mean Item Score—Avolition/Anhedonia	2.0 (0.8)		

AA = African-American; AS = Asian; BPRS = Brief Psychiatric Rating Scale; C = Caucasian; F = female; M = male; M/O = mixed/other; MATRICES = Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery; WASI = Wechsler Abbreviated Scale of Intelligence; WRAT-Reading = Wide-Ranging Achievement Test, Reading Subtest; WTAR = Wechsler Test of Adult Reading.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Predictive Inference Task

All participants performed a modified predictive inference task programmed in MATLAB (The Mathworks, Natick, MA), using the MGL (<https://github.com/justingardner/mgl>), SnowDots and Tower-Of-Psych toolboxes (<https://github.com/TheGoldLab/Lab-Matlab-Control/>). The task was an extension of previous predictive inference tasks (Nassar *et al.*, 2010, 2012) in which participants were asked predict the next in a series of outcomes, but cast in terms of catching bags dropped from a helicopter (McGuire *et al.*, 2014). Ours incorporated an additional condition in which surprising outcomes were not indicative of an underlying change in the generative process (d'Acremont and Bossaerts, 2016; Nassar *et al.*, 2019), and in fact uninformative about the underlying mean of the distribution. This condition allowed us to measure the degree to which participants appropriately attributed surprising information according to its most likely source (Nassar *et al.*, 2019).

Participants were instructed to place a bucket at a horizontal position on the ground in order to maximize the chances of catching bags that would be dropped from a helicopter (Fig. 1A). On each trial, a bag was dropped from the helicopter at the top of the screen with a probabilistic horizontal displacement described in detail below. During training, the helicopter was visible to participants and they were capable of observing the variability from trial to trial in the distribution of bags around it. Bucket placement was accomplished using a gamepad that contained right and left buttons that controlled movement

of the bucket from its previous position and a separate button to confirm bucket placement. Both patients and controls tended to place buckets under the helicopter in this training task, rather than placing it at the most recent bag location, suggesting that they had a basic understanding of the generative process (Supplementary Fig. 1). During the subsequent test phase, the helicopter was obscured by clouds, forcing participants to infer its position based on previously observed bag locations. To minimize the need for working memory necessary to represent the previous bag location the task provided a visual depiction of the prediction error from the previous trial, spanning the range from the middle of the previous bucket position to the most recent bag location (small red line in Fig. 1A). Here we report only data from the test phase as the training data simply required participants to move a bucket underneath a visible helicopter.

A key manipulation in our design was how the helicopter location evolved from one trial to the next. The task involved two distinct statistical contexts that were capable of disambiguating surprise from learning and the methods and mathematical justification for this approach have previously been reported (Nassar *et al.*, 2019). In all conditions, helicopter, bag, and bucket locations were generated and recorded on a scale ranging from 0 (left side of screen) to 300 (right side of screen). Depending on the condition, the helicopter would either (i) remain stationary on the majority of trials and re-aim to a random location on the 0–300 interval, with a hazard rate of 0.125

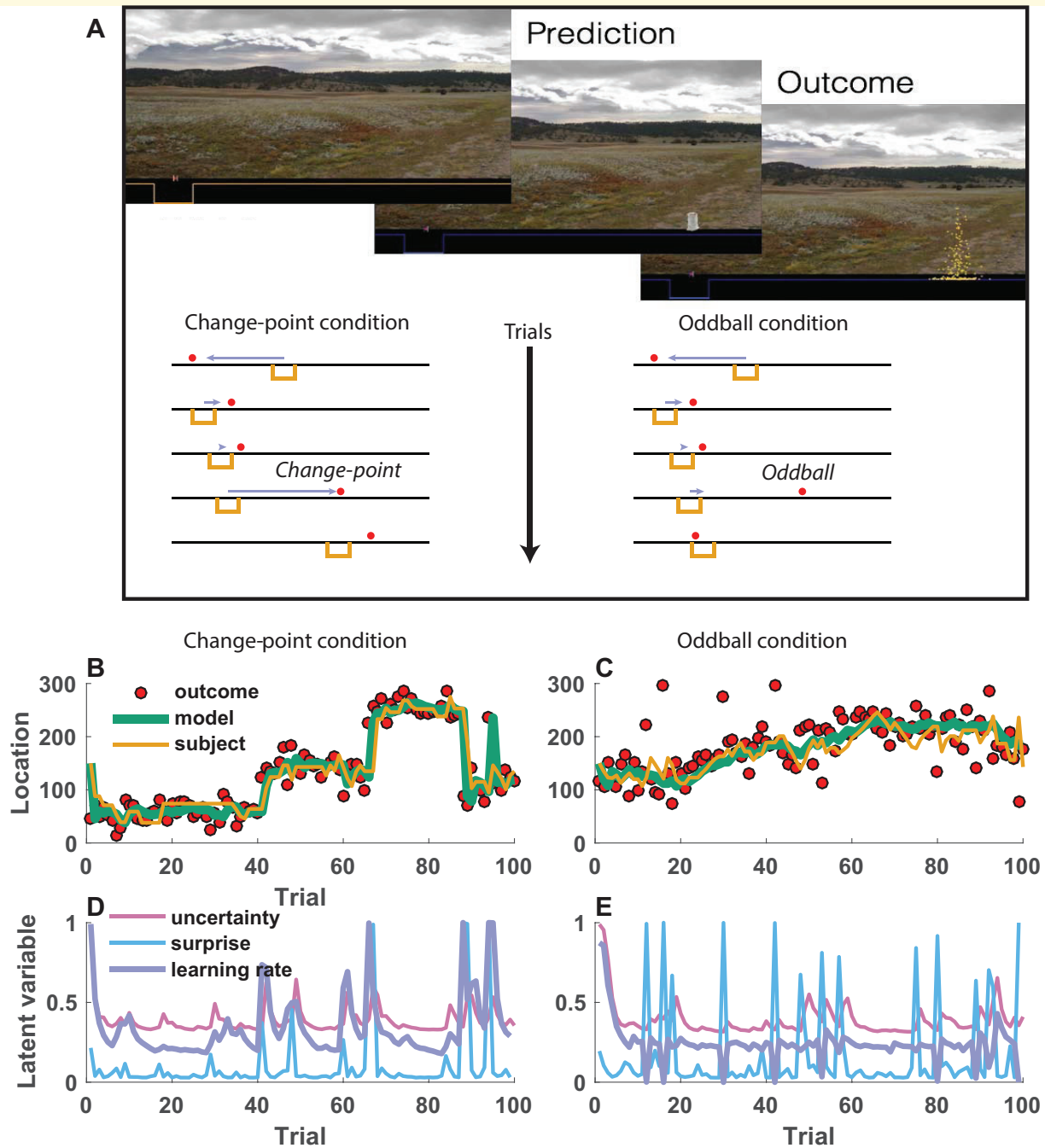


Figure 1 A modified predictive inference task to measure behavioural adjustment in response to surprising outcomes. **(A)** In each trial, participants moved a bucket to the location at which they expected a helicopter (obscured by clouds) to drop a bag of potentially rewarding contents (*top left*). After the participant indicated satisfaction with the bucket placement, a bag fell from the top of the screen, which provided new information about the true location of the helicopter (in the bag location; *top middle*) and about the reward attained on that trial (in the amount of the bag contents that landed in the bucket; *top right*). All participants completed the task under two different generative conditions that were explicitly instructed. In the change-point condition (*bottom left*) the helicopter position occasionally underwent change-points, leading to a persistent change in the location of bags (red circles) across trials (vertical axis). In the oddball condition (*bottom right*) the bag location was occasionally unrelated to the actual helicopter location, giving rise to oddball events that were unrelated to bag positions preceding or following them. **(B and C)** Bucket placements made by an example subject (yellow) and the normative model (green) for each trial (abscissa) of a task block in which bag locations (red points) were generated using either a change-point **(B)** or oddball **(C)** structure. Note that model and example subject behaviour includes rapid behavioural adjustment in response to large errors in predicted bag location for the change-point condition **(B)**, but not in the oddball condition **(C)**. **(D and E)** Behaviour of the normative model is described by an error-driven learning rule in which the learning rate (purple) is adjusted in each trial according to uncertainty about the current helicopter position (pink) and surprise (blue), as indexed by the posterior probability with which a particular outcome was generated as a consequence of an unlikely generative event such as a change-point **(D)** or oddball **(E)**. The model is fully aware of the generative environments and thus increases learning from surprising information when in the change-point context **(D)** but decreases learning from surprising information in the oddball context **(E)**.

(‘change-point condition’); or (ii) change position slightly from one trial to the next according to a normally distributed random walk with mean zero and standard deviation (SD) 7.5 (‘oddball condition’). In both conditions, bag positions were typically drawn from a normal distribution surrounding the actual helicopter location (SD = 20). However, in the oddball condition, bag positions were occasionally sampled from a uniform distribution across the entire screen space (hazard rate = 0.125). Thus, in the change-point condition, unexpected bag locations were indicative of a new helicopter position, thereby incentivizing updating the bucket position (Fig. 1B). In contrast, in the oddball condition, unexpected bag locations were typically indicative of a one-off outlier (oddball), thereby incentivizing participants to leave the bucket in its previous position (Fig. 1C).

Task incentives were provided in two ways that together controlled total payment for the task. In one incentive condition, participants could accumulate rewards by catching coins of a specific colour (blue, green or red; counterbalanced across participants). In the other incentive condition, participants were endowed with a fixed number of rewards that would be lost in proportion to the number of dropped items that were missed on a given trial. In both conditions, the total rewards were displayed as a pile of coloured tokens inside of the bucket. Thus, our task included two statistical conditions (change-point and oddball) and two incentive conditions (appetitive versus aversive framing) and each participant completed at least 100 trials of each combination of these factors for a total of 400 trials. We found no main effect or group difference in the effects of the incentive condition and thus for all analyses we combine data from the appetitive and aversive conditions (Supplementary material). The majority of participants (124 of 134) completed exactly 100 trials of each condition, but the first 10 participants (seven patients, three control subjects) completed 200 trials of each condition. Here we pool data from all participants, but results were similar when excluding the participants that completed 200 trials per condition rather than 100.

Normative model of learning rate adjustment

To understand how participants should update beliefs in our different conditions, we used a normative learning model that has been described previously (Nassar *et al.*, 2010, 2016, 2019). The model is derived from the full Bayesian ideal observer (Adams and MacKay, 2007; Wilson *et al.*, 2010; Stephan *et al.*, 2016) by approximating the optimal predictive distribution with a Gaussian distribution that has a matched mean and variance (Nassar *et al.*, 2010, 2019; Kaplan *et al.*, 2016). One key advantage of this approximation to optimal belief updating is that it leads to an error-driven learning rule in which the influence of incoming prediction errors, which we refer to as the learning rate and use throughout the text to quantify the degree of belief updating, is adjusted from trial to trial. Normative learning rates are adjusted according to two latent variables: surprise and uncertainty. Surprise indexes the probability with which the model believes a new observation to have come from an alternate process (either change-point or oddball, depending on condition). Surprise is estimated as a posterior probability that the event was driven by an alternate process than that expected, and thus depends on a likelihood term that grows with (i) prediction error magnitude (Nassar *et al.*, 2010); and (ii) the prior

probability or ‘hazard rate’ assigned to surprising events (change-points/oddballs). Uncertainty indexes the degree to which the model is uncertain about the current helicopter location, and is analogous to the Kalman gain in a Kalman filter: intuitively, when the current estimate is uncertain, any deviant observation should have a larger influence on updating that estimate. High levels of uncertainty drive the normative model to learn more rapidly in both conditions (Fig. 1D and E). However, surprise affects model behaviour differentially in the two conditions. In the change-point condition, where surprising errors are indicative of change-points and thereby predictive of future outcomes, high levels of surprise drive the model to use high learning rates (Fig. 1D). In contrast, in the oddball condition, where surprising errors are indicative of one-off outliers that do not predict future outcomes, high levels of surprise lead to reductions in prescribed learning rate (Fig. 1E).

Performance-based exclusion measures

In general, participant data indicated compliance with the basic task objectives. However, in a small number of participants with extremely poor performance, it was not completely clear whether the participants were genuinely attempting to perform the task. To remove participants on this extreme, we set a criterion on the mean absolute trial error (distance between centre of bucket and true helicopter position) and excluded participants who failed to meet this benchmark from our analysis (mean distance of helicopter < 32; Supplementary Fig. 2). This led to exclusion of seven patients and one control participant. Including these participants in our analysis did not substantially change our key findings.

Predictions of increased hazard rate in normative model

To test a recent computational instantiation of the aberrant salience hypothesis (Kaplan *et al.*, 2016; Stephan *et al.*, 2016), we simulated behaviour in our task from a version of the normative model that incorporated an overestimation of the frequency of abnormal events. Specifically, we changed the hazard rate from its ground truth value 0.125 to 0.4. We simulated data for all task sessions from models containing ground truth and elevated hazard rates.

We examined the average updating behaviour of simulated models, along with human participants, using a sliding window regression approach. Data for each simulated subject were binned according to absolute prediction error and bucket updates (e.g. the signed change in bucket location from one trial to the next) from trials in the bin were regressed onto an explanatory matrix that included trial prediction errors (e.g. the signed difference between the bucket position and the observed bag location). Bins were set according to a sliding window that began at a minimum absolute prediction error (smallest 5% of absolute prediction errors) and ended at the largest prediction errors (largest 5% of absolute prediction errors). The central bins included 25% of the total absolute prediction errors; however bins were narrower on the extremes in order to visualize changes in updating at extremely small or extremely large prediction errors. The slope of the relationship between prediction errors and subsequent updates provides a measure of the

average learning rate, and thus our sliding window procedure allowed us to test how learning rate depends on absolute prediction error.

Extracting trial-to-trial learning rates

In addition to examining learning rate by averaging across trials we assessed the influence of information presented on individual trials as trial-wise learning rates (Nassar *et al.*, 2010). Specifically, we defined learning rates as the bucket update on each trial divided by the prediction error observed prior to that update. Learning rates >1 were rounded to 1 and learning rates <0 were rounded to 0. Learning rates were binned into 20 equally spaced bins between 0 and 1 for display—and for learning rate histograms displayed in figure three trials that did not contain an appreciable prediction error (absolute prediction error < 20) were omitted—although results were similar when these trials were included.

Trial-to-trial learning rates were used to identify three update categories: (i) non-updates (learning rate ≤ 0.1); (ii) moderate updates (learning rate >0.1 and <0.9); and (iii) total updates (learning rate ≥ 0.9). Overall relative frequency of each update category was computed for all participants and compared across groups using two sample t -tests. Each single trial update category (non, total, moderate) was regressed onto an explanatory matrix that included factors likely to affect learning rate including surprise, condition (change-point/oddball), surprise \times condition interaction, uncertainty, and trial value.

The proportion of learning attributable to total updates for each subject was calculated by dividing the frequency of total updates, scaled by the average learning rate, by the average learning rate across all trials for a given subject. Individual differences in this proportion were regressed onto an explanatory matrix that included a binary patient category variable as well as a continuous variable constituting the average learning rate for each participant. This same regression was applied to individual differences in the frequency with which subjects used learning rates in each of the 20 discrete learning rate bins to examine which specific learning rates were over- and underused by patients, after controlling for average learning rate.

Characterizing flexibility and precision of participant beliefs

Previous work that examined how different belief updating strategies affect the precision and flexibility of beliefs has primarily focused on changes in average updating behaviour. Here we develop a new method to examine the precision and flexibility of beliefs that makes use of the entire sequence of observed trial-to-trial learning rates described above.

The method first involved re-representing each belief as a weighted mixture of previous outcomes. Specifically, we stepped through the sequence of single trial learning rates and for each trial: (i) assigned weight to the newest outcome proportional to the learning rate on that trial; and (ii) assumed that the remaining weight $(1 - \text{learning rate})$ was divided in proportion to the weight assignments from the previous trial. Through this process, the subject belief on each trial is recast as a weighted average of previous outcomes through the following equivalency (Sutton and Barto, 1998):

$$B_t = (1 - \alpha)^t B_0 + \sum_{i=1}^{t-1} \alpha_i (1 - \alpha_{i+1})(1 - \alpha_{i+2}) \dots (1 - \alpha_t) X_i \quad (1)$$

where X_i is the position of the bag on the i th trial, α_i is the learning rate on the i th trial, and B_t is the bucket updated position on trial t . In short, the belief on trial t is a weighted average of previous outcomes, where the weight of each previous outcome is related to the learning rate describing the update immediately after that outcome (α_i), and negatively related to learning rate describing updates to subsequent outcomes $[(1 - \alpha_{i+1})]$. This procedure can be thought of as projecting the participant belief into the space of the previous outcomes that contributed to it, whereby the dot product of the weight and the corresponding outcome history perfectly reproduces the participant belief for each trial.

Note that a rational agent should flexibly alter the weight assigned to previous outcomes depending on whether those outcomes are perceived to have occurred before or after a change-point. We therefore assessed participants' flexibility of beliefs by quantifying the fraction of the weights that correspond to trials occurring during the relevant context. In the change-point condition this corresponds to the fraction of weight assigned to outcomes occurring since the most recent change-point. The theoretical precision of beliefs (under the assumption that all observations were drawn from the same mean with independent variance) was also computed for each trial as follows:

$$\tau = \frac{1}{\sigma_{tot}^2} = \frac{1}{(w_1 \sigma_{samp})^2 + (w_2 \sigma_{samp})^2 \dots + (w_{t-1} \sigma_{samp})^2} \quad (2)$$

where σ_{tot}^2 is the variance on the weighted mean of samples, σ_{sample}^2 is the variance on each sample, and w reflects the weight given to that sample during updating. If all weight falls on a single outcome, then precision goes to one over the sample variance, or to the precision of a single observation. In contrast, if N weights go to one over N (e.g. N outcomes equal weight), then the precision goes to N over the sample variance, corroborating the intuition that averaging across a greater number of trials should produce higher precision. Since the denominator of this expression always includes the sample variance, we calculate effective sample size by expressing precision as a function of the precision of a single sample:

$$\text{Effective samples} = \frac{\tau}{\sigma_{samp}^2} \quad (3)$$

While this measure can be computed for all trials, of particular interest is the degree to which this measure of belief precision grows during stable periods of the task, when participants could, in fact, be integrating information over a large number of outcomes. Thus, for statistical testing we examine precision of trials in which at least eight observations had been made since the most recent outcome, and we refer to this value as asymptotic precision.

Model fitting

To understand the computational differences between the patients and controls, we fit an extended version of the normative model to participant behaviour (Nassar *et al.*, 2016, 2019). In short, the model updated beliefs about the helicopter location

as described above, and produced bucket positions from a normal distribution centred on the inferred helicopter location. The model is described completely in the [Supplementary material](#) and included the following free parameters, which have been described previously: (i) hazard rate: frequency that the model expected extreme events (change-points/oddballs) (Nassar *et al.*, 2010, 2019; d'Acromont and Bossaerts, 2016); (ii) likelihood weight: the degree to which extremeness of an outcome factored into identification of change-points/oddballs (Nassar *et al.*, 2010); (iii) uncertainty underestimation: degree to which uncertainty is inappropriately reduced on each trial (Nassar *et al.*, 2016); (iv) drift scale: the rate at which the helicopter was assumed to be drifting in the oddball condition (Nassar *et al.*, 2010); (v) update variance: the base width of the distribution over possible bucket positions centred on the inferred helicopter location (Nassar *et al.*, 2016); and (vi) update variance slope: the degree to which the width of the distribution over bucket positions increases with larger normative updates (Nassar *et al.*, 2016; Findling *et al.*, 2019).

The model also included two additional parameters that modelled (i) the frequency with which prescribed updates in the oddball condition were generated as if in the change-point condition [proportion context error (oddball)]; and (ii) the frequency with which prescribed updates in the change-point condition were generated as if in the oddball condition [proportion context error (change-point)].

The model included one additional change to capture the prevalence of non-updates in participant data. Specifically, the model included two terms to model a probability that a given trial would include a perseverative response, i.e. updates were set to zero for the trial. To capture the selectivity of perseverative responses apparent in the participant data, perseveration probability was determined by the probability density of the prescribed update on a scaled normal distribution (mean = 0, SD = perseveration width), where the scale term was set such that the perseveration probability on a trial with a prescribed update of zero would be equal to perseveration max, which was fit as an additional free parameter in the model.

The extended model, along with several simpler models, were fit to participant data through likelihood maximization using `fmincon` in MATLAB. Model comparison was conducted through Bayesian model selection (Stephan *et al.*, 2009) using $-1/2$ AIC as the model evidence. Parameter estimates for the best fitting model were regularized by refitting the model using posterior probability maximization and an informed prior over parameters derived from the original maximum likelihood fits. Predictive checking was performed by simulating task performance (one-step look ahead) for each participant using the maximum *a posteriori* model parameters fit to that participant.

Individual differences analysis

Model estimated parameters were included in a logistic regression to determine whether they can be used to predict (classify) patient status (Wiecki *et al.*, 2015; Huys *et al.*, 2016). The logit-transformed predictions from this model corresponded to continuous patient scores, with higher values corresponding to participants who had parameter profiles more similar to patients, and lower values corresponding to participants who had parameter profiles more similar to controls. We examined how these continuous parameter scores related to disease symptomology

by correlating them with measures of positive symptoms (the average rating on the four psychosis items from the BPRS: grandiosity, suspiciousness, unusual thought content, and hallucinations), general negative symptoms (the average rating on all items from the SANS), motivational deficits (the average rating on items from the avolition/role-functioning and anhedonia/asociality subscales of the SANS), and a composite measure of cognitive function (from the MATRICS battery).

Patient status classification

Binary patient status (schizophrenia, control) was predicted using a leave-one-subject-out logistic regression using three separate sets of predictors: (i) non-update and moderate update frequencies (two predictors); (ii) information theoretic predictors including mean per cent relevant context and mean effective samples for each condition (four predictors); and (iii) model parameters from quantitative model fits (10 parameters). For each participant, the classifier was trained on all participants but one, and a prediction score for the left out participant was computed as the dot product of the model coefficients and left out participant prediction scores. Out of sample prediction scores were sorted and used to construct a receiver operating characteristic (ROC) curve. Area under the ROC curve (AUC) was estimated using a trapezoidal approximation (`trapz` in MATLAB).

Permutation testing was conducted by permuting the patient labels 1000 times and repeating the same analysis described above. *P*-values were assigned to AUC scores as $1 - \text{the frequency of encountering an AUC value as large as the observed value in the permutation distribution}$.

Data availability

All code and behavioural data in this manuscript will be made available on the corresponding author's website (<https://sites.brown.edu/mattlab/resources/>).

Results

To characterize alterations in belief updating we examined the behaviour of patients and controls in a computerized predictive inference task framed as an attempt to infer the location of a helicopter (Fig. 1A). Participants moved a bucket to the location at which they believed a helicopter to be hovering overhead through a series of button presses. After indicating satisfaction with their bucket placement, the helicopter would drop a bag containing potentially valuable contents and participants would collect the contents that landed in their bucket. During training, participants could see the helicopter and place the bucket accordingly. However, in the testing phase, the helicopter was covered with clouds and participants were required to infer the position of the helicopter based on the bags that had previously fallen from it. Participants completed both training and testing in two separate statistical contexts. In the first condition, the helicopter was typically stationary but occasionally relocated to an alternate screen position (change-point condition; Fig. 1B). In the second condition, the helicopter drifted

slowly from one trial to the next, but occasionally a bag location was chosen uniformly across the entire range of possible screen positions, rather than being sampled from a location nearby to the helicopter (oddball condition; Fig. 1C).

Behaviour of participants and a normative learning model (McGuire *et al.*, 2014; Nassar *et al.*, 2019) were highly sensitive to the statistical context manipulation. In the change-point condition, the normative model rapidly adjusted beliefs in response to outlying bag locations (Fig. 1B, green) and captured the behaviour of an example participant that did the same (Fig. 1B, yellow). In the oddball condition, the normative model was insensitive to outlying bag locations (Fig. 1C, green) allowing it to capture the same tendency in an example participant (Fig. 1C, yellow). The normative model achieved this context sensitivity by adjusting its sensitivity to new bag locations according to two latent factors, uncertainty and surprise (Fig. 1D and E, pink and blue). ‘Uncertainty’ quantifies the model’s degree of uncertainty about the current helicopter location, with higher levels of uncertainty evoking greater sensitivity to new bag locations, or (in the language of error-driven learning) a higher learning rate (Fig. 1D and E, purple). ‘Surprise’ is defined contextually. In the change-point condition it indicates the likelihood of a change-point, and therefore dictates faster learning to facilitate flexibility in the face of change (Fig. 1D, compare blue and purple). In the oddball condition, surprise indicates the probability that the event is an oddball and thus spikes at times corresponding to the outlying bag observations (*cf.* Fig. 1C and E). Normative learning requires ignoring oddballs, as by definition they do not predict future bag locations, and thus surprise dictates a normative reduction in learning rate in the oddball condition (Fig. 1D, *cf.* blue and purple). Thus, a predisposition towards interpreting deviant events with aberrant salience, such as has previously been suggested to occur in schizophrenia patients suffering from delusions (Kaplan *et al.*, 2016; Stephan *et al.*, 2016) would lead to higher rates of learning in the change-point condition (where detected abnormal events drive learning) but lower rates of learning in the oddball condition (where detected abnormal events prevent learning).

Patients do not display high hazard rate updating behaviours

To quantify this prediction, we simulated belief updates from the normative model equipped with either realistic or unrealistically high expectations about the rate of abnormal events (hazard rate; Fig. 2A). Simulations from both models reveal the general tendency to increase learning rate with unexpectedly large errors in the change-point condition (Fig. 2A) and to decrease learning rate with unexpectedly large errors in the oddball condition (Fig. 2A). The high hazard rate model that expects more change-points and oddballs learns more rapidly in the change-point condition, but more slowly in the oddball condition, when compared to a

model equipped with the appropriate hazard rate (Fig. 2A; compare blue to yellow). However, belief updating of schizophrenia patients does not match the qualitative predictions of this high hazard rate model (Fig. 2B), suggesting that patients do not ascribe heightened salience to all observations.

Patients less frequently combine information to form integrated beliefs

While only minimal differences in belief updating were apparent in trial-averaged data, a key advantage of our task is that it allows us to measure the influence of individual outcomes on beliefs, by computing single trial learning rates (Nassar *et al.*, 2010). Single trial learning rates reflect the degree of belief update on a given trial as a fraction of the prediction error observed on that trial. Thus, a learning rate of 1 indicates that the participant moved the bucket to the exact location of the most recent bag (total updating), whereas a learning rate of 0 indicates that the bucket was maintained in its previous position (non-updating). Moderate learning rates between these two extremes indicate an updated belief that combines the prior belief with new outcome information, and thereby facilitate the integration of new and old information. Patients and controls used a wide range of learning rates in the change-point (Fig. 3A and C) and oddball conditions (Fig. 3B and D). However, the patient group included fewer moderate learning rates and more zero learning rates than did controls.

To quantify this difference, learning rates were categorized into discrete bins corresponding to non-updates (learning rate near 0), moderate updates (between dotted lines in Fig. 3A–D), and total updates (learning rate near 1). On average, patients used more non-updates [mean/standard error of the mean (SEM) for patients: 0.49/0.02 and controls: 0.37/0.03, $t(124) = 3.3$, $P = 0.001$; Fig. 3E and F] and fewer moderate updates [mean/SEM for patients: 0.33/0.02 and controls: 0.45/0.02, $t(124) = -4.1$, $P = 7 \times 10^{-5}$; Fig. 3G and H] than did controls, whereas the frequency of total updates did not differ between the two groups [mean/SEM for patients: 0.15/0.02 and controls: 0.16/0.02, $t(124) = -0.4$, $P = 0.66$]. The frequency of moderate and non-updates differed consistently across groups (Fig. 3K and L) and could be used to classify patient status (AUC = 0.73, permutation $P = 0.001$).

While total updates did not differ between the groups (Fig. 3I and J), they seemed to account for a larger fraction of the total learning in patients, relative to controls. The proportion of learning attributable to total updates was higher for subjects that used higher learning rates on average and for a given average learning rate tended to be higher in patients than controls (Fig. 4A). A regression model constructed to explain individual differences in the proportion of learning that was attributable to total updates based on (i) patient status; and (ii) average learning rate revealed

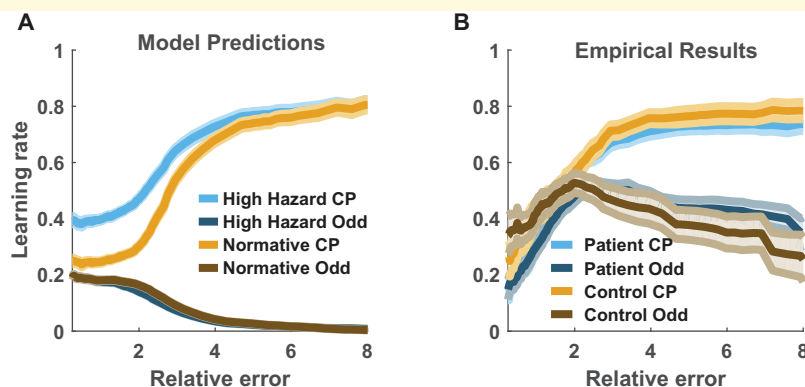


Figure 2 Schizophrenia patients do not display heightened sensitivity to unlikely events. **(A)** Synthetic updating behaviour generated by a normative model (yellow) and the same model equipped with a heightened sensitivity to detect unlikely events, implemented in the normative framework as an abnormally high prior on such events (hazard rate; blue) was regressed onto prediction errors in sliding windows of absolute prediction error magnitude (x-axis). The resulting slope, termed the learning rate (y-axis), increases with prediction error magnitude in the change-point condition (lighter colours) but decreases with prediction error magnitude in the oddball condition (darker colours). Higher hazard rate (blue) leads to a leftward shift in both curves, reflecting a higher sensitivity to small changes in prediction error, particularly for moderate prediction error magnitudes. **(B)** Patient (blue) and control (yellow) participant learning (y-axis), assessed in the same manner, displays a qualitatively similar bifurcation of learning in the two conditions (dark = oddball, light = change-point) with increased prediction error magnitude (x-axis); however, patient curves are not shifted leftward with respect to control curves, as would be predicted by an increased hazard rate. There is not a leftward shift of the blue curves relative to the yellow (as would be expected under the high hazard rate hypothesis) nor is there a consistent offset in the learning rate of patients relative to controls (compare blue and yellow on ordinate). However, across conditions there was a modest reduction of learning rates in patients relative to controls [mean/SEM learning rate for patients: 0.34/0.02 and controls: 0.41/0.03, $t(124) = -2.0$, $P = 0.04$]. CP = change-point.

significant positive coefficients for both patient status and average learning rate (patient coefficient = 0.78, t statistic = 2.59, degrees of freedom = 123, $P = 0.01$). When the same regression model was applied to explain individual differences in the frequency of learning rates ranging from 0 to 1, it revealed that patients tend to overuse both very small and very large learning rates when compared to controls of a similar average learning rate (Fig. 4B). In contrast, patients less frequently used moderate learning rates, particularly small moderate learning rates, than did control subjects.

Moderate updates, needed to integrate information across multiple samples, are most important during periods of uncertainty when existing beliefs are based on a small number of observations (Nassar *et al.*, 2010). As might be expected based on this idea, control participants selectively increased their use of moderate updates during periods of uncertainty [Fig. 4C; mean (95% confidence interval, CI) moderate update slope = 0.032 (0.021, 0.042), $t(30) = 6.33$, $P = 5.5 \times 10^{-7}$ mean (95% CI) total update slope = -0.0006 (-0.008, 0.007), $t(30) = -0.17$, $P = 0.87$]. While patients did increase moderate updates somewhat during periods of uncertainty [Fig. 4D; blue line; mean (95% CI) slope = 0.024 (0.017–0.030), $t(94) = 6.60$, $P = 2.4 \times 10^{-9}$], they also increased their use of total updates [Fig. 4D; red line; mean (95% CI) slope = 0.013 (0.008–0.018), $t(94) = 5.00$, $P = 3.2 \times 10^{-6}$]. Thus, while both groups adjusted learning rate according to uncertainty, the patient group often did so by completely replacing their prior belief, rather than combining it with newly arriving information.

Patient beliefs are both less flexible and less precise

Based on the observed differences in single trial learning rates, and their modulation by uncertainty, we sought to examine how the exact sequence of learning rates might affect beliefs. Typically, learning rate is measured on average across a large numbers of trials, and under such circumstances there is a well-established stability/flexibility trade-off: faster learning leads to better performance after change-points (high flexibility) but worse performance during periods of extended stability (low precision) (Behrens *et al.*, 2007; Nassar *et al.*, 2010; Franklin and Frank, 2015). Access to single trial learning rates allows us to examine this trade-off by considering not just the mean learning rate, but the exact sequence of learning rates employed in the task.

To do so, we developed a novel method for computing the flexibility and precision that relies on the key insight that beliefs can be recast as a weighted average of previously observed outcomes (see ‘Materials and methods’ section). To gain an intuition for this, consider a situation in which the participant has observed only two outcomes, the bucket is positioned at the location of the first of those outcomes, and the participant needs to update the bucket in response to the second outcome (Fig. 5A). If the participant does not update the bucket position at all, his updated belief is equivalent to a weighted average that gives all weight to the first of the two outcomes (Fig. 5A, left). In contrast, if the participant updates the bucket position completely to the most recent

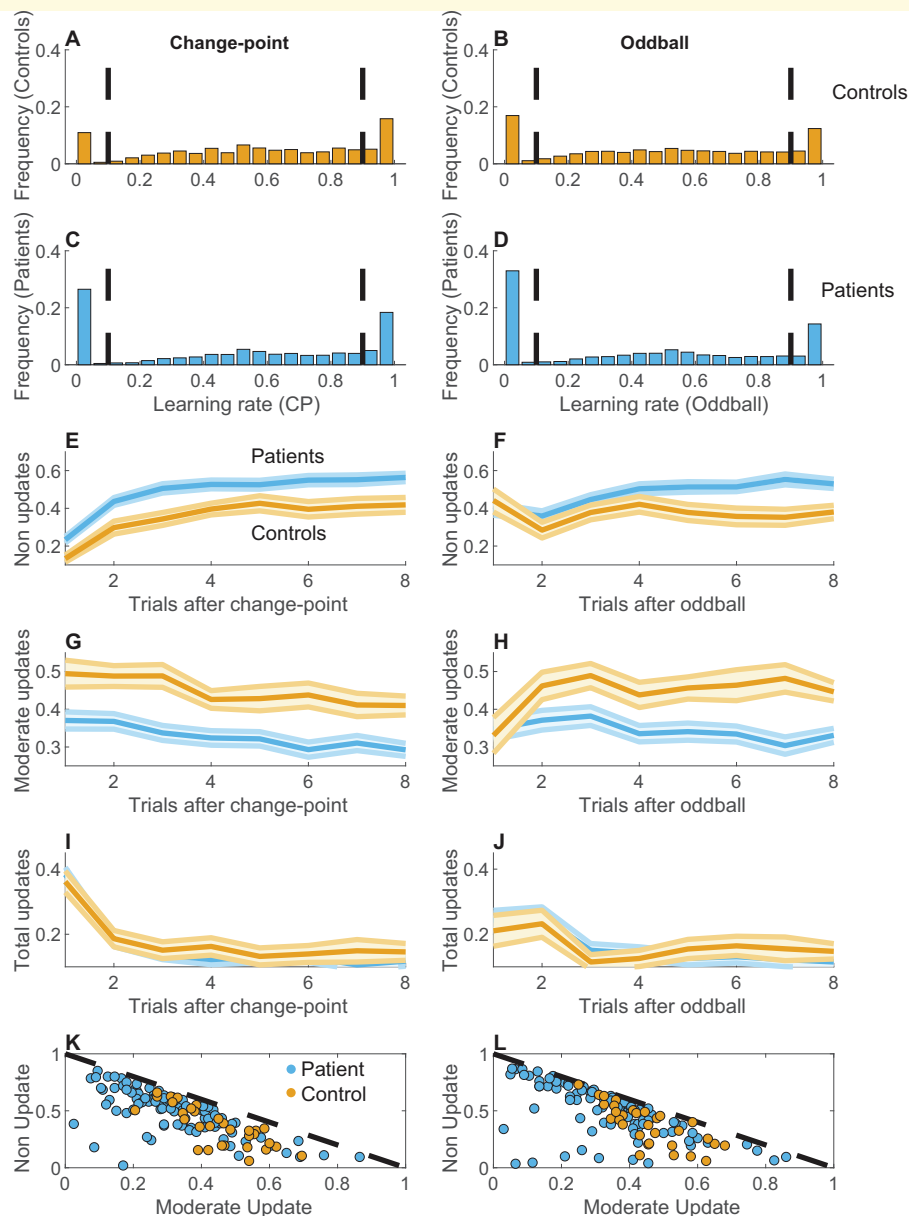


Figure 3 Schizophrenia patients make moderate updates less frequently than matched controls. (A–D) Learning rate frequency histograms depicting the relative frequency with which controls (A and B) and patients (C and D) used specific single trial learning rates (x-axis) in the change-point (CP) (A and C) and oddball (B and D) conditions. Dotted lines indicate thresholds used to group single trial learning rates into non-updating (left), moderate updating (middle), and total updating (right) categories. (E and F) Mean/SEM (lines/shading) frequency of non-updating in patients (blue) and controls (yellow) is plotted as a function of the number of trials after a change-point (E) or oddball (F). (G and H) Mean/SEM (lines/shading) frequency of moderate updating in patients (blue) and controls (yellow) is plotted as a function of the number of trials after a change-point (G) or oddball (H). Patients used moderate updates less frequently, and non-updates more frequently than controls [mean/SEM moderate updates for patients: 0.33/0.02 and controls: 0.45/0.02, $t(124) = -4.1$, $P = 7 \times 10^{-5}$; mean/SEM non updates for patients: 0.49/0.02 and controls: 0.37/0.03, $t(124) = 3.3$, $P = 0.001$]. (I and J) Mean/SEM (lines/shading) frequency of total updating in patients (blue) and controls (yellow) is plotted as a function of the number of trials after a change-point (I) or oddball (J). There was no statistical difference in total updates between the groups [mean/SEM total updates for patients: 0.15/0.02 and controls: 0.16/0.02, $t(124) = -0.4$, $P = 0.66$]. (K and L) Mean non-update frequency (y-axis) is plotted against moderate update frequency (x-axis) for individual patients (blue points) and controls (yellow points) in change-point (K) and oddball (L) conditions.

outcome, then his belief can be recast as a weighted history where all weight lies on the most recent outcome (Fig. 5A, right). If the participant updates moderately (e.g. learning

rate = 0.5) then weight will be attributed to each of the two outcomes (Fig. 5A, middle). Under the assumption that all weight is attributed to outcomes from the same generative

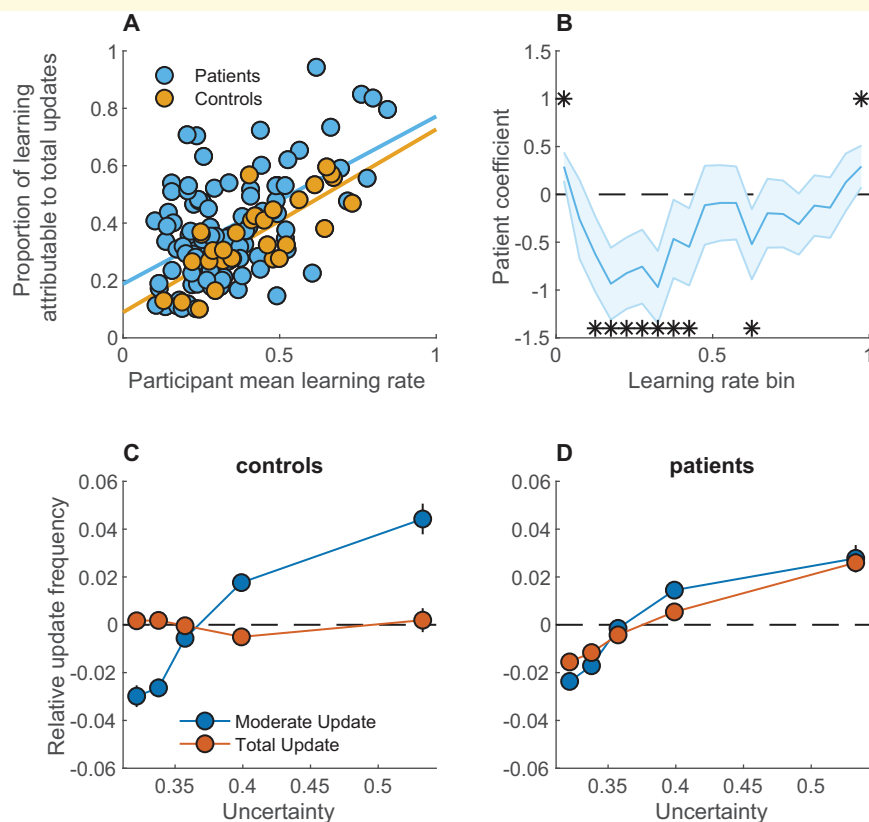


Figure 4 Patients rely more on total updates, particularly when uncertain. (A) For participants who used higher learning rates on average (*x*-axis), a greater proportion of learning was attributable to discrete total updates (single trial learning rates > 0.9; *y*-axis). For any given average learning rate, patients (blue) tended to be more reliant on total updates than were controls (orange). Points reflect individual subject and lines reflect least squares fits to separate groups. (B) Coefficients (*y*-axis) reflecting the contribution of patient status to predictions about the frequency of discrete bins of single trial learning rate (*x*-axis). Positive/negative values indicate more/less frequent use of a particular category of learning rate by patients after controlling for average rate of learning. Line/shading reflects mean/95% confidence intervals and asterisks reflect significant differences from zero after false discovery rate correction ($P < 0.005$). (C) Control participants used more moderate updates (blue, *y*-axis), but not more total updates (red) on trials in which the normative model indicated a high level of uncertainty (*x*-axis). (D) Patients increased both moderate (blue) and total (red) updates as a function of uncertainty.

process (e.g. helicopter location) the precision of beliefs can be computed analytically from the weight profile, in this case revealing that beliefs resulting from the moderate update are twice as precise as those resulting from either the non-update or total update (Fig. 5A, bottom).

When this method is applied to simulated behaviour from fixed learning rate models, it yields weight profiles in which higher weight is attributed to recent outcomes for a high learning rate model (Fig. 5A, blue) and higher weight attributed to outcomes observed in the distant past for a low learning rate model (Fig. 5A, yellow). Moreover, the exact profile of weights attributed to outcomes in the past depends on the sequence of learning rates; for example, the normative learning model gives nearly equal weight to all observations since the most recent change-point, but no weight to outcomes occurring before that (Fig. 5A, green). In principle, an equal weighting of all outcomes having occurred since the most recent change-point would be optimal in that it

would yield the highest possible belief precision without incorporation of irrelevant outcomes having occurred prior to the most recent change-point.

The profile of weights for a given trial provides insight into the flexibility and precision of beliefs. The flexibility of an updating strategy can be assessed by examining the proportion of weights that are attributed to outcomes that were observed in the current context (e.g. since the most recent change-point). Thus, for the example trial depicted in Fig. 5A, we would conclude that the normative and high learning rate models are flexible, in that they do not incorporate information from the previous context, whereas the low learning rate model does assign some weight to the most recent outcome from the previous context. The weight profile can also be used to infer the effective number of samples incorporated in beliefs, thereby providing a measure belief precision. For example, if all weight was attributed to a single outcome, the effective sample size is one and the belief

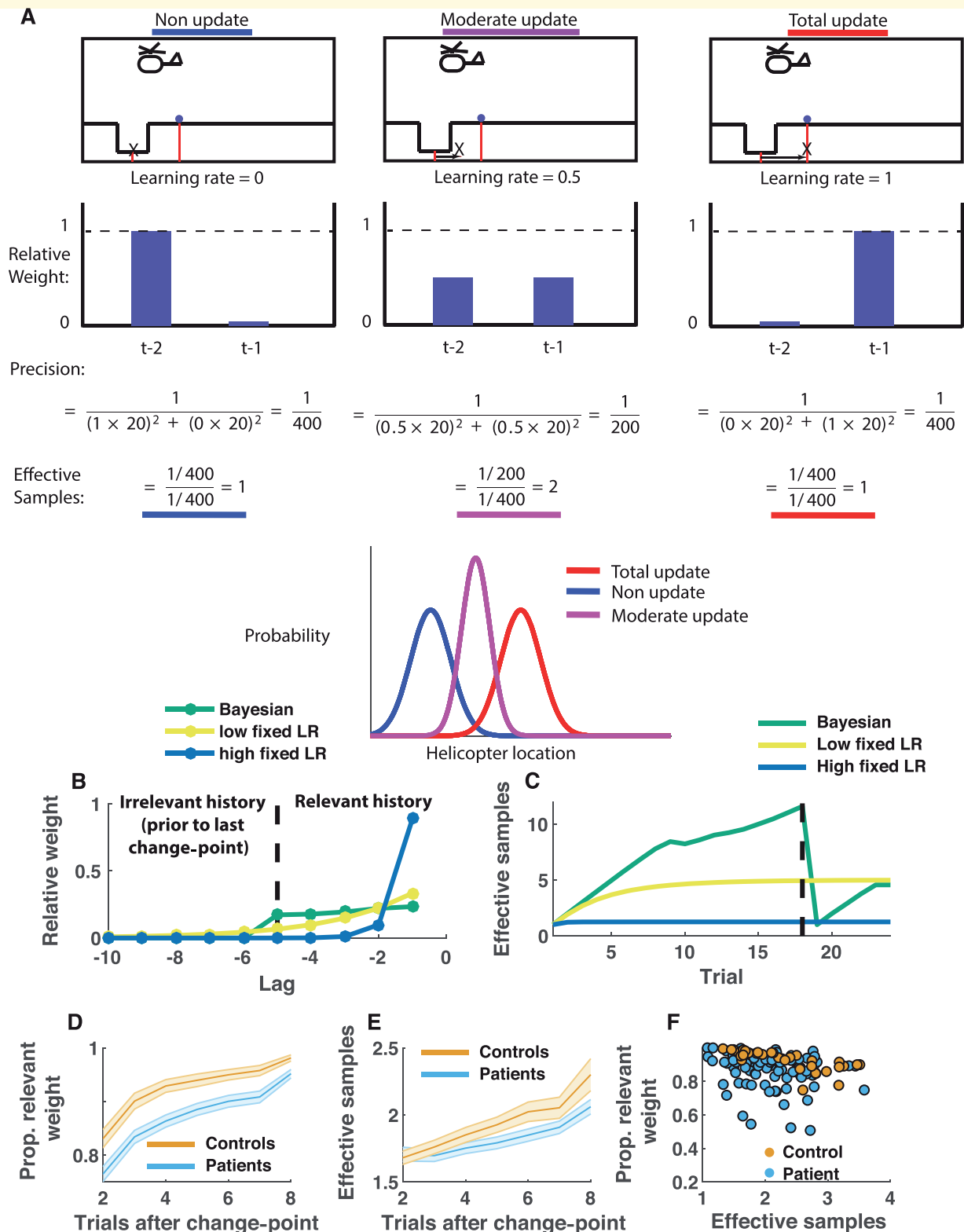


Figure 5 Learning rate sequences used by schizophrenia patients yield beliefs that are both less flexible and less precise than those of control subjects. (A) Schematic depicting the effects of a non-update (left), moderate update (middle), and total update (right) on the precision of an underlying belief distribution. In all cases bucket placement is initialized to a prior outcome (x ; $t-2$) and is updated in accordance with the most recent one (blue dot; $t-1$). The degree of updating used by the agent affects the weight of previous outcomes on the updated bucket position, with the non-update leading to complete reliance on the $t-2$ bag position, the total update relying completely on the $t-1$ bag position, and the moderate update equally weighting these two sources of information (second row). Computing precision of the resulting belief distribution yields a value twice as large for the moderate update than the other two updating strategies, which we quantify as containing two

(continued)

is relatively imprecise. In contrast, if weight were distributed equally across two outcomes, the effective sample size would be two, and the precision increased, although not dramatically. Here we used a generalization of this idea to infer the effective precision of beliefs for any arbitrary weight profile (see ‘Materials and methods’ section). When applied to simple model simulations, our method reveals lower precision beliefs for the high learning rate model and higher precision beliefs for the low learning rate model, consistent with the standard stability/flexibility trade-off (Fig. 5B, blue and yellow). However, the normative learning model achieves even higher levels of precision than the low learning rate model during periods of stability—demonstrating that it is both flexible and precise (Fig. 5B, green).

Applying the same method to participant data revealed that schizophrenia patients are neither flexible nor precise. Patients attributed less weight to the current context than did the controls [Fig. 5C; mean/SEM weight in relevant context for patients: 0.87/0.01 and controls: 0.92/0.01, $t(124) = -2.8$, $P = 0.006$], but contrary to the idea of a flexibility stability trade-off, also formed beliefs that asymptotically contained fewer effective samples than controls [Fig. 5D, mean/SEM effective samples in stable beliefs: 2.1/0.05 and controls: 2.3/0.12, $t(124) = -2.1$, $P = 0.04$]. While there were large individual differences in both measures, our measures of flexibility and precision were capable of classifying participants out of sample with reasonable accuracy (Fig. 5E, AUC = 0.74, permutation $P = 0.001$).

Quantitative model fitting with selective non-updating

Given the differences in the single trial learning rates used by patients and control subjects (Figs 3 and 4) and their apparent effects on the precision and flexibility of beliefs (Fig. 5), we sought to extend our normative model of behaviour to better capture these aspects of patient updating behaviour. To do so, we added two additional parameters that defined the probability with which the model would use a ‘non-update’—implemented as a

learning rate of exactly zero (see ‘Materials and methods’ section). Furthermore, to capture other aspects of behaviour we also added two additional terms in the model to account for potential context errors in which participants used the updating rules from the wrong context (e.g. responding in the change-point condition as if it were the oddball condition). The resulting model provided an improved fit over our original normative model and several other models that were tested (Supplementary Fig. 3), estimated parameters that were recoverable (Supplementary Fig. 4) and simulated updates that qualitatively matched the empirical updating behaviour (Fig. 6A and B).

Parameter estimates from the model also discriminated patients from controls. A logistic multiple regression model that attempted to predict patient category based on each participants’ parameters extracted from our extended normative model provided a reasonably good prediction accuracy (AUC = 0.67, permutation $P = 0.002$). The two parameters contributing most to the identification of patients were related to non-updating (Fig. 6C), with both the peak non-updating probability, and the width of the non-updating function across prescribed updates, being higher in patients (mean/SEM beta for peak and width = 0.15/0.05, 0.16/0.05, t -values = 3.0, 3.3; P -values = 0.003, 0.002). Together, these parameter differences lead to a selective propensity for non-updating in patients for prescribed updates on the same scale as the standard deviation of the noise distribution (Fig. 6D). It is noteworthy that this difference does not persist in trials where the largest updates are prescribed by the normative equations, and thus that patients are able to overcome the perseverative tendency in the situations in which it would be penalized most. Thus, patients do not simply have a greater proportion of lapse trials in which they ignore outcomes altogether, but instead preferentially persevere when moderate updates would be dictated. This selective perseveration was not related to either positive or negative symptoms across the patient group, but was related to cognitive measures (Supplementary Figs 5 and 6).

Figure 5 Continued

effective samples, as opposed to only a single effective sample in the case of a non-update or total update. (B) Single trial learning rates (LR) can be used to calculate the relative weight (y-axis) attributed to previous outcomes at any lag (x-axis). Applying this method to synthetic learning behaviour yields a geometric distribution for fixed learning rate models (blue, yellow) that is unaffected by change-points in the generative structure of the task (depicted by the dotted line at lag -5). In contrast, normative learning models (green) approximate a uniform weight distribution across all lags occurring after the previous change-point but do not assign weight to trials occurring prior to the most recent change-point (dotted line). Flexible belief updating requires that beliefs are based on only relevant information, that is, that all weight is given to trials occurring since the last change-point. (C) The precision of a belief on a given trial can be computed according to weight attribution profile that gave rise to it. The precision, which can be measured in units of effective samples, increases to an asymptotic value for fixed learning rate models (yellow and blue) but changes dynamically in the normative model (green)—growing almost linearly during periods of stability but rapidly collapsing to one after a recognized change-point. (D) Flexibility, as assessed by the proportion of belief weights attributed to outcomes in the relevant context (y-axis), increased as a function of trials after a change-point (x-axis) for controls (yellow) and patients (blue)—but was consistently higher for controls. (E) Precision, as assessed by the effective number of samples contributing to the reported belief (y-axis), also increased as a function of trials after a change-point (x-axis), and did so more rapidly for controls (yellow) than for patients (blue). (F) Differences in flexibility (proportion relevant weight) and precision (effective samples) were prominent in a large number of individual patients.

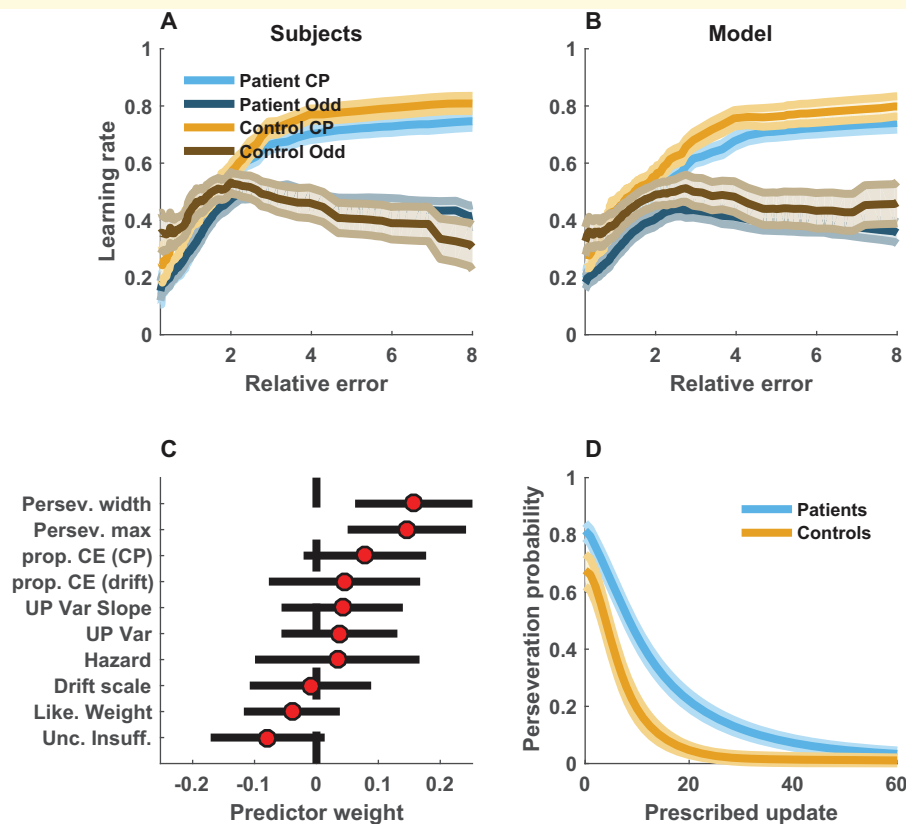


Figure 6 Direct model fitting suggests that patients use more non-updates than control participants. **(A)** Patients (blue) and controls (yellow) both tended to increase learning rate (y-axis) in response to surprising information (higher relative errors; x-axis) in the change-point (CP) condition (light colours), but decrease learning rate in response to surprising information in the oddball condition (dark colours). **(B)** Synthetic data from an extension of the normative model that was fit to patients (blue) and controls (yellow) mimic the reduced learning rate from small errors and the less extreme bifurcation observed in the empirical patient data. **(C)** Regression coefficients and 95% confidence intervals (points and lines; sorted by value) stipulating the contribution of each parameter estimated by the normative model to a logistic regression classifier of patient status. The two parameters governing the magnitude and shape of the perseverative response profile (Persev. Width, Persev. Max) made significant positive contributions to the classifier. **(D)** Perseveration probability as a function of the model-prescribed update is plotted separately for patients (blue) and controls (yellow). Note that perseveration did not differ uniformly across task conditions, but most prominently when the model prescribed making a relatively small update.

Discussion

Schizophrenia is characterized by persisting abnormal beliefs, or delusions. Previous work has theorized that such delusions might emerge through aberrant salience assigned to incoming information (Kapur, 2003), and previous behavioural and neuroimaging work has supported this idea by formalizing aberrant salience in terms of a heightened predisposition towards believing that new observations come from an alternative process (Kaplan et al., 2016; Stephan et al., 2016). Here, we directly test key predictions of this formalization, and decouple them from related cognitive processes, including the learning rate itself. We found no evidence that patients are more likely to categorize new information as a signal rather than noise (e.g. high hazard rate), nor did we see pronounced differences in the average learning rate in patients relative to control subjects (Fig. 2). Instead, we observed that patients update beliefs more often in a binary

fashion, infrequently relying on moderate learning rates that allow integration of new and old information (Fig. 3). This subtle difference in belief updating, which is masked when learning is averaged across trials, has negative consequences for both the flexibility and precision of stored beliefs (Fig. 5). We can account for these differences by extending a normative model to include a non-updating function that probabilistically converts small updates into non-updates, with the parameters of this function elevated to describe patient behaviour (Fig. 6). In addition to providing categorical discrimination of patients from controls, these parameters relate to measures of overall cognitive function, but not to clinical measures of positive and negative symptoms (Supplementary Fig. 5). Taken together, our results argue against a computational formalization of aberrant salience theory, demonstrate the importance of how learning is patterned in time, and reveal that the primary belief updating deficit in schizophrenia is in the integration of new and old

information through moderate learning. Interestingly, this lack of integration could lead to either over-learning or under-learning, depending on the specific task, and therefore may resolve tension between observations that patients learn more slowly in some tasks (Goldberg *et al.*, 1987) but more quickly in others (Kaplan *et al.*, 2016).

Belief updating abnormalities in schizophrenia

Expression of persisting abnormal beliefs is a common positive symptom in schizophrenia. Clinical pharmacology studies suggest that this and other positive symptoms might be related to abnormal dopamine signalling. However, the cognitive mechanisms through which abnormal beliefs arise remain elusive. A major roadblock has been in the identification of behavioural tasks that are capable of eliciting abnormal beliefs in patients and simultaneously distinguishing between candidate mechanisms.

One line of research has suggested that abnormal beliefs might arise from so-called ‘jumping to conclusions’—forming beliefs based on a small amount of evidence. This idea was spurred by research studies using the ‘Beads Task’ (Phillips and Edwards, 1966) where participants are able to draw any number of beads from an urn before reporting a belief as to the predominant bead colour in the urn. These studies suggest that patients, as well as healthy control subjects who are susceptible to delusions, tend to draw very few beads before making a judgement on the predominant colour (Evans *et al.*, 2015). However, a recent study that improved on the standard beads paradigm, to control for potential confounds, arrived at a very different conclusion: patients with more severe delusions tended to seek more information (Baker *et al.*, 2019). Interestingly, computational analysis of these severely delusional patients revealed that they were over-using information presented at the beginning of each task trial, as if they were relying too heavily on prior information and underutilizing contradictory evidence (Baker *et al.*, 2019). This finding, along with a related observation in a conditioned hallucination paradigm (Powers *et al.*, 2017), has suggested that overly strong priors might be the cognitive abnormality that gives rise to delusions (Corlett *et al.*, 2019).

Our data, in broad strokes, are consistent with this idea. An extremely strong prior about the helicopter location in our task would lead to non-updating for small prediction errors that are assumed to be dropped from the ‘well-known’ helicopter location, and complete updating for trials in which a change-point in helicopter location had occurred. In short, an extremely narrow prior distribution in our change-point task condition should lead to the sort of binary updating behaviour that we observe in patients. However, a closer look reveals some discrepancies between our results and this idea. First, we do not see between-group differences in the ‘uncertainty underestimation’ model parameter designed to capture individual differences in uncertainty, but

rather on the perseveration parameters (Fig. 6). Admittedly, these parameters capture related aspects of the behaviour, although it is noteworthy that our patients did adjust learning according to trial-to-trial differences in uncertainty, although they did so differently than the controls, with a greater tendency to update beliefs completely according to new information during periods of uncertainty (Fig. 4). This would not be expected of an overly narrow prior in our task (Nassar *et al.*, 2016). A second discrepancy between our results and the strong prior account of delusions is that we did not see any relationships between our model parameters and positive symptoms of schizophrenia. Instead, we see relationships with a broad array of cognitive measures, suggesting that our indices are tapping into processes different from those assessed in the studies mentioned above. However, one important consideration is that our patients were stably medicated, leading to lower positive symptom profiles and potentially limiting our ability to detect relationships between our task measures and positive symptoms.

One potential issue related to interpreting our findings is that the level of explicit understanding of our task may have differed between patients and controls. While we can rule out the simplest version of this idea from the observation that patients tended to place buckets appropriately when the helicopter was visible, there are more nuanced versions of this concern that might be more difficult to discount. In particular, we cannot guarantee that participants always remembered which condition they were in. Indeed, the context error parameters in our model suggest that both groups occasionally updated bucket locations in a manner more appropriate for the alternate context. There was a trend for the context error parameters to take larger values when fit to patients (Fig. 6C), and while these terms did not differ significantly across groups, it does seem that our best fitting model did not completely capture the discrepancy between patients and controls in asymptotic updating for large errors in the two conditions (*cf.* Fig. 6A and B), and it is possible that a more complex model may be better able to tease this difference out. Nonetheless, it is hard to imagine any such context confusion effects accounting for our primary observation, which included a reduced frequency of moderate learning rates in the patient group in both change-point and oddball conditions.

Patterns of learning and the stability flexibility trade-off

Previous studies examining belief-updating behaviour in schizophrenia have relied on computational modelling to infer participant beliefs based on choices. Here we measured beliefs directly, which allowed us to characterize the weighted history of influences on each belief. This allowed us to examine the degree to which the patterns of learning in patients differed from controls, revealing that patients tend to rely on fewer effective samples than controls, and that the samples they do rely on are more frequently irrelevant to the active statistical context.

This set of results deviates from a common interpretation of learning rate, or even the hazard rate, as mediating a trade-off between stability and flexibility (Behrens *et al.*, 2007; Glaze *et al.*, 2018). Instead, our results highlight the importance of specific learning rates (in particular, moderate learning rates that mediate integration of new and old information) and the manner in which learning is distributed across trials (Gallistel *et al.*, 2014). While healthy young adults and normative learning models demonstrate a trade-off in belief stability and flexibility (Nassar *et al.*, 2010), we show that patients and controls do not differ in this sort of continuum; instead patients inefficiently distribute learning across time so as to form beliefs that are both less flexible and less precise than those held by control subjects. This feature would have been missed by averaging learning across trials, as it was largely attributable to the overall change in the distribution of learning rates in patients, with learning taking more of an all-or-none nature, thereby limiting the degree to which information can be integrated across multiple observations. Our extended normative model was capable of capturing both binary and continuous aspects of belief updating, potentially bridging an important gap between existing models of learning (Nassar *et al.*, 2010; Gallistel *et al.*, 2014). To the best of our knowledge, our study is the first to examine the implications of how learning rates are sequenced in time, thus avoiding the pitfalls of previous studies that estimated learning rates for entire sessions and, as a consequence, might have masked such learning differences across individuals, age groups, and clinical populations.

Indeed, learning has been assessed in schizophrenia patients in a large number of studies using a large number of paradigms. In some of these studies, patients were characterized as switching more frequently than control subjects, suggesting an over-responsiveness to feedback (Yogev *et al.*, 2004; Li *et al.*, 2014; Kaplan *et al.*, 2016). However, in other cases patients have learned more slowly than control subjects and been characterized by perseverative responding (Goldberg *et al.*, 1987; Laws, 1999; Leeson *et al.*, 2009; Reddy *et al.*, 2016; Baker *et al.*, 2019). Previous work and theories have posited that the dominant behavioural feature (over-learning or under-learning) may depend on symptom profiles. However, here we show that both behavioural features can co-occur within individuals.

One important question motivated by this work is what are the biological and cognitive mechanisms through which the extreme updating strategies observed in patients occur? Recent work has suggested that unstructured variability in learning, much like that which we observe in patients, is related to blood oxygen level-dependent activity in regions of frontal cortex including dorsal anterior cingulate cortex (dACC) and ventromedial prefrontal cortex (vmPFC) (Findling *et al.*, 2019). However, it is not entirely clear what those signals might be conveying. One possibility is that the variability arises through the use of multiple systems for learning, with a working memory system sometimes overriding associative learning to contribute a total update (Collins and Frank, 2012; Collins *et al.*, 2017). Existing models of working memory-based systems assume that only

a single memory is selected to generate a response (Collins, 2018). However, in principle, a belief report in our task might be constructed by reading out multiple outcomes stored in working memory. In such a system, moderate updates would require having a large memory that could be used to store previous outcomes such that replacing a stored outcome with a new observation only changes one of multiple stored outcomes. However, by the same token, having an extremely limited capacity (e.g. only capable of storing a single outcome) might force learning into a binary regime (if stored memory is updated, then update is total, otherwise it will be perseverative). Previous work implicating working memory deficits in patients, and suggesting that patient learning deficits are attributable to change in working memory (Collins *et al.*, 2017), provide at least indirect support for this mechanism, and should motivate future work.

Taken together, our results suggest that patients with schizophrenia are more extreme in their belief updates, limiting the degree to which information is integrated across time, and giving rise to beliefs that are both inflexible and imprecise (incorporating fewer data). Our results shed light on why previous reports have noted both over- and under-sensitivity to feedback as core features of schizophrenia and provide a common lens through which these aspects of behaviour can be viewed. Furthermore, our results motivate future work to better understand the cognitive operations underlying moderate belief updates, and how these operations are impaired in schizophrenia.

Acknowledgements

We would like to thank Leeka Hubzin, Sharon August, and Ruba Mateen for help collecting behavioural data, Tiantian Li for assembling demographics information and Ben Heasley for programming the original helicopter task.

Funding

This work was funded by National Institutes of Health (NIH) grants F32MH102009 and K99AG054732 (M.R.N.), National Institute of Mental Health (NIMH) R01 MH080066-01. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

- Adams RP, MacKay DJC. Bayesian online change point detection. eprint arXiv:0710.3742 2007. Available from <https://arxiv.org/abs/0710.3742>.
- Andreasen NC. Scale for the Assessment of Positive Symptoms (SAPS). Iowa City, Iowa: The University of Iowa; 1984.
- Baker SC, Konova AB, Daw ND, Horga G. A distinct inferential mechanism for delusions in schizophrenia. *Brain* 2019; 142: 1797–812.
- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. *Nat Neurosci* 2007; 10: 1214–21.
- Collins AGE, Albrecht MA, Waltz JA, Gold JM, Frank MJ. Archival report. *Biol Psychiatry* 2017; 82: 431–9.
- Collins AGE, Frank MJ. How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur J Neurosci* 2012; 35: 1024–35.
- Collins AGE. The tortoise and the hare: interactions between reinforcement learning and working memory. *J Cogn Neurosci* 2018; 30: 1422–32.
- Corlett PR, Horga G, Fletcher PC, Alderson-Day B, Schmack K, Powers AR. Hallucinations and strong priors. *Trends Cogn Sci* 2019; 23: 114–27.
- d’Acremont M, Bossaerts P. Neural mechanisms behind identification of leptokurtic noise and adaptive behavioral response. *Cereb Cortex* 2016; 26: 1818–30.
- Evans S, Averbach B, Furl N. Jumping to conclusions in schizophrenia. *Neuropsychiatr Dis Treat* 2015; 11: 1615.
- Findling C, Skvortsova V, Drommelle R, Palminteri S, Wyart V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat Neurosci* 2019; 22: 2066–2077.
- First MB, Spitzer LR, Gibbon M, Williams JBW. Structured clinical interview for DSM-IV axis I disorders. Washington, DC: American Psychiatric Press; 1997.
- Franklin NT, Frank MJ. A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife* 2015; 4: e12029.
- Gallistel CR, Krishan M, Liu Y, Miller R. The perception of probability. *Psychol Rev* 2014; 121: 96–123.
- Glaze CM, Filipowicz ALS, Kable JW, Balasubramanian V, Gold JJ. A bias-variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nat Hum Behav* 2018; 2: 213–24.
- Goldberg TE, Weinberger DR, Berman KF, Pliskin NH, Podd MH. Further evidence for dementia of the prefrontal type in schizophrenia? A controlled study of teaching the Wisconsin Card Sorting Test. *Arch Gen Psychiatry* 1987; 44: 1008–14.
- Green MF, Nuechterlein KH, Gold JM, Barch DM, Cohen J, Essock S, et al. Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICES conference to select cognitive domains and test criteria. *Biol Psychiatry* 2004; 56: 301–7.
- Horga G, Abi-Dargham A. An integrative framework for perceptual disturbances in psychosis. *Nat Rev Neurosci* 2019; 20: 763–16.
- Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III—the final common pathway. *Schizophr Bull* 2009; 35: 549–62.
- Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci* 2016; 19: 404–13.
- Kaplan CM, Saha D, Molina JL, Hockeimer WD, Postell EM, Apud JA, et al. Estimating changing contexts in schizophrenia. *Brain* 2016; 139: 2082–95.
- Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 2003; 160: 13–23.
- Laws KR. A meta-analytic review of Wisconsin Card Sort studies in schizophrenia: general intellectual deficit in disguise? *Cognitive Neuropsychiatry* 1999; 4: 1–30. discussion 31–5.
- Leeson VC, Robbins TW, Matheson E, Hutton SB, Ron MA, Barnes TRE, et al. Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. *Biol Psychiatry* 2009; 66: 586–93.
- Li C-T, Lai W-S, Liu C-M, Hsu Y-F. Inferring reward prediction errors in patients with schizophrenia: a dynamic reward task for reinforcement learning. *Front Psychol* 2014; 5: 1282.
- McGuire JT, Nassar MR, Gold JJ, Kable JW. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 2014; 84: 870–81.
- Nassar MR, Bruckner R, Frank MJ. Statistical context dictates the relationship between feedback-related EEG signals and learning. *eLife* 2019; 8: e46975.
- Nassar MR, Bruckner R, Gold JJ, Li S-C, Heekeren HR, Eppinger B. Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nat Commun* 2016; 7: 11609.
- Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasley B, Gold JJ. Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat Neurosci* 2012; 15: 1040–6.
- Nassar MR, Wilson RC, Heasley B, Gold JJ. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 2010; 30: 12366–78.
- Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol Rep* 1962; 10: 799–812.
- Phillips LD, Edwards W. Conservatism in a simple probability inference task. *J Exp Psychol* 1966; 72: 346–54.
- Powers AR, Mathys C, Corlett PR. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* 2017; 357: 596–600.
- Reddy LF, Waltz JA, Green MF, Wynn JK, Horan WP. Probabilistic reversal learning in schizophrenia: stability of deficits and potential causal mechanisms. *Schizophr Bull* 2016; 42: 942–51.
- Stephan KE, Diaconescu AO, Iglesias S. Bayesian inference, dysconnectivity and neuromodulation in schizophrenia. *Brain* 2016; 139: 1874–6.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. *NeuroImage* 2009; 46: 1004–17.
- Sutton R, Barto A. Reinforcement learning: an introduction. Cambridge, MA: MIT Press; 1998.
- The Psychological Corporation. Wechsler test of adult reading. San Antonio, TX: Harcourt Assessment; 2001.
- Wechsler D. Abbreviated scale of intelligence. San Antonio, TX: Psychological Corporation 1999.
- Wiecki TV, Poland J, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. *Clin Psychol Sci* 2015; 3: 378–99.
- Wilson RC, Nassar MR, Gold JJ. Bayesian online learning of the hazard rate in change-point problems. *Neural Comput* 2010; 22: 2452–76.
- Yogev H, Sirota P, Gutman Y, Hadar U. Latent inhibition and overswitching in schizophrenia. *Schizophr Bull* 2004; 30: 713–26.