CrossMark

Motivational deficits in schizophrenia relate to abnormalities in cortical learning rate signals

D. Hernaus[1] · Z. Xu[2] · E. C. Brown[3] · R. Ruiz[1] · M. J. Frank[4] · J. M. Gold[1] · J. A. Waltz[1]

Abstract

Individuals from across the psychosis spectrum display impairments in reinforcement learning. In some individuals, these deficits may result from aberrations in reward prediction error (RPE) signaling, conveyed by dopaminergic projections to the ventral striatum (VS). However, there is mounting evidence that VS RPE signals are relatively intact in medicated people with schizophrenia (PSZ). We hypothesized that, in PSZ, reinforcement learning deficits often are not related to RPE signaling *per se* but rather their impact on learning and behavior (i.e., *learning rate modulation*), due to dysfunction in anterior cingulate and dorsomedial prefrontal cortex (dmPFC). Twenty-six PSZ and 23 healthy volunteers completed a probabilistic reinforcement learning paradigm with occasional, sudden, shifts in contingencies. Using computational modeling, we found evidence of an impairment in trial-wise learning rate modulation (α) in PSZ before and after a reinforcement contingency shift, expressed most in PSZ with more severe motivational deficits. In a subsample of 22 PSZ and 22 healthy volunteers, we found little evidence for between-group differences in VS RPE and dmPFC learning rate signals, as measured with fMRI. However, a follow-up psychophysiological interaction analysis revealed decreased dmPFC-VS connectivity concurrent with learning rate modulation, most prominently in individuals with the most severe motivational deficits. These findings point to an impairment in learning rate modulation in PSZ, leading to a reduced ability to adjust task behavior in response to unexpected outcomes. At the level of the brain, learning rate modulation deficits may be associated with decreased involvement of the dmPFC within a greater RL network.

Keywords Schizophrenia · Decision-making · Reinforcement learning · Functional magnetic resonance imaging · Striatum · Prefrontal cortex · Motivational deficits

✉ D. Hernaus
dennis.hernaus@maastrichtuniversiy.nl

1 Department of Psychiatry; Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, MD, USA

2 Duke University School of Medicine, Duke Neurobiology, Durham, NC, USA

3 Hotchkiss Brain Institute, The University of Calgary, Calgary, AB, Canada

4 Department of Cognitive, Linguistic & Psychological Sciences, Dept of Psychiatry and Human Behavior, Brown University, Providence, RI, USA

Introduction

Reinforcement learning (RL), the ability to learn from actions through reward and punishment, is an essential mechanism underlying optimal decision-making (Dayan & Berridge, 2014). A driving factor in learning from the environment is the *reward prediction error (RPE) signal*—the mismatch between outcome and expectation, thought to be signaled by midbrain dopaminergic neurons projecting to the ventral striatum (VS) (Schultz, Dayan, & Montague, 1997; Steinberg et al., 2013). There is now appreciable evidence suggesting that RL deficits exist in individuals from across the psychosis spectrum, ranging from adolescents and young adults at clinical high risk for psychotic illness to people with chronic, multi-episode schizophrenia (PSZ) (Barch et al., 2017; Waltz, Demro, et al., 2015b).

Although results have been mixed (Gold et al., 2012; Hartmann-Riemer et al., 2017; Reddy, Waltz, Green, Wynn,

& Horan, 2016), there are a number of studies to date that have reported an association between RL deficits and the severity of (both clinical and subclinical) negative symptoms, suggesting to some extent that changes in adaptively responding to environmental stimuli may play an important role in the onset of motivational deficits (Barch et al., 2017; Strauss, Waltz, & Gold, 2014). In some, but not all, individuals with psychotic symptoms, RL deficits can be linked to aberrant VS RPE signals at the neural level (Maia & Frank, 2017; Murray et al., 2008) and, potentially, to increased presynaptic dopamine function at the neurochemical level (Boehme et al., 2015). Especially in individuals suffering from motivational deficits, RL impairments may result from a decrease in anticipatory pleasure (Engel, Fritzsche, & Lincoln, 2013; Frost & Strauss, 2016), which depends on VS function and has been shown to be affected in PSZ (Radua et al., 2015).

Perhaps surprisingly, there is some evidence that VS RPE signals in medicated PSZ are relatively intact (Dowd, Frank, Collins, Gold, & Barch, 2016; Gradin et al., 2011; Waltz et al., 2010); an observation most recently replicated in the largest sample of patients in this literature to date (Culbreth, Westbrook, Xu, Barch, & Waltz, 2016). In addition, work from our lab has revealed that RL deficits in medicated PSZ are associated with specific impairments in representations of expected value to drive choice, putatively linked to functioning of orbitofrontal cortex (OFC), while RPE-based learning, linked to basal ganglia and dopamine, is less affected in PSZ (Collins et al., 2014; Gold et al., 2012). This latter observation aligns well with reports of normalized reward-related signals in striatum following antipsychotic medication administration (Nielsen et al., 2012).

Thus, although previous work has reported altered striatal RPE signals in the psychosis spectrum (Murray et al., 2008; Schlagenhauf et al., 2014), there also are multiple reports of (generally medicated) PSZ with RL deficits but intact striatal RPE signals (Culbreth et al., 2016; Waltz et al., 2010). Based on these observations, we hypothesized that a mechanism determining the *impact* of RPEs on learning and behavior, rather than signaling the RPE *per se*, might contribute to RL deficits in PSZ. In computational models of RL (Sutton & Barto, 1998), a parameter called learning rate (α) acts as the multiplier of the RPE (denoted as δ) to determine how much each RPE signal is weighted in the updating of representations of stimulus or action value or stimulus-response association strength. Higher learning rates are not always adaptive. In stationary, but probabilistic, environments it is useful to decrease learning rates to avoid being overly sensitive to spurious outcomes. When contingencies are changing or uncertain, however, higher learning rates are desirable (Behrens, Woolrich, Walton, & Rushworth, 2007; Franklin & Frank, 2015). Evidence suggests that learning rate is dynamic, varying as a function of the volatility of the learning environment, and that the circuits that underlie the modulation of learning rate incorporate several frontal cortical regions, including dorsal anterior cingulate cortex (dACC) (Behrens et al., 2007) and dorsomedial prefrontal cortex (dmPFC) (Krugel, Biele, Mohr, Li, & Heekeren, 2009; McGuire, Nassar, Gold, & Kable, 2014), brain regions also implicated in learning from outcomes (Mars et al., 2005), expectation updating (Behrens et al., 2007; McGuire et al., 2014), and rapid, flexible decision-making (Krugel et al., 2009).

In medicated PSZ, we have previously reported abnormal outcome-related signals in frontal cortex (Waltz et al., 2010) and impaired expected value representation (Gold et al., 2012), thought to depend on intact orbitofrontal cortex function (Metereau & Dreher, 2015). In both studies, these impairments scaled with the severity of motivational deficits. Multiple other groups have reported abnormal prefrontal activity (or fronto-parietal connectivity) in association with belief updating in uncertain environments (Kaplan et al., 2016; Koch et al., 2010; Paulus, Frank, Brown, & Braff, 2003). While these findings *suggest* that abnormalities in learning rate modulation may contribute to RL deficits in PSZ, even when RPE signals are intact, no systematic investigation of dynamic learning rate modulation has been conducted in this population.

To directly test this hypothesis, we administered a probabilistic RL paradigm to medicated PSZ and healthy volunteers (HV), which involved choosing from three decks of cards with different reinforcement rates. Using computational models of decision making, we quantified subjects' dynamic modulation of learning rate on trials surrounding shifts in reinforcement contingencies. Employing functional magnetic resonance imaging (fMRI), we also isolated RPE and learning rate modulation signals in the brain using model-based analyses, expecting intact VS but altered dmPFC signals in medicated PSZ relative to HV. Finally, if deficient learning rate modulation contributes to RL deficits in medicated PSZ, then this may be the consequence of decreased coupling between regions that signal and utilize RPEs. We therefore conducted a psycho-physiological interaction (PPI) analysis, investigating correlations between dmPFC and regions that are thought to signal RPEs, during stable and volatile phases of the task. For all these analyses, we expected to observe the greatest deficits in learning-rate-associated activity in PSZ with the most severe motivational deficits.

Methods

Participants

Twenty-seven participants meeting the diagnosis for schizophrenia or a schizoaffective disorder and 25 HV matched on age, gender, ethnicity, and parental education were recruited. All participants provided written, informed consent to

protocols approved by the Institutional Review Board of the University of Maryland School of Medicine (Protocol HP-00051996). We recruited PSZ through clinics at the Maryland Psychiatric Research Center or other nearby community mental health centers. The presence of a diagnosis of schizophrenia or a schizoaffective disorder in PSZ, as well as the absence of a clinical disorder in HV, was confirmed using the SCID-I (First, Spitzer, Gibbon, & Williams, 1997). The absence of an Axis II personality disorder in HV was confirmed using the SIDP-R (Pfohl, Blum, Zimmerman, & Stangl, 1989). All PSZ were on a stable antipsychotic medication regimen (no changes in medication dose or type in the 4 weeks leading up to study participation; antipsychotic medication details and haloperidol equivalents can be found in Table S1). Major exclusion criteria included: pregnancy, current illegal drug use (verified using a urine screen), history of substance dependence (SCID-I), a neurological disorder, and/or a medical condition affecting study participation (such as chronic, uncontrolled hypertension or diabetes). Participants were asked to abstain from alcohol 24 hours before the study session, which was verified by a breathalyzer. In order to avoid any potential effects of nicotine withdrawal, smokers were allowed to smoke before the study session.

Clinical and cognitive assessment

To investigate a potential effect of motivational deficit severity on learning rate modulation, we used the Scale for the Assessment of Negative Symptoms (SANS). SANS data were collected by a trained and experienced clinical research associate. Brief Psychiatric Rating Scale (BPRS) ratings were collected in the same session and the positive symptom factor (suspiciousness, hallucinations, unusual thought content, grandiosity) was used as a measure of positive symptom severity (McMahon et al., 2002). The same clinical research associate assessed general cognitive ability (intelligence quotient; IQ) using the Wechsler Abbreviated Scale of Intelligence (WASI-II) (Wechsler, 2011) and performance on the MATRICS Consensus Cognitive Battery (Nuechterlein et al., 2008). Antipsychotic regimen doses were converted to haloperidol equivalents according to Andreasen et al. (2010).

Reinforcement Learning Paradigm

We used a probabilistic RL task adapted from a previous study by Krugel et al. (2009) (Figure S1 for graphical overview). Participants selected one out of three card decks, identified by colors (black, red, and blue) using the index finger of their right hand (2,000 ms). After a brief inter-stimulus interval (pseudo-randomized to 2,000, 4,000, or 6,000 ms) and a card flip animation (~200 ms), participants were informed about the outcome of their choice (1,000 ms), which could be a win (+100 points) or loss (−50 points). Choices were

rewarded probabilistically, with a choice of the "optimal deck" leading to a 100-point gain on 90% of trials (and a loss of 50 points on 10% of trials). Choices of two nonoptimal decks led to 100-point gains on 50% and 10% of trials (and losses of 50 points on 50% and 90% of trials), respectively. Participants were instructed to try to identify the optimal deck (i.e., the one with the highest expected value) as quickly as possible; they also were informed that, occasionally, a new deck would become the optimal one. In fact, this occurred after subjects selected the optimal deck 8 times in a run of 9 trials (1 nonoptimal choice was allowed). The task consisted of 160 trials, subdivided into 4 runs of 40 trials (each run lasting ca. 7 minutes; total task duration just over 28 min). Deck locations only changed between runs, not within them. Prior to the task, participants underwent a training session where they received task instructions, completed a practice session, and were given the opportunity to ask questions. After the experiment, the total of amount of earned points was divided by 1,000 (total earnings for each group reported in Table S2).

Computational modeling

We modeled performance data using an algorithm optimized for sudden and dynamic shifts in reinforcement contingencies (Krugel et al., 2009; Sutton & Barto, 1998). Using the expected reward $q_i(t)$ for each deck, the probability of choosing the deck $p_t(t)$ was calculated according to the softmax function described in Eq. 1.

$$p_i(t) = \frac{e[\gamma \cdot q_i(t)]}{\sum_{j=1}^{n} e[\gamma \cdot q_j(t)]} \tag{1}$$

In accordance with previous work (Rescorla & Wagner, 1972; Sutton & Barto, 1998), the RPE $\delta_i(t)$ was operationalized as the difference between the expected reward $q_i(t)$ and the actual reward $r_i(t)$, as shown in Eq. 2.

$$\delta_i(t) = r_i(t) - q_i(t) \tag{2}$$

Reward expectation updating following an RPE can then be formulated as Equation 3, where learning rate, α, acts as the multiplier of the RPE. In other words, α determines how strongly the RPE is used to update expectations.

$$q_i(t) = q_i(t-1) + \alpha \cdot \delta_i(t-1) \tag{3}$$

Given the dynamic nature of the task, where reward expectation updating is especially likely to occur following a shift in reward contingencies, learning rates $\alpha(t)$, in addition to RPEs, $\delta(t)$, were estimated on a trial-by-trial basis, similar to Krugel et al. (2009). A useful heuristic is that to prevent overly updating to a single spurious RPE, a learner can keep track of the recent history of RPEs and increase the learning rate when errors in prediction are increasing across trials, and vice

versa when errors are decreasing. Thus, we first computed the absolute RPE values (degree of unsigned surprise) and accumulated these (Eq. 4) to represent the recent average prediction error:

$$|\delta(t)| = |\delta(t-1) \cdot (1-\alpha(1))| + \delta(t) \cdot \alpha(t) \qquad (4)$$

where $\alpha(1)$ is a parameter that determines the initial learning rate on trial 1. Next, the slope m of the absolute smoothed RPEs $|\delta_i(t)|$ was calculated to obtain a metric for whether the averaged errors are increasing or decreasing, normalized to the average RPE (so that it is independent of the scale):

$$m(t) = \frac{|\delta(t)| - |\delta(t-1)|}{[(|\delta(t)| + |\delta(t-1)|)/2]} \qquad (5)$$

Learning rate $\alpha(t)$ was fixed between 0 and 1, as shown in Eq. 6. Here, the β parameter estimates the effect of RPE slope m on learning rate α. Small β values are indicative of a dynamic learning rate, whereas with greater learning rates the dynamics are negligible [i.e., learning rates do not deviate much from $\alpha(1)$] (Krugel et al., 2009).

$$f(m) = \text{sign}(m) \cdot \left[1 - e^{(-(m/\beta))^2} \right] \qquad (6)$$

Finally, trial-by-trial estimates of α were calculated using function f, where $\alpha(t)$ increased and decreased as a function of a positive and negative slopes m, respectively (Eq. 7).

$$\begin{aligned} &\text{If } m > 0; \alpha(t) = \alpha(t-1) + f[m(t)] \cdot [1-\alpha(t-1)] \\ &\text{If } m < 0; \alpha(t) = \alpha(t-1) + f[m(t)] \cdot \alpha(t-1) \end{aligned} \qquad (7)$$

Thus, if the slope m is positive (errors are increasing), the learning rate is increased toward 1, whereas if it is negative, the learning rate is driven toward 0. Note that this model reduces to a constant learning rate model when β values are large, i.e., f(m) ~0. Parameter initialization bounds for α (0.01>, <0.9), β (>0.01), and γ (>0.001) were fixed in accordance with previous work (Krugel et al., 2009). Supplemental Text 1 contains a detailed description of the model selection procedure, posterior predictions, and a demonstration of the effect of different β values on simulated performance. Model fit and individual parameters for the winning model are reported in Table S3.

FMRI acquisition and pre-processing

Whole-brain functional EPI images were acquired on a 3T Siemens Trio scanner (Erlangen, Germany) while participants completed the RL task. We acquired 852 T2*-weighted images with the following parameters: n slices = 81; TR = 2 s; TE = 30 ms; FA = 90°; voxel size = 1.5 mm³; FOV = 22 x 22 cm; matrix size = 128 x 128. Additionally, we acquired T1-weighted structural images (MPRAGE) with standard parameters (n slices = 192; TR = 8.6 s; TE = 4 ms; FA = 20; and

voxel size = 1 mm³) for anatomical reference. To minimize head movement, foam padding was used.

Data were preprocessed and analyzed using the AFNI software package (Cox, 1996). Pre-processing steps consisted of co-registration of EPI and anatomical images, warping to Talairach space (using the 452 International Consortium for Brain Mapping template, which caused the images to be upsampled to 1.5-mm isotropic voxels), and smoothing with a 6-mm FWHM kernel. In keeping with other work (Shine et al., 2016), volumes with >0.5-mm displacement in any plane were excluded from analysis, and participants with >20% excluded volumes were excluded from analyses altogether. Three participants were excluded on this basis. Anatomical images were segmented using SPM 12's segmentation algorithm (default settings) and were inspected for fit. In a manner similar to previous work (Hernaus, Casales Santa, Offermann, & Van Amelsvoort, 2017), an average grey matter image of the entire sample was constructed, which was later used to inclusively mask grey matter in group-level analyses.

FMRI data analysis

Regressors of interest for the model-based fMRI GLM were outcome onset times, amplitude-modulated (in separate GLMs) by trial-by-trial estimates of RPE [$\delta(t)$] and learning rate [$\alpha(t)$] (see Section 2.4). Eight regressors of no interest (presentation times of stimuli evoking nonresponses, "too slow" feedback, and the 6 demeaned motion parameters) were additionally added. All regressors of interest and the "too slow" feedback regressor of no interest were boxcar functions of 2s. We used a linear regression model with ARMA(1,1) modeling of serial correlation.

The strength of learning rate-dependent functional connectivity between dmPFC and other brain areas was investigated using a PPI analysis (O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). We selected the 8 trials before and 8 trials following every reinforcement contingency shift for every participant and distinguished between "learning rate modulation trials" and "nonlearning rate modulation trials." Trials 1-4 pre-shift and trials 1-4 post-shift were labeled as "learning rate modulation trials," whereas trials 5-8 pre-shift and trials 5-8 post-shift were labeled as "nonlearning rate modulation trials." Learning rate modulation trials were believed to represent a volatile environment, in which participants were unlikely to be certain that they were sampling from the optimal deck and therefore needed to utilize the RPE signal to rapidly update their expectations (with a large learning rate). Conversely, nonlearning rate trials were believed to represent a stable environment, in which participants were likely aware that they were sampling from the best deck and, hence, could update expectations more gradually updating (with a smaller learning rate). Note that an 8-trial interval was the minimum amount of trials before a sudden shift in reinforcement

contingencies could occur. The GLM for the PPI analysis contained the psychological regressor (learning rate modulation trials = 1, nonlearning rate modulation trials = −1), the time-series extracted from the dmPFC ROI, the psychophysiological interaction, and eight regressors of no interest.

At the group level, one-sample t tests were conducted to identify significant activations in regions associated with regressors of interest in the entire sample, without respect to group; independent two-samples t tests were conducted to test for between-group differences. To estimate the minimum cluster size required to correct for multiple comparisons across the whole brain at a level of p(FWE-corrected) = 0.05, 10,000 Monte Carlo simulations of the probability distribution were run using the AFNI 3dClustSim command (-autocorrelation option), with a voxel-wise threshold of $p < 0.001$.

In addition to whole-brain analyses, unbiased ROIs were selected from the literature. As an area consistently associated with learning rate modulation (McGuire et al., 2014), we decided on a 1026-mm³ (304 functional voxels) region of interest in the dmPFC (Figure S3A). This dmPFC area (peak voxel x = 3, y = 8, z = 51) was identified in a conjunction analysis of surprise-, uncertainty-, and reward-driven learning, suggesting reliable involvement in learning rate modulation (McGuire et al., 2014). Additionally, this region showed a high degree of overlap with regions previously associated with outcome-driven action-selection (Mars et al., 2005). At a voxel level threshold of $p < 0.005$, activity in a large cluster encompassing the dmPFC (cluster size = 381 voxels, peak voxel coordinates x = 1, y = 13, z = 52) was observed in the entire sample, suggesting that dmPFC activity tracked trial-by-trial learning rate in our sample.

Given the well-established involvement of the VS in signaling RPEs, we additionally decided on a 1073-mm³ (318 voxels) region of nucleus accumbens, defined by the Desai Lab AFNI atlas (https://afni.nimh.nih.gov/pub/dist/doc/program_help/whereami.html; Figure S3B). Thus, our ROIs encompassed regions implicated in signalling RPEs, as well as regions thought to use RPEs to update expectations. For group-level analyses, mean beta values were extracted from the above-mentioned ROIs and converted to Z-scores. ROI activity thresholds were set to $p < 0.005$ (against 0) and group differences in activity were Bonferroni-corrected for the number of statistical ROI tests (n = 2).

Statistical analyses

Of the total sample of 52 participants, two HVs were excluded, because they never experienced a sudden shift in reinforcement contingencies; one PSZ produced no responses on approximately half of all trials (48.75%) and also was excluded. Thus, RL performance data from 26 PSZ and 23 HV were subjected to further analyses. Conform previous work (den Ouden et al., 2013; Waltz & Gold, 2007), these analyses included: number of reversals achieved, total amount of earnings, switching between decks on subsequent trials regardless of outcome, resampling a deck following a win (win-stay), switching decks following a loss (lose-shift), and the tendency to choose the previous-best deck following a reversal regardless of outcome (post-shift perseveration).

For fMRI analyses another five participants were excluded due to phase-instability artifacts associated with the acquisition protocol (n = 3), a technical error leading to an incomplete fMRI dataset (n = 1) and excessive motion (n = 1; see Section 2.5 for details). Thus, fMRI analyses were conducted for a sample of 22 HV and 22 PSZ.

In accordance with previous work from our group (Gold et al., 2012; Waltz, Brown, et al., 2015a), we used mean item scores from the SANS avolition subscale as measures of a motivational deficit severity, in the service of assessing potential relationships between motivational deficit severity and neural signals associated with learning rate modulation. In light of a bi-modal distribution for many of our SANS items, as well as previous work reporting the greatest performance impairments in individuals with high motivational deficits (Gold et al., 2012), we created high and low motivational deficit subgroups. The median score (in PSZ) on the SANS avolition subscale was 1.50, yielding a reasonable subgroup cutoff score (13 vs. 13 in the sample of all PSZ; 11 vs. 11 in the sample of all PSZ with usable fMRI data). We additionally investigated the effect of positive symptom severity. The median of the average BPRS positive symptom factor score was also 1.50, leading to a subgroup of 14 vs. 12 in the sample of all PSZ and 13 vs. 9 in the sample of all PSZ with usable fMRI data.

Results

Sample Demographics

Participant groups were matched on age, gender, race, and parental education level. There was a significant group difference in participant education, IQ score as measured by the WASI-II, and some subscales of the MATRICS Consensus Cognitive Battery performance (Table 1).

Reinforcement Learning performance

When comparing groups on all trials, HV and PSZ did not significantly differ in the number of stages achieved (Figure S4), total earnings, switching, win-stay, or lose-shift behavior (Table S2). However, there was evidence for subtle performance deficits in PSZ following reversals. On trials 2-8 following the first 5 stages (which 79.59% of all participants achieved), PSZ demonstrated increased post-shift perseveration. That is, following a sudden shift in contingencies, PSZ more often sampled from the deck that was optimal *before the*

Table 1 Sample demographics

	HV (n = 23)	PSZ (n = 26)	t/X^2	p
Age	34.92 (10.80)	39.61 (11.88)	-1.44	0.16
Gender [F, M]	[9, 14]	[7, 19]	0.83	0.36
Race				
African American, Caucasian, other	[8,14,1]	[8,16,2]	0.28	0.87
Education level (years)	15.39 (1.70)	13.19 (2.50)	3.55	<0.01
Maternal education level	15.00 (2.86)	13.62 (2.79)	1.69	0.10
Paternal education level	13.70 (3.25)	14.08 (3.01)	-0.42	0.67
WASI-II IQ score	114.04 (12.81)	103.93 (13.00)	2.75	<0.01
MATRICS domains*				
Processing speed	52.71 (9.77)	42.41 (9.43)	3.52	<0.01
Attention/vigilance	49.38 (9.63)	46.59 (11.80)	0.85	0.40
Working memory	50.67 (9.89)	42.32 (11.03)	2.61	0.01
Verbal learning	51.48 (8.84)	43.23 (10.14)	2.84	<0.01
Visual learning	44.48 (11.70)	42.59 (13.35)	0.49	0.63
Reasoning	50.43 (9.57)	49.14 (8.59)	0.47	0.64
Social cognition	51.71 (8.98)	41.05 (11.75)	3.34	<0.01
Smoking status [yes, no]	[6, 17]	[13, 13]	2.94	0.09
Antipsychotic medication				
Total haloperidol		11.70 (7.18)		
Clinical ratings				
BPRS positive		1.97 (1.02)		
BPRS negative		1.76 (.68)		
BPRS disorganization		1.18 (.31)		
BPRS total		32.58 (7.77)		
SANS avolition		1.86 (1.43)		
SANS anhedonia		2.44 (1.01)		
SANS total		23.92 (13.54)		

*MATRICS data were available for 21/23 HV and 22/26 PSZ

shift (t_{47} = −2.63, p = 0.01; Fig. 1a; Figure S5 for trial-by-trial overview), which was especially pronounced on early trials (trials 2-5; t_{47} = −2.26, p = 0.03; Fig. 1a). A between-group difference in post-shift perseveration was still present at trend level when comparing across all stages achieved, which included participants that performed particularly well (t_{47} = −1.9, p = 0.06). Increased post-shift perseveration in PSZ may have lead to insufficient sampling of other decks following a reversal, reflected by a decrease in win-stay choices on early trials (t_{47} = 2.19, p = 0.03; Fig. 1b). While lose-staying— that is, resampling a deck following a loss—on early trials after a reversal was numerically greater in PSZ (M = 10.26%, SD = 14.87) compared with HV (M = 5.46%, SD = 8.64), this did not reach significance (t_{47} = 1.34, p = 0.19).

Computational modeling parameters

Other than changes in trial-by-trial estimates of learning rate reported in Section 3.4, no significant group differences were observed in free parameters, including learning rate on the first trial (Table 2).

Learning rate modulation

Next, we extracted trial-by-trial learning rate estimates 8 trials before and after a sudden shift in reinforcement contingencies. A repeated measures ANOVA with group as between-subjects factor and trials 1-8 pre-shift as a within-subjects factor revealed a significant group difference in the slope of learning rate pre-shift ($F_{2,116}$ = 4.27, p = 0.01), as HV showed a greater pre-shift decline in learning rate than PSZ (Fig. 2a). A learning rate slope difference was also apparent when comparing the groups on all 16 time-points ($F_{6,257}$ = 2.74, p = 0.02; Fig. 2a), where learning rate in HV followed a significant (p < 0.001) quadratic trend. These results suggest that HV showed dynamic changes in learning rate modulation as a function of environmental stability, which was not present in PSZ.
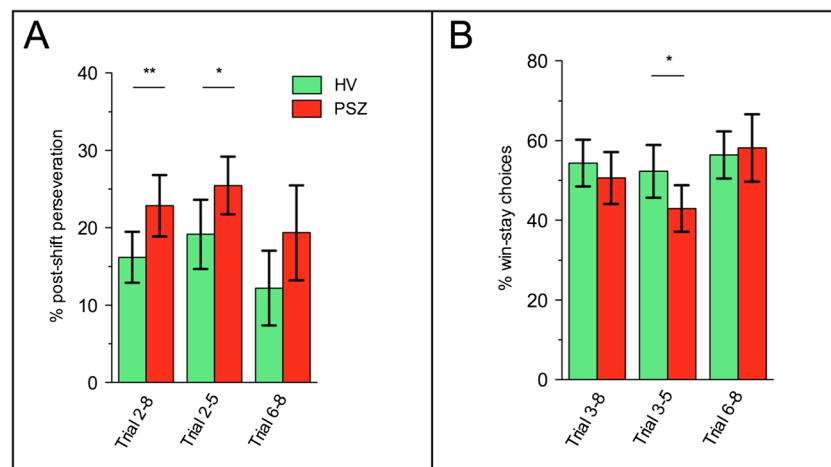
Fig. 1 Increased post-shift perseveration and decreased win-stay choices in PSZ. **a** Relative to HV, PSZ showed greater perseveration (i.e., selecting the previous-best card deck) following a sudden shift in reinforcement contingencies, especially in trials immediately following the shift. **b** Additionally, a decreased tendency to resample from the card deck that resulted in a win on the previous trial was observed immediately following the contingency shift. Note that post-shift perseveration was calculated across trials 2-8 post-shift; trial 2 being the first trial after participants have received feedback regarding a contingency shift. Win-stay choices were calculated for trial 3-8 post-shift; trial 3 being the earliest opportunity at which participants could have made a win-stay choice (following a shift to a new deck). Estimates were calculated on the basis of the first 5 stages. **p < 0.01, *p < 0.05, bars represent 95% confidence intervals

Interestingly, the group difference in pre-shift learning rate slope seemed to be driven by PSZ with the highest ratings for motivational deficits; a significant group-by-time interaction was observed when repeating the analysis with high and low avolition subgroups ($F_{5,114} = 2.73$, $p = 0.02$; Fig. 2b). Comparing individual slopes (overall group main effect $F_{2,46} = 4.44$, $p = 0.02$), we observed significant differences between HV and the high avolition subgroup ($p_{\text{Bonferroni-corrected}} = 0.02$) but not between HV and the low avolition subgroup ($p_{\text{Bonferroni-corrected}} = 0.37$). Pre-shift learning rate slope did not differ between low and high positive symptom subgroups ($p_{\text{Bonferroni-corrected}} = 0.99$).

We additionally observed an influence of avolition severity on post-shift learning rates. When we looked at the change in learning rate from trial 8 pre-shift, where participants have evidently identified the optimal deck, to mean learning rate on trials 10-17 post-shift, where participants need to locate the new optimal deck, we observed a main effect of participant group ($F_{2,46} = 4.58$, $p = 0.02$; Fig. 2c), such that HV showed greater pre- to post-shift learning rate changes than high-avolition PSZ ($p_{\text{Bonferroni-corrected}} = 0.03$), while low-avolition PSZ showed similar, yet nonsignificant, deficits compared with HV ($p_{\text{Bonferroni-corrected}} = 0.09$). No differences in post-

shift learning rates were observed between low and high positive symptom subgroups ($p_{\text{Bonferroni-corrected}} = 0.99$).

In summary, these results demonstrate that PSZ show deficits in the ability to update expectations dynamically through learning rate modulation, with this deficit being more pronounced in PSZ with more severe motivational deficits, but not in PSZ with greater positive symptoms.

Whole-brain analyses

As expected, trial-wise estimates of RPE were strongly associated with BOLD signal time-courses in VS (Fig. 3a; Table 3A) in the entire sample of participants with usable MRI data (n = 44). Trial-wise estimates of learning rate, on the other hand, were associated with BOLD signal time-courses in frontal and parietal cortex (Fig. 3b; Table 3B), in line with previous work demonstrating activity increases in the fronto-parietal network during decision-making tasks in volatile environments (Behrens et al., 2007; Koch et al., 2010; McGuire et al., 2014). At a voxel level threshold of $p < 0.005$, learning rate-related activity in dmPFC (cluster size: 381, peak voxel: x = 3, y = 8, z = 51) also was observed. Whole-brain analyses revealed no between-group differences in RPE- or LR-associated activity in any clusters large enough to survive correction for multiple comparisons across the whole brain.

Region of interest analyses

In the entire sample, VS RPE ($t_{43} = 5.75$, $p < 0.001$) and dmPFC learning rate signals ($t_{43} = 3.04$, $p = 0.004$) were

Table 2 No group differences in free parameters

Parameter	HV (n = 23)	PSZ (n = 26)	t	P p
$\alpha(1)$	0.68 (0.25)	0.64 (0.27)	0.52	0.61
β	3.77 (4.03)	3.70 (4.33)	0.06	0.95
γ	1.76 (0.69)	1.50 (0.50)	1.57	0.12
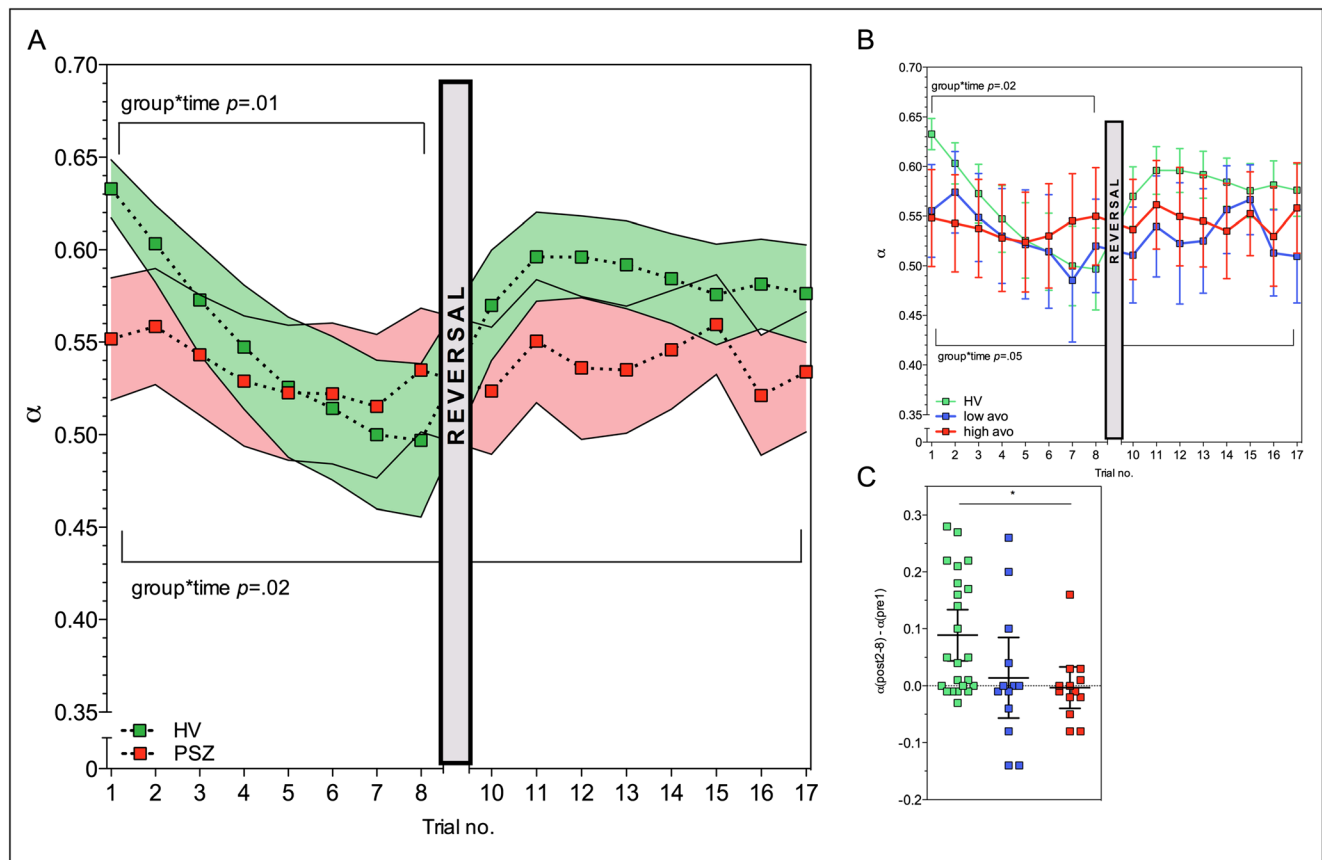
Fig. 2 Learning rate modulation deficits increase with motivational deficit severity. **a** PSZ, relative to HV, demonstrated a decrease in learning rate modulation, especially in trials leading up to a contingency shift. Solid bars represent SEM. *$p < 0.05$. **b** Decreased learning rate modulation on trials leading up to a contingency shift was especially apparent for PSZ with the most severe motivational deficits. These PSZ showed little to no learning rate modulation across all trials. **c** Relative to HV, PSZ with greater motivational deficits additionally demonstrated smaller increases in learning rate following a contingency shift, defined as the difference between trial 8 pre-shift and trial 2-8 ("11-17") post-shift

significantly greater than zero, consistent with our whole-brain analysis results. Moreover, greater dmPFC LR signals correlated positively with the percentage of optimal choices in the entire sample (Pearson's $r = 0.36$ $p = 0.02$, $n = 41$, 3 participants >2.5 SDs removed from mean). In line with previous work, VS RPE signals did not differ between groups ($t_{42} = -1.61$, $p = 0.12$; Fig. 3c). For model-based LR analyses, there was a trend for a group difference in dmPFC, with HV displaying a greater association between trial-wise estimates of learning rate and activity in dmPFC than PSZ ($t_{42} = 1.93$, $p = 0.06$; Fig. 3d). Despite the observed trend in dmPFC, these analyses do not provide clear evidence for impairments in regions associated with expectation updating or signaling of the RPE.

Learning rate modulation-dependent functional connectivity

Next, we conducted a PPI analyses to investigate learning rate modulation-dependent functional connectivity between dmPFC and the rest of the brain. In the entire sample, functional connectivity between dmPFC and superior parietal lobule increased from nonlearning rate-modulation to learning rate-modulation trials (Table 2C; Fig. 4a). In a follow-up ROI analysis, we observed an overall group difference in dmPFC-VS functional connectivity ($F_{2,46} = 3.74$, $p = 0.03$), which became highly significant after removing one HV outlier with a value smaller than 2.5*IQR ($F_{2,45} = 3.95$, $p = 0.008$). Specifically, we observed that HV and the low avolition subgroup displayed *increased* dmPFC-VS functional from nonlearning to learning rate-modulation trials. DMPFC-VS functional connectivity *decreased* in the high-avolition group (HV vs. high-avolition subgroup $p_{Bonferroni-corrected} = 0.01$; low- vs. high-avolition subgroup $p_{Bonferroni-corrected} = 0.02$; Fig. 4b). DMPFC-VS functional connectivity did not differ between low and high positive symptom subgroups $p_{Bonferroni-corrected} = 0.99$). Taken together, these results suggest decreased coupling between regions that *signal* the RPE and regions that *utilize* the RPE to update predictions. In line with our modeling results, connectivity changes were expressed most in PSZ with high motivational deficits.
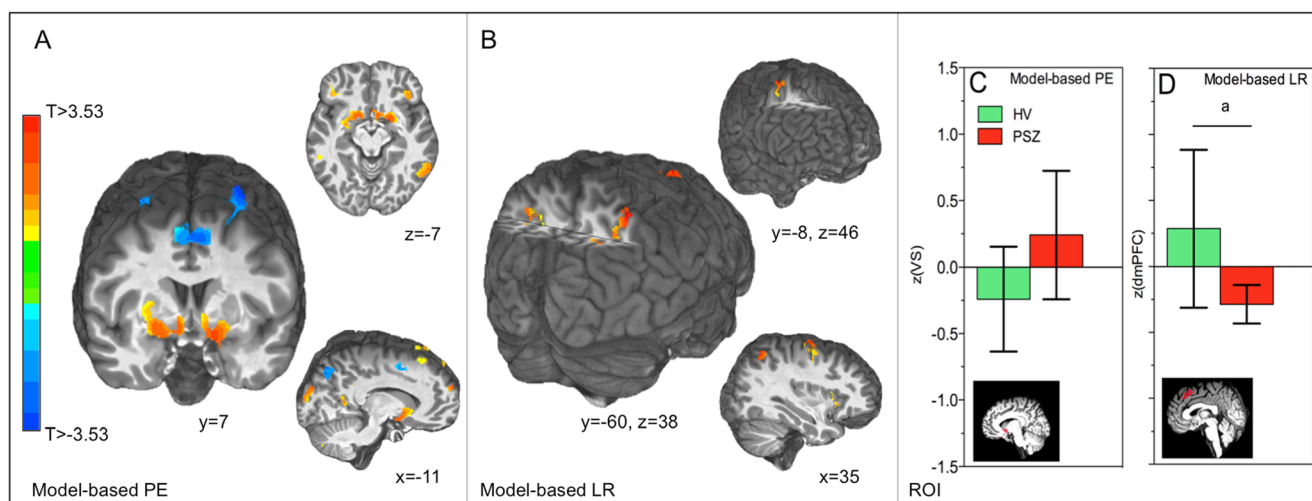
Fig. 3 Whole brain and ROI fMRI analyses. Robust RPE and LR signals were observed in: (**a**) VS, (**b**) superior parietal lobule, and multiple other brain regions (Table 3). In subsequent ROI analyses, (**c**) no group differences in VS RPE signals were observed, while (**d**) a trend for a group difference was observed for dmPFC learning rate signals. a = trend-significant ($p = 0.06$), bars represent 95% confidence intervals. PE = prediction error, LR = learning rate

Correlations with clinical variables

Total haloperidol-equivalent antipsychotic levels did not correlate with performance measures or fMRI analyses reported above (all $p < 0.46$). In the entire sample of participants with MATRICS data ($n = 43$), correlations between overall MATRICS scores and total numbers of stages achieved trended toward significance (Spearman's rho = 0.29, $p = 0.06$), but we observed no significant relationships between overall MATRICS scores and the slope of pre-reversal learning rate modulation (Spearman's rho = 0.01, $p = 0.98$). In the entire sample ($n = 49$), age correlated negatively with stages

Table 3. Whole-brain one-sample t tests

Region	Coordinates (MNI)					
	X	Y	Z	Direction	Cluster	Min. clust. size
Model-based PE (3A)						459
L inf. parietal lobule	-43	-38	41	<0	2052	
L mid. temporal gyrus	-53	-68	17	>0	1291	
L precentral gyrus	-29	-18	77	<0	1007	
L putamen	-14	5	-11	>0	904	
L cerebellum	-24	-75	-62	>0	888	
L pos. cingulate	-2	-50	14	>0	804	
R putamen	19	3	-11	>0	773	
L sup. frontal gyrus	-5	5	53	<0	639	
R precentral gyrus	30	-16	69	<0	589	
R inf. parietal lobule	36	-49	41	<0	562	
R inf. temporal gyrus	60	-57	-15	>0	497	
L inf. frontal gyrus	-49	30	15	>0	472	
Model-based LR (3B)						167
R sup. parietal lobule	39	-65	55	>0	306	
L sup. parietal lobule	-38	-68	48	>0	267	
R sup. frontal gyrus	32	-11	70	>0	218	
dmPFC PPI (3C)						222
R inf. Parietal lobule	48	-46	59	>0	340	

Coordinates are in MNI space; cluster size in voxels
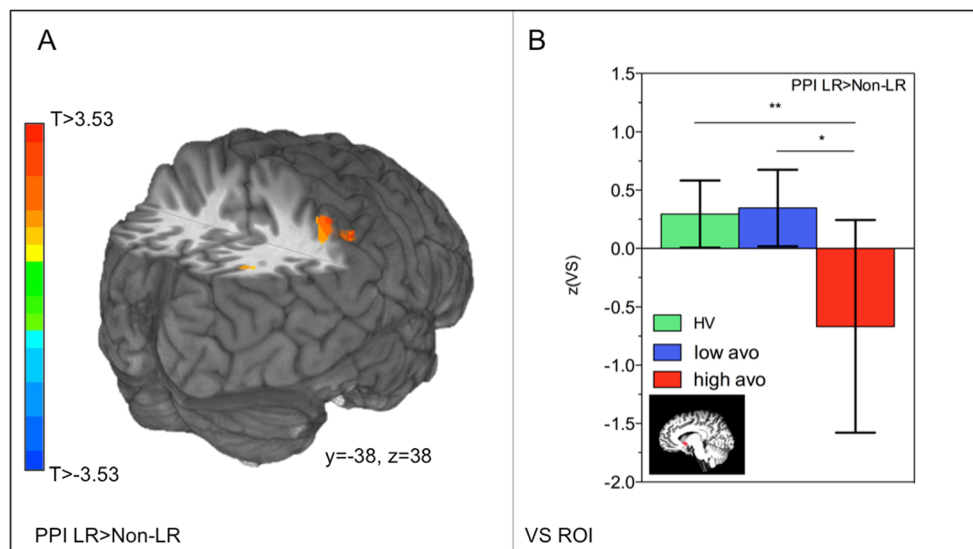
Fig. 4 Decreased learning rate modulation-related dmPFC-VS coupling in PSZ. **a** Whole-brain functional connectivity between dmPFC and inferior parietal lobule increased from nonlearning rate to learning rate trials in the entire sample. **b** In a follow-up ROI analysis in VS, functional connectivity increases were observed from nonlearning to learning-rate trials for HV and PSZ with more mild motivational deficits, while dmPFC-VS connectivity decreases were observed in PSZ with more severe motivational deficits. **$p < 0.01$, *$p < 0.05$, bars represent 95% confidence intervals. LR = learning rate

achieved (Spearman's rho $= -0.41$, $p = 0.003$) but not with the slope of pre-reversal learning rate modulation (Spearman's rho $= -0.18$, $p = 0.22$). Adding MATRICS score, smoking status, or age as a covariate did not change the results.

Discussion

In this report, we present evidence of an impairment in learning rate modulation in PSZ, leading to a reduced ability to adjust task behavior in response to unexpected outcomes. We moreover observed abnormal dmPFC-VS functional connectivity during changing task environments, potentially suggesting that the absence of learning rate dynamics in PSZ may be associated with a decrease in the integrity of the RL network. At both the behavioral and neural levels of enquiry, we observed greater deficits in PSZ with high motivational deficits, once again linking aspects of impaired reinforcement learning to negative symptom severity.

Schizophrenia has long been associated with increased perseveration (Lysaker, Bell, Bryson, & Kaplan, 1998) and a decreased ability to adaptively respond to performance feedback (Cicero, Martin, Becker, & Kerns, 2014; Mahurin, Velligan, & Miller, 1998). This inability to adaptively modulate behavior has previously been linked to increased reliance on response history when planning future decisions (Paulus, Geyer, & Braff, 1999), as well as to degraded representations of the expected value of choices (Gold et al., 2012; Waltz & Gold, 2016). Here, we expand on this work and offer a mechanism by which expected value could *become* degraded. That is, underutilization of accurately propagated teaching signals

may underlie choice perseveration, thereby hampering the formation of accurate and adaptive representations of expected value. Our observations of smaller *decreases in* learning rate modulation under relatively stable circumstances (preshift) points to a reduction in *dynamic* learning rate modulation, in PSZ, relative to HV. This idea is further supported by our observation of smaller *increases* in learning rate modulation under volatile circumstances (post-shift), in PSZ, relative to HV. Our subgroup-specific effects indicate that learning rate modulation deficits may be especially relevant in the endurance of motivational deficits, but not positive symptoms. Importantly, it should be noted that these analyses were conducted in subgroups of PSZ, and the specificity of symptom effects should therefore be interpreted with caution.

Consistent with the results of multiple previous studies (Culbreth et al., 2016; Dowd et al., 2016; Waltz et al., 2013; Waltz et al., 2010), we found that striatal RPE signals did not differ between HV and PSZ. Specifically, no significant group differences were observed in VS for a model-based analysis using trial-by-trial estimates of the RPE. As we have argued in the introduction, and in light of previous of work showing deficits in RPE coding (Murray et al., 2008; Reinen et al., 2016; Schlagenhauf et al., 2014), these results may suggest that RL impairments across the psychosis continuum are underlain by different mechanisms. Although we did not observe any associations between antipsychotic medication dose and performance, modeling or fMRI outcome measures in the current study, the possibility still exists that mechanisms of RL impairments may be dependent on illness phase (at-risk, first episode, or PSZ) or relate to the effects of antipsychotic medication on the reward system (Diederen et al., 2017; Nielsen

et al., 2012). To summarize, our results related to VS RPE signals could add to the notion that the basic machinery that signals unexpected outcomes is unaffected in certain groups of PSZ. Direct comparisons of different patient groups or a systematic investigation of antipsychotic medication effects will be essential for developing a better understanding of the different neural mechanisms that may underlie RL impairments in the psychosis continuum.

Of note, we found subtle evidence of between-group differences in dmPFC activity related to learning rate modulation in our ROI analyses. In our model-based analysis, we observed a trend for HV to show greater learning rate modulation signals than PSZ, which should be interpreted with caution in light of the sample size and correction for multiple comparisons. Observations of reductions in error- and response conflict-related activity in dACC (Culbreth et al., 2016; Dowd et al., 2016), overlapping with dmPFC, and altered outcome-related signals in ACC and inferior frontal gyrus (Kerns et al., 2005; Polli et al., 2008) have been observed before in PSZ. Moreover, a study by Koch et al. (2010) revealed that frontal cortical activity changes in response to environmental volatility were absent in medicated PSZ. Taken together, there is growing evidence for disrupted frontal cortical signaling in association with expectation updating in PSZ. Here we show, for the first time, that some of these signals may relate specifically to impaired learning rate modulation. Given that learning rate signals have been shown to be dependent on catechol-o-methyl transferase (COMT) genotype (Krugel et al., 2009) and affected by catecholamine enhancement (Jepma et al., 2016), these deficits may arise from changes in brain catecholamine function. Finally, Collins et al. have previously shown in HV (2017) and PSZ (2014) that RL learning rate is critically affected by working memory, suggesting that some of our observed finding may be explained by working memory deficits. Future studies should investigate how dynamic learning rate modulation is affected by working memory demands, such as load and delay.

The current results additionally provide preliminary support for the idea that a reduced ability to adaptively update expectations in PSZ could be related to decreased coupling between regions that signal and utilize the RPE to update behavior; the strength of learning rate modulation-dependent dmPFC-VS connectivity was decreased in PSZ, and especially so in PSZ with more severe motivational deficits. This finding echoes those of Kaplan et al. (2016) who observed changes in ACC effective connectivity within a change-detection network. In that study, increased PFC-midbrain effective connectivity tracked with delusional severity, whereas we observed decreased dmPFC-VS connectivity as a function of motivational deficit severity. Thus, changes in frontal cortical connectivity may relate closely to RL deficits, with this mechanism potentially being differentially affected by certain symptom dimensions. The latter claim is supported

by positive and negative symptom-specific alterations in cortico-striatal connectivity at rest (Sarpal et al., 2015; Wang et al., 2016; White et al., 2016).

To summarize, our current results suggest that aberrant learning rate modulation is a central feature of RL deficits in medicated PSZ, which can account for such deficits even in the presence of relatively unaffected RPE signaling. At the level of behavior, learning rate modulation impairments might manifest themselves as perseveration and/or a decreased tendency to sample the environment, thereby interfering with the formation of expected value. Functionally, we have provided initial support for the idea that learning rate modulation deficits might involve decreased embedding of the dmPFC within a greater RL network. Given its sensitivity to motivational deficits, learning rate modulation might serve as a mechanistic probe to evaluate the efficacy of future treatments aimed at alleviating negative symptoms of schizophrenia.

Some limitations to this study should be acknowledged. First, some of the reported fMRI results were only observed in ROIs or in whole-brain analyses at a lower voxel-wise threshold. Our ability to detect whole-brain group differences at a stringent whole-brain threshold may have been limited by our sample size and the use of patient subgroups (based on symptom severity). These results should therefore be considered preliminary, and replication is necessary. Moreover, while we found consistent evidence for learning rate modulation deficits at the neural (PPI) and behavioral (post-switch perseveration) in individuals with high negative symptoms, these measure were not directly correlated. This might suggest that our PPI phenotype may not underlie the observed behavioral deficits, although we would argue that this also strongly depends on our analysis approach. Specifically, our PPI analysis used all available experiment trials, while our modeling results pertained to highly specific trials surrounding contingency shifts.

Second, the ability to detect group differences in the neural signals of RL directly relates to how free parameters are calculated (Wilson & Niv, 2015). While model-based fMRI analyses of learning rate are robust to changes in parameter estimation it is not known exactly how this impacts estimates of RPE. Nevertheless, it is thought that poor model fitting should still be able to address questions related to the *location* of brain correlates of RL (and between-group differences therein).

Third, although previous studies have found PSZ to be impaired on reversal learning tasks (Schlagenhauf et al., 2014; Waltz & Gold, 2007), we only found subtle evidence for impairments. While our three-option task may be suitable to study learning rate dynamics, participants, on average, only achieved 6-7 stages, suggesting that on the majority of trials they were searching for the optimal deck. An increased number of decks to sample from may have made it more difficult for participants to achieve a reversal and, subsequently, reduced our ability to detect group differences in performance.

While the specific aim of this study was to disentangle RPE and LR signals in the brain, and how these may be altered in PSZ, there is a possibility that PSZ may show abnormal precision-weighted RPEs (the product of the LE and RPE), which has been shown to depend on dopamine function (Diederen et al., 2017). Future studies may wish to address how changes in learning rate dynamics in PSZ (with motivational deficits) may affect RPE-weighting.

Finally, this study was conducted in a sample of chronic, medicated PSZ. Although we did not find antipsychotic medication dose to be related to any of the reported outcome measures, direct comparison to a first-episode or prodromal cohort may be informative of the role of learning rate modulation deficits across the psychosis continuum.

Conclusions

RL deficits are common across the psychosis continuum. We show that aberrations in learning rate modulation in PSZ may drive RL deficits in some PSZ, even in the presence of accurately signaled RPEs. Our fMRI results additionally hint at the idea that learning rate modulation impairments relate to changes in dmPFC function and connectivity between dmPFC and other regions of the brain RL network. Furthermore, we observed greater abnormalities in learning rate modulation and associated neural signals in individuals with high motivational deficits. Abnormal learning rate modulation and associated brain function therefore might be an important mechanism involved in motivational deficits.

References

Andreasen, N. C., Pressler, M., Nopoulos, P., Miller, D., & Ho, B. C. (2010). Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs. *Biological Psychiatry, 67*(3), 255-262. https://doi.org/10.1016/j.biopsych.2009.08.040

Barch, D. M., Carter, C. S., Gold, J. M., Johnson, S. L., Kring, A. M., MacDonald, A. W., … Strauss, M. E. (2017). Explicit and Implicit Reinforcement Learning Across the Psychosis Spectrum. *Journal of Abnormal Psychology* https://doi.org/10.1037/abn0000259

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*(9), 1214-1221. https://doi.org/10.1038/nn1954

Boehme, R., Deserno, L., Gleich, T., Katthagen, T., Pankow, A., Behr, J., … Schlagenhauf, F. (2015). Aberrant Salience Is Related to Reduced Reinforcement Learning Signals and Elevated Dopamine Synthesis Capacity in Healthy Adults. *Journal of Neuroscience, 35*(28), 10103-10111. https://doi.org/10.1523/JNEUROSCI.0805-15.2015

Cicero, D. C., Martin, E. A., Becker, T. M., & Kerns, J. G. (2014). Reinforcement learning deficits in people with schizophrenia persist after extended trials. *Psychiatry Research, 220*(3), 760-764. https://doi.org/10.1016/j.psychres.2014.08.013

Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *Journal of Neuroscience, 34*(41), 13747-13756. https://doi.org/10.1523/JNEUROSCI.0989-14.2014

Collins, A. G. E., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working Memory Load Strengthens Reward Prediction Errors. *Journal of Neuroscience, 37*(16), 4332-4342. https://doi.org/10.1523/JNEUROSCI.2700-16.2017

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162-173. https://doi.org/10.1006/Cbmr.1996.0014

Culbreth, A. J., Westbrook, A., Xu, Z., Barch, D. M., & Waltz, J. A. (2016). Intact Ventral Striatal Prediction Error Signaling in Medicated Schizophrenia Patients. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(5), 474-483. https://doi.org/10.1016/j.bpsc.2016.07.007

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience, 14*(2), 473-492. https://doi.org/10.3758/s13415-014-0277-8

Diederen, K. M., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine Modulates Adaptive Prediction Error Coding in the Human Midbrain and Striatum. *Journal of Neuroscience, 37*(7), 1708-1720. https://doi.org/10.1523/JNEUROSCI.1979-16.2016

Dowd, E. C., Frank, M. J., Collins, A., Gold, J. M., & Barch, D. M. (2016). Probabilistic Reinforcement Learning in Patients With Schizophrenia: Relationships to Anhedonia and Avolition. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(5), 460-473. https://doi.org/10.1016/j.bpsc.2016.05.005

Engel, M., Fritzsche, A., & Lincoln, T. M. (2013). Anticipatory pleasure and approach motivation in schizophrenia-like negative symptoms. *Psychiatry Research, 210*(2), 422-426. https://doi.org/10.1016/j.psychres.2013.07.025

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). Structured Clinical Interview for DSM-IV- Axis I Disorders (SCID-I). Washington, DC: American Psychiatric Press.

Franklin, N. T., & Frank, M. J. (2015). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *Elife, 4*. https://doi.org/10.7554/eLife.12029

Frost, K. H., & Strauss, G. P. (2016). A Review of Anticipatory Pleasure in Schizophrenia. *Current Behavioral Neuroscience Reports, 3*(3), 232-247. https://doi.org/10.1007/s40473-016-0082-5

Gold, J. M., Waltz, J. A., Matveeva, T. M., Kasanova, Z., Strauss, G. P., Herbener, E. S., … Frank, M. J. (2012). Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Archives of General Psychiatry, 69*(2), 129-138. https://doi.org/10.1001/archgenpsychiatry.2011.1269

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., … Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain, 134*(Pt 6), 1751-1764. https://doi.org/10.1093/brain/awr059

Hartmann-Riemer, M. N., Aschenbrenner, S., Bossert, M., Westermann, C., Seifritz, E., Tobler, P. N., … Kaiser, S. (2017). Deficits in reinforcement learning but no link to apathy in patients with

schizophrenia (40352). *Scientific Reports, 7*. https://doi.org/10.1038/Srep44510

Hernaus, D., Casales Santa, M. M., Offermann, J. S., & Van Amelsvoort, T. (2017). Noradrenaline transporter blockade increases fronto-parietal functional connectivity relevant for working memory. *European Neuropsychopharmacology* https://doi.org/10.1016/j.euroneuro.2017.02.004

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. *PLoS Computational Biology, 12*(10), e1005171. https://doi.org/10.1371/journal.pcbi.1005171

Kaplan, C. M., Saha, D., Molina, J. L., Hockeimer, W. D., Postell, E. M., Apud, J. A., … Tan, H. Y. (2016). Estimating changing contexts in schizophrenia. *Brain, 139*(Pt 7), 2082-2095. https://doi.org/10.1093/brain/aww095

Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Johnson, M. K., Stenger, V. A., Aizenstein, H., & Carter, C. S. (2005). Decreased conflict- and error-related activity in the anterior cingulate cortex in subjects with schizophrenia. *The American Journal of Psychiatry, 162*(10), 1833-1839. https://doi.org/10.1176/appi.ajp.162.10.1833

Koch, K., Schachtzabel, C., Wagner, G., Schikora, J., Schultz, C., Reichenbach, J. R., … Schlosser, R. G. (2010). Altered activation in association with reward-related trial-and-error learning in patients with schizophrenia. *Neuroimage, 50*(1), 223-232. https://doi.org/10.1016/j.neuroimage.2009.12.031

Krugel, L. K., Biele, G., Mohr, P. N., Li, S. C., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of America, 106*(42), 17951-17956. https://doi.org/10.1073/pnas.0905191106

Lysaker, P. H., Bell, M. D., Bryson, G., & Kaplan, E. (1998). Neurocognitive function and insight in schizophrenia: support for an association with impairments in executive function but not with impairments in global function. *Acta Psychiatrica Scandinavica, 97*(4), 297-301.

Mahurin, R. K., Velligan, D. I., & Miller, A. L. (1998). Executive-frontal lobe cognitive dysfunction in schizophrenia: a symptom subtype analysis. *Psychiatry Research, 79*(2), 139-149.

Maia, T. V., & Frank, M. J. (2017). An Integrative Perspective on the Role of Dopamine in Schizophrenia. *Biological Psychiatry, 81*(1), 52-66. https://doi.org/10.1016/j.biopsych.2016.05.021

Mars, R. B., Coles, M. G., Grol, M. J., Holroyd, C. B., Nieuwenhuis, S., Hulstijn, W., & Toni, I. (2005). Neural dynamics of error processing in medial frontal cortex. *Neuroimage, 28*(4), 1007-1013. https://doi.org/10.1016/j.neuroimage.2005.06.041

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron, 84*(4), 870-881. https://doi.org/10.1016/j.neuron.2014.10.013

McMahon, R. P., Kelly, D. L., Kreyenbuhl, J., Kirkpatrick, B., Love, R. C., & Conley, R. R. (2002). Novel factor-based symptom scores in treatment resistant schizophrenia: implications for clinical trials. *Neuropsychopharmacology, 26*(4), 537-545. https://doi.org/10.1016/S0893-133X(01)00387-6

Metereau, E., & Dreher, J. C. (2015). The medial orbitofrontal cortex encodes a general unsigned value signal during anticipation of both appetitive and aversive events. *Cortex, 63*, 42-54. https://doi.org/10.1016/j.cortex.2014.08.012

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., … Fletcher, P. C. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry, 13*(3), 239, 267-276. https://doi.org/10.1038/sj.mp.4002058

Nielsen, M. O., Rostrup, E., Wulff, S., Bak, N., Broberg, B. V., Lublin, H., … Glenthoj, B. (2012). Improvement of brain reward

abnormalities by antipsychotic monotherapy in schizophrenia. *Archives of General Psychiatry, 69*(12), 1195-1204. https://doi.org/10.1001/archgenpsychiatry.2012.847

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., … Marder, S. R. (2008). The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *The American Journal of Psychiatry, 165*(2), 203-213. https://doi.org/10.1176/appi.ajp.2007.07010042

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience, 7*(5), 604-609. https://doi.org/10.1093/scan/nss055

den Ouden, H. E., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., … Cools, R. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron, 80*(4), 1090-1100. https://doi.org/10.1016/j.neuron.2013.08.030

Paulus, M. P., Frank, L., Brown, G. G., & Braff, D. L. (2003). Schizophrenia subjects show intact success-related neural activation but impaired uncertainty processing during decision-making. *Neuropsychopharmacology, 28*(4), 795-806. https://doi.org/10.1038/sj.npp.1300108

Paulus, M. P., Geyer, M. A., & Braff, D. L. (1999). Long-range correlations in choice sequences of schizophrenic patients. *Schizophrenia Research, 35*(1), 69-75.

Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. 1989. Structured Interview for DSM-III-R Personality Disorders (SIDP-R). Iowa City: University of Iowa, Department of Psychiatry.

Polli, F. E., Barton, J. J., Thakkar, K. N., Greve, D. N., Goff, D. C., Rauch, S. L., & Manoach, D. S. (2008). Reduced error-related activation in two anterior cingulate circuits is related to impaired performance in schizophrenia. *Brain, 131*(Pt 4), 971-986. https://doi.org/10.1093/brain/awm307

Radua, J., Schmidt, A., Borgwardt, S., Heinz, A., Schlagenhauf, F., McGuire, P., & Fusar-Poli, P. (2015). Ventral Striatal Activation During Reward Processing in Psychosis: A Neurofunctional Meta-Analysis. *JAMA Psychiatry, 72*(12), 1243-1251. https://doi.org/10.1001/jamapsychiatry.2015.2196

Reddy, L. F., Waltz, J. A., Green, M. F., Wynn, J. K., & Horan, W. P. (2016). Probabilistic Reversal Learning in Schizophrenia: Stability of Deficits and Potential Causal Mechanisms. *Schizophrenia Bulletin, 42*(4), 942-951. https://doi.org/10.1093/schbul/sbv226

Reinen, J. M., Van Snellenberg, J. X., Horga, G., Abi-Dargham, A., Daw, N. D., & Shohamy, D. (2016). Motivational Context Modulates Prediction Error Response in Schizophrenia. *Schizophrenia Bulletin, 42*(6), 1467-1475. https://doi.org/10.1093/schbul/sbw045

Rescorla, R. A., & Wagner, A. R. (1972). *in* Classical Conditioning II: Current Research and Theory, *eds Black AH, Prokasy WF*. New York City: Appleton–Century Crofts.

Sarpal, D. K., Robinson, D. G., Lencz, T., Argyelan, M., Ikuta, T., Karlsgodt, K., … Malhotra, A. K. (2015). Antipsychotic treatment and functional connectivity of the striatum in first-episode schizophrenia. *JAMA Psychiatry, 72*(1), 5-13. https://doi.org/10.1001/jamapsychiatry.2014.1734

Schlagenhauf, F., Huys, Q. J., Deserno, L., Rapp, M. A., Beck, A., Heinze, H. J., … Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage, 89*, 171-180. https://doi.org/10.1016/j.neuroimage.2013.11.034

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., … Poldrack, R. A. (2016). The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron, 92*(2), 544-554. https://doi.org/10.1016/j.neuron.2016.09.018

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience, 16*(7), 966-973. https://doi.org/10.1038/nn.3413

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

Sutton, R. S., & Barto, A. G. (1998). Reinforcement Learning: An Introduction. Cambridge: MIT Press.

Waltz, J. A., Brown, J. K., Gold, J. M., Ross, T. J., Salmeron, B. J., & Stein, E. A. (2015a). Probing the Dynamic Updating of Value in Schizophrenia Using a Sensory-Specific Satiety Paradigm. *Schizophrenia Bulletin, 41*(5), 1115-1122. https://doi.org/10.1093/schbul/sbv034

Waltz, J. A., Demro, C., Schiffman, J., Thompson, E., Kline, E., Reeves, G., … Gold, J. (2015b). Reinforcement Learning Performance and Risk for Psychosis in Youth. *The Journal of Nervous and Mental Disease, 203*(12), 919-926. https://doi.org/10.1097/NMD.0000000000000420

Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophrenia Research, 93*(1-3), 296-303. https://doi.org/10.1016/j.schres.2007.03.010

Waltz, J. A., & Gold, J. M. (2016). Motivational Deficits in Schizophrenia and the Representation of Expected Value. *Current Topics in Behavioral Neurosciences, 27*, 375-410. https://doi.org/10.1007/7854_2015_385

Waltz, J. A., Kasanova, Z., Ross, T. J., Salmeron, B. J., McMahon, R. P., Gold, J. M., & Stein, E. A. (2013). The roles of reward, default, and executive control networks in set-shifting impairments in schizophrenia. *PLoS One, 8*(2), e57257. https://doi.org/10.1371/journal.pone.0057257

Waltz, J. A., Schweitzer, J. B., Ross, T. J., Kurup, P. K., Salmeron, B. J., Rose, E. J., … Stein, E. A. (2010). Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology, 35*(12), 2427-2439. https://doi.org/10.1038/npp.2010.126

Wang, Y., Liu, W. H., Li, Z., Wei, X. H., Jiang, X. Q., Geng, F. L., … Chan, R. C. (2016). Altered corticostriatal functional connectivity in individuals with high social anhedonia. *Psychological Medicine, 46*(1), 125-135. https://doi.org/10.1017/S0033291715001592

Wechsler, D. (2011). Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). San Antonio: NCS Pearson.

White, T. P., Wigton, R., Joyce, D. W., Collier, T., Fornito, A., & Shergill, S. S. (2016). Dysfunctional Striatal Systems in Treatment-Resistant Schizophrenia. *Neuropsychopharmacology, 41*(5), 1274-1285. https://doi.org/10.1038/npp.2015.277

Wilson, R. C., & Niv, Y. (2015). Is Model Fitting Necessary for Model-Based fMRI? *PLoS Computational Biology, 11*(6), e1004237. https://doi.org/10.1371/journal.pcbi.1004237