Schizophrenia Bulletin vol. 46 no. 6 pp. 1558–1566, 2020
doi:10.1093/schbul/sbaa069
Advance Access publication 20 May 2020

Impairments in Probabilistic Prediction and Bayesian Learning Can Explain Reduced Neural Semantic Priming in Schizophrenia

Victoria Sharpe*[,1], Kirsten Weber[2,3], and Gina R. Kuperberg[1,4]

[1]Department of Psychology, Tufts University, Medford, MA; [2]Department of Neurobiology of Language, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [3]Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands; [4]Department of Psychiatry, Massachusetts General Hospital, Boston, MA

*To whom correspondence should be addressed; 490 Boston Ave. Medford, MA 02155; tel: +1 (617)-627-2198, fax: 617-627-3181, e-mail: victoria.sharpe@tufts.edu

It has been proposed that abnormalities in probabilistic prediction and dynamic belief updating explain the multiple features of schizophrenia. Here, we used electroencephalography (EEG) to ask whether these abnormalities can account for the well-established reduction in semantic priming observed in schizophrenia under nonautomatic conditions. We isolated predictive contributions to the neural semantic priming effect by manipulating the prime's predictive validity and minimizing retroactive semantic matching mechanisms. We additionally examined the link between prediction and learning using a Bayesian model that probed dynamic belief updating as participants adapted to the increase in predictive validity. We found that patients were less likely than healthy controls to use the prime to predictively facilitate semantic processing on the target, resulting in a reduced N400 effect. Moreover, the trial-by-trial output of our Bayesian computational model explained between-group differences in trial-by-trial N400 amplitudes as participants transitioned from conditions of lower to higher predictive validity. These findings suggest that, compared with healthy controls, people with schizophrenia are less able to mobilize predictive mechanisms to facilitate processing at the earliest stages of accessing the meanings of incoming words. This deficit may be linked to a failure to adapt to changes in the broader environment. This reciprocal relationship between impairments in probabilistic prediction and Bayesian learning/adaptation may drive a vicious cycle that maintains cognitive disturbances in schizophrenia.

Key words: language/N400/precision/statistical learning

Introduction

Prediction plays a crucial role in efficient, flexible cognition.[1,2] Sensory inputs that match our prior probabilistic predictions are easier to process than unpredictable inputs. Moreover, we are able to seamlessly *adapt* to changes in the statistical structures of our environments, *learning* from new inputs so that our predictions remain optimal.[3,4] It has been proposed that abnormalities in prediction can explain multiple features of schizophrenia, including positive and negative symptoms,[5–7] perceptual abnormalities,[8,9] impairments of proactive cognitive control,[10,11] and abnormalities of language comprehension and production.[12,13] Here we used event-related potentials (ERPs), a direct measure of neurocognitive processing, to show that abnormalities in probabilistic semantic prediction can account for the well-established reduction of the neural semantic priming effect in schizophrenia observed under non-automatic experimental conditions and that this, in turn, is linked to impaired learning/adaptation.

Semantic priming is a classic paradigm that can tease apart the different mechanisms by which we use long-term semantic knowledge, together with context, to facilitate semantic processing of incoming words. The semantic priming effect describes the facilitated processing of target words that are preceded by semantically associated prime words.[14,15] In the brain, semantic priming manifests as a reduction of the N400—a negative-going ERP component that peaks around 400 ms post-stimulus onset and indexes lexico-semantic processing.[16] Target words that are semantically associated with their primes elicit smaller N400 amplitudes than unrelated targets.[17,18]

In schizophrenia, both behavioral[19] and neural[20] semantic priming effects are reduced under experimental conditions that encourage controlled processing (in contrast with the preserved, or increased, semantic priming effect observed in some patients under more automatic conditions).[21–23] In the present study, we asked whether this reduced neural priming effect results from impairments in predictive processing.

Previous semantic priming studies in schizophrenia have been unable to address this question because they were carried out under conditions that encouraged not only predictive processes but also retroactive matching processes. For example, many behavioral[24–26] and ERP[27,28] studies used a lexical decision task, which encourages a retrospective evaluation of the semantic relationship between target and prime to bias the decision about whether the target is a word or nonword.[15,29,30] Moreover, in these previous behavioral[24–26,31,32] and ERP[22,27,28] studies, the proportion of directly associated word-pairs was less than 30% (sometimes because of the inclusion of nonwords), which discourages prediction.[15] One previous behavioral study used a high relatedness proportion (50%), but combined this with a pronunciation task, which does not require deep semantic processing.[33]

Here, we examined neural semantic priming in schizophrenia under experimental conditions that specifically probed prediction. We also examined the computational mechanisms underlying prediction and learning (the ability to adapt to the statistical structure of a new environment) in schizophrenia. To this end, we used a paradigm developed by Lau et al,[34] which manipulated the predictive validity of the prime, within-subjects, by varying the proportion of semantically associated prime-target pairs. Healthy participants saw a lower predictive validity block (10% associated pairs) and then a higher predictive validity block (50% associated pairs). Crucially, participants were not told that the proportion of associated trials would change halfway through. Throughout the experiment, participants monitored for animal words, which appeared in random filler trials. This task discouraged retrospective matching mechanisms (as there was no lexical decision to be made) but encouraged deep semantic processing. We found that healthy adults were able to take advantage of the increased predictive validity of the second block to generate stronger predictions, enhancing the neural semantic priming effect—a finding we replicated using magnetoencephalography[35] and functional magnetic resonance imaging.[36]

Because participants were not alerted to the change in predictive validity halfway through the experiment, these findings suggest that predictive semantic priming, even under non-automatic experimental conditions, can engage mechanisms that are more probabilistic and less strategic than had previously been assumed[37–39] (see supplementary material, section 1). This finding also highlighted the bidirectional relationship between probabilistic prediction and learning/adaptation. Specifically, to generate stronger predictions in the second block, participants needed to adapt to the change in the statistical structure of the environment—that is, they needed to dynamically update their estimate of the prime's predictive validity. In a follow-up study, we used a Bayesian model to formalize this link between prediction and learning/adaptation. We showed that in healthy adults, Bayesian principles could explain trial-by-trial variance in the N400 as participants adapted to the higher predictive validity of the second block.[40]

In the present study, we used this paradigm to test the hypothesis that people with schizophrenia would be less likely than healthy control participants to use the prime to predictively facilitate semantic processing of the target under conditions of higher predictive validity. That is, we hypothesized a significant Group by Relatedness interaction, driven by a smaller N400 semantic priming effect in the patients than in the controls. We also used our Bayesian adaptor model to explore the computational principles underlying abnormalities in probabilistic prediction and its relationship with learning/adaptation in schizophrenia.

Methods

Participants

Here, we report data from 18 outpatients with schizophrenia and 19 control participants (we excluded 5 additional datasets based on a priori exclusion criteria, see supplementary material, section 3). Patients were recruited from the Lindemann Mental Health Center in Boston, MA, and met DSM-IV criteria for schizophrenia or schizoaffective disorder (confirmed using the Structured Clinical Interview for DSM-III-R).[41] All participants gave written informed consent to participate, approved by the Massachusetts General Hospital Human Subjects Research Committee. Patients' symptoms were assessed using the Scale for the Assessment of Positive Symptoms[42] and the Scale for the Assessment of Negative Symptoms.[43] Premorbid verbal IQ was assessed using the North American Adult Reading Test.[44] All participants were right-handed,[45] monolingual English speakers, with normal/corrected-to-normal vision, no history of neurological impairment, and no substance abuse or dependence within 6 months. All patients were taking stable doses of antipsychotic medication (see supplementary material, section 3 for details).

The final schizophrenia and control groups were matched for age, gender, race, parental socioeconomic status,[46] and years of education (table 1). Premorbid verbal IQ was lower in the schizophrenia group than in the control group ($t(32.568) = 2.847$, $P < .05$). However, adding premorbid verbal IQ as a covariate in our analyses did not change the pattern of results (supplementary tables 4–6).

Stimuli and Task

Details of the design have been previously reported[34,48] and are described in supplementary material, section 2. Briefly, we crossed Relatedness and Predictive Validity in a 2 × 2 design. Relatedness was operationalized as

Forward Association Strength (FAS)[49] between the prime and the target. To manipulate the Predictive Validity, we added different numbers of associated word-pairs (FAS > .32) and unrelated word-pairs (FAS = 0) to the 2 blocks. In the lower predictive validity block, 10% of word-pairs (40/400) were associated; in the higher predictive validity block, which always followed the lower predictive validity block, 50% of word-pairs (200/400) were associated. The experiment was divided into 8 runs of 100 trials. Participants were allowed breaks between runs and were not told that the relatedness proportion would change. Participants' task was to press a button as quickly as possible upon seeing an animal word, which appeared as either prime or target in 80 unrelated filler trials in each block. These fillers were not included in any analyses. See figure 1 for the trial structure.

EEG Recording and Preprocessing

Electroencephalography (EEG) data were acquired using a 70-electrode cap (BrainProducts): sampling rate: 600 Hz; impedance: <10 k Ohms; bandpass: 0.1–30 Hz, referenced to the left mastoid. During preprocessing using EEGLab,[50] the EEG was segmented around the prime and target separately and artifact-corrected using Independent Components Analysis (infomax algorithm).[51] The remaining artifacts were rejected using standard algorithms. Slightly more trials were rejected in the patients than in controls (patients: 15.04%, controls: 10.00%, $F(1, 144) = 15.845$, $P < .05$), but this did not differ by Predictive Validity or Relatedness (Fs < 3.02, Ps > .08).

Table 1. Demographic information and clinical characterization of patients

	Control Group	Schizophrenia Group
N	19	18
Gender (M | F)	15 | 4	14 | 4
Race (C | AA | Other)[a]	9 | 9 | 1	15 | 3 | 0
Age	45.63 (6.20)	42.84 (9.21)
Parental SES[b]	2.84 (1.02)	2.78 (0.88)
Education (years)	12.58 (1.04)	12.78 (1.80)
Premorbid Verbal IQ[c]	109.68 (8.10)	100.52 (10.78)
CPZ Equivalent (mg)[d]	N/A	592.35 (300.70)
Duration of illness (years)	N/A	18.28 (8.99)
Age of illness onset (years)	N/A	23.78 (6.98)
SAPS[e]	N/A	3.44 (3.58)
SANS[f]	N/A	5.06 (3.81)

Note: Standard deviations in parentheses. [a]C, Caucasian; AA, African American; [b]Socioeconomic Status[46]; [c]NAART[44]; [d]Chlorpromazine Equivalents[47]; [e]summed global scores[43]; [f]summed global scores.[42]

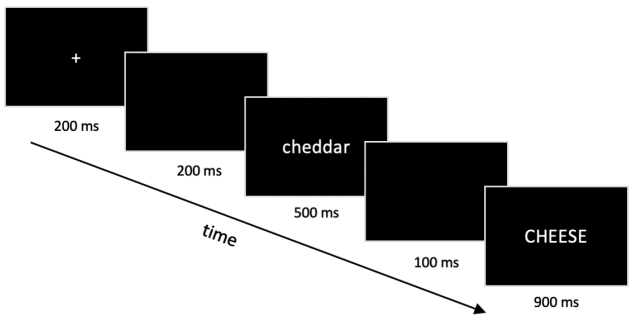

Fig. 1. Structure of each trial. Prime: "cheddar"; target: "cheese." See supplementary material, section 2 for details.

Analysis

Following our previous study using these materials,[34] we visualized the ERP grand-averaged plots using a matched, counterbalanced subset of stimuli (figure 2A). However, to maximize power, and to examine trial-by-trial adaptation, we carried out linear mixed-effects regressions (LMERs) on the N400 evoked by targets in *all* trials (except animal probe filler trials) using R[52] (*lme4 version 1.1-21*[53] and *lmerTest version 3.1–0*[54]). To explore nonspecific attentional effects, we also carried out analyses on the N400 evoked by primes. For all analyses, we operationalized the N400 as the average voltage across all sampling points between 300 and 500 ms, across 7 central-posterior electrodes (CP3, CP1, CP2, CP4, P1, PZ, and P2). This spatiotemporal region was selected a priori based on Lau et al.[34] Voltages were extracted for each trial using ERPLab.[55] Outliers were rejected using quantile trimming.

Predictors of interest were Group and Relatedness (FAS). "Nuisance" variables were log-transformed frequency,[57] orthographic length, concreteness,[58] semantic neighborhood size (number of unique word association responses[49]), and orthographic neighborhood size (Coltheart's N[59]). Continuous predictors were z-transformed. Significance was assessed using a type-III sums of squares estimation, with P-values estimated using the Satterthwaite approximation.[60] Random intercepts for items and subjects were included in all models, as were random slopes for all predictors of interest that varied by item or by subject (see supplementary material, section 5, for full model specifications).

Finally, we used our Bayesian model[40] to output a log-transformed probability of encountering each target in each participant in the higher predictive block (figure 3A). In this model, the probability of encountering an associated target is updated using Bayes' rule on each trial, assuming a beta-binomial distribution over associated/unrelated trials. The model then uses this probability estimate to weight prime-target FAS and target frequency to yield a probability estimate on each target, which is converted to the information-theoretic measure surprisal (see supplementary material, section 4, supplementary figure 1). The model output for a given trial is calculated as follows:

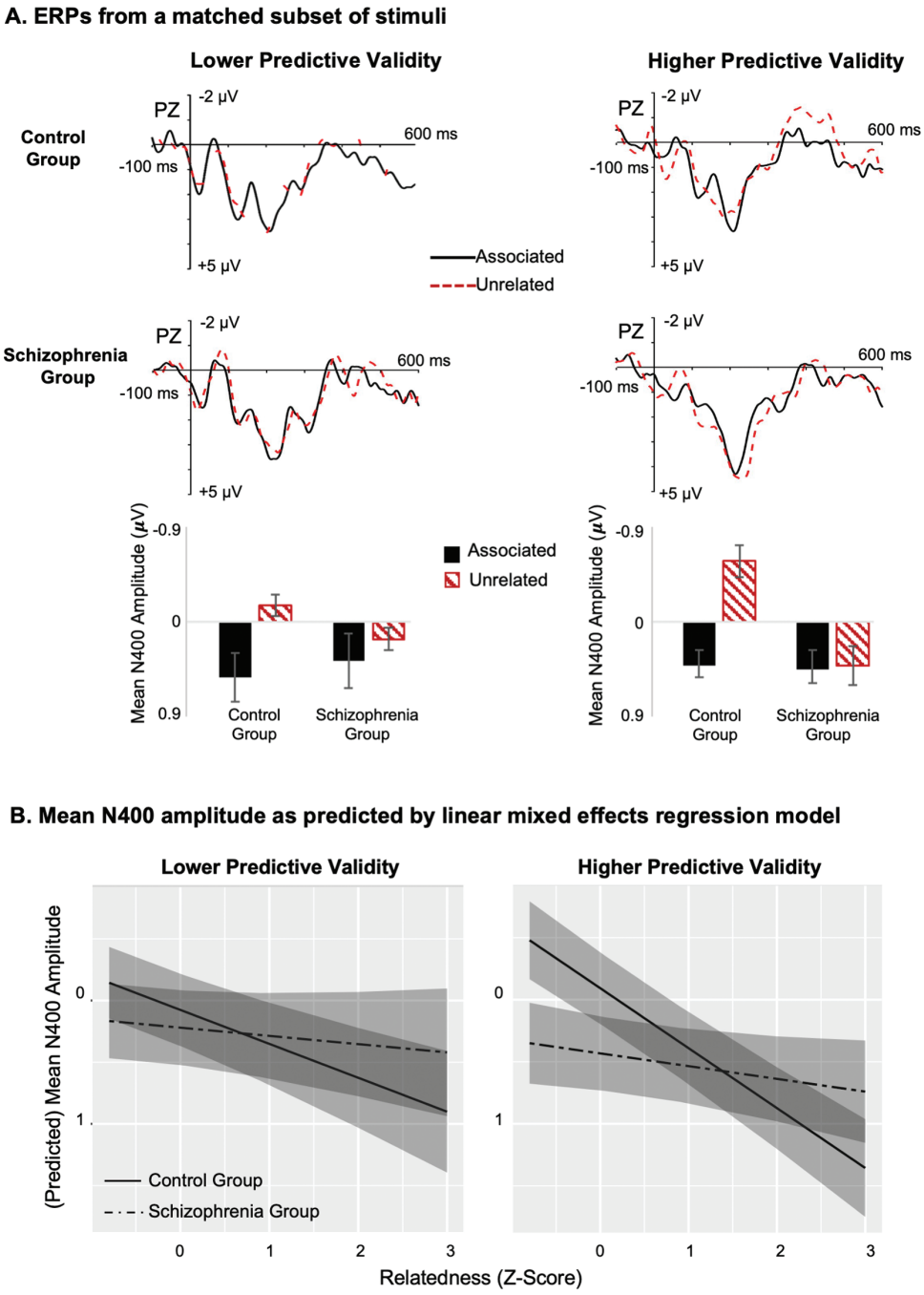

Fig. 2. (A) Grand-average ERPs, shown at Pz, time-locked to target word onset, from a subset of the data matched on lexical variables. (B) Predicted N400 amplitude as a function of Group, Relatedness (FAS[14]), and Predictive Validity. Using the effects package in R,[52,56] we used the coefficients from the LMER (tables 2A and 2B) to generate a "predicted" N400 over a range of FAS values, with nuisance variables held constant at their means. Gray ribbons represent one standard error. Negative is plotted up. ERP, event-related potentials; FAS, Forward Association Strength; LMER, linear mixed-effects regressions.

Model output = $-\log_2[\mu * p(\text{word}|\text{prime}) + (1-\mu) * p(\text{word}|\text{average context})]$, where μ is the expected probability of encountering an associated target.

To test our hypothesis that, compared with healthy controls, the schizophrenia group would be less likely to adapt to the higher predictive validity of the second block, we carried out another regression analysis in which we included the trial-by-trial output of this model as a predictor.

Results

The schizophrenia group was less accurate in detecting animal probes than the control group (controls: 81.31%; patients: 68.99%, $F(1, 102) = 6.595$, $P < .05$). However, both groups detected the majority of probes, with few false positives (controls $d' = 5.09$, SD = .74; patients $d' = 4.34$, SD = .92). There was no main effect of Predictive Validity

and no interaction between Predictive Validity and Group (Fs < 1.5, Ps > .2).

Confirming our first a priori hypothesis, we found a significant Group by Relatedness interaction in the higher predictive validity block (table 2A, figure 2B). Follow-ups showed that Relatedness significantly predicted N400 amplitude in controls (Est. = .487, t = 4.132, P < .05), but not in patients (Est. = .10, t = .795, P = .427). In the lower predictive validity block, there was only a main effect of Relatedness (table 2B).

In contrast to the N400 evoked by targets, there was no significant difference between groups in the amplitude of the N400 evoked by the primes in the higher predictive validity block (Est. = .45, P = .13), providing an important control for nonspecific attentional effects.

Finally, confirming our second hypothesis, in the higher predictive validity block, the trial-by-trial output of our computational model[40] interacted with Group (table 2C, figure 3B). Importantly, this interaction accounted for variance in trial-by-trial N400 amplitude *beyond* the variance accounted for by the interaction between Group and Relatedness (FAS). Follow-ups showed that trial-by-trial Model Output significantly predicted N400 amplitude in the controls (Est. = −.587, t = −2.618, P < .05) but not in patients (Est. = .408, t = 1.591, P = .112).

Discussion

We used EEG to show that people with chronic schizophrenia were impaired in their use of single word contexts to predictively facilitate neural processing of incoming words. While many previous studies have reported reduced behavioral and neural semantic priming effects in schizophrenia, our findings are the first to show (*a*) that this reduction is evident under experimental conditions that isolate *prediction* and (*b*) that it is linked to impairments in Bayesian trial-by-trial adaptation to changes in the statistics of the broader environment.

As expected, both patients and controls showed minimal semantic priming in the lower predictive validity block where there was a little utility in predicting the upcoming target based on the prime (predictions would have been incorrect on most trials). However, in the higher predictive validity block, where there was a substantial chance of generating a correct prediction about the target, people with schizophrenia showed a significantly smaller N400 semantic priming effect than the control participants.

One possibility is that the reduced predictive N400 effect in the schizophrenia group was driven by a failure to attend to the prime words or to engage in the task at all. To address this possibility, we compared the amplitude of the N400 evoked by the primes across the 2 groups in the higher predictive validity block. This analysis revealed no difference between patients and controls. Given that the amplitude of the N400 is known to decrease to non-attended words,[61] and when participants engage in shallow non-semantic processing,[62] this suggests that, like controls, patients attended to the meaning of the primes. Moreover, although patients' performance on the behavioral task was worse than controls, they performed well above chance, and their performance did not worsen across the 2 blocks. Thus, rather than reflecting a *general* disengagement from semantic processing, we suggest that people with schizophrenia were less likely than controls to *use* the prime to predict the meaning of the target.

Table 2. The results of LMERs Examining Modulation of the N400 Evoked by Target Words

A. Higher Predictive Validity Block: Group * Relatedness					
	Estimate (mV)	Std. Error	t-value	P-value	Sig.
Group	0.53	0.41	1.27	.21	
Relatedness	0.48	0.11	4.53	.00	***
Group*Relatedness	−0.38	0.15	−2.50	.02	*
B. Lower Predictive Validity Block: Group * Relatedness					
	Estimate (mV)	Std. Error	t-value	P-value	Sig.
Group	0.14	0.42	0.34	.73	
Relatedness	0.28	0.12	2.37	.02	*
Group*Relatedness	−0.21	0.16	−1.27	.21	
C. Effects of Bayesian Adaptor Model Output * Group (controlling for Relatedness * Group)					
	Estimate (mV)	Std. Error	t-value	P-value	Sig.
Group	0.12	0.44	0.28	.78	
Model Output	−0.57	0.22	−2.60	.01	**
Relatedness	−0.06	0.23	−0.27	.79	
Group*Model Output	0.93	0.31	3.02	.00	**
Group*Relatedness	0.52	0.33	1.55	.12	

Note: Fixed effects of predictors of interest are Shown. See supplementary tables 1–3 for the effects of nuisance variables. *.05 > P > .01, ** .01 ≥ P > .001, *** .001 ≥ P.

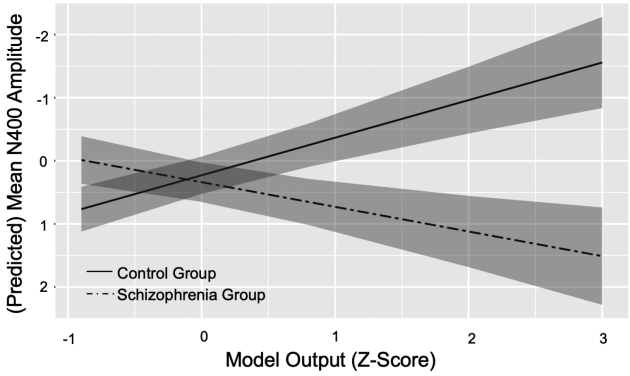

Fig. 3. Predicted N400 amplitude in the higher predictive validity block as a function of Group and Model Output. Using the *effects* package in R,[52,56] we used the coefficients from the LMER (table 2C) to generate a "predicted" N400 over a range of FAS values, with nuisance variables held constant at their means. Gray ribbons represent one standard error. Negative is plotted up. Note that Model Output indexes Bayesian surprisal (the unexpectedness of the target, given the prime); thus, greater surprisal elicits *larger* (more negative) N400 amplitudes in healthy controls. FAS, Forward Association Strength; LMER, linear mixed-effects regression.

This interpretation links the large semantic priming and N400 literatures to more general theories of impaired prediction in schizophrenia, which have thus far mainly focused on perception[5,8,63] and executive function.[10,11] It also has general implications for the interpretation of previous findings of abnormal N400 modulation in people with chronic schizophrenia. In healthy adults, there is a large body of evidence that the N400 during language comprehension is driven by probabilistic predictive mechanisms.[37,64] One possibility, therefore, is that reduced N400 effects seen in schizophrenia during sentence and discourse processing reflect impairments in proactively using the broader context to generate probabilistic semantic predictions about upcoming words (see supplementary material, section 1).

Our findings also highlight the link between probabilistic prediction and learning/adaptation. As healthy controls transitioned from conditions of lower to higher predictive validity, despite never being explicitly told that the statistical structure of the environment had changed, they implicitly *learned* that there was utility in using the prime to predict the target as they saw more and more associated pairs. Mathematically, we formalized this trial-by-trial learning using a dynamic Bayesian model.[40] We found that, within the higher predictive validity block, the trial-by-trial output of this model interacted with Group, and that this interaction explained more variance than the interaction between Group and Relatedness alone. This suggests that patients were less likely than controls to dynamically track their uncertainty about the statistical structure of the environment and use this uncertainty to (*a*) modulate their rate of learning and (*b*) *weight* the associative strength of the prime to generate predictions about the upcoming target, as specified by our model.[40]

These findings are in keeping with previous studies and modeling frameworks.[63,65,66] In many of these previous studies, however, learning was indexed by the changes in *behavior* in response to explicit trial-by-trial feedback,[65,66] and/or participants provided subjective confidence ratings after each trial.[67] An important feature of the present study, and our computational model, is that it indexed *implicit* trial-by-trial learning at a neural level, without any overt learning task. Moreover, because the N400 is itself a neural index of semantic probability, no explicit ratings of probability were required. This type of implicit learning is, of course, highly relevant to the ability to adapt to different statistical environments in the real world.[3]

The present study has several limitations. First, the patient group was limited to people with chronic schizophrenia who were taking medication, so we cannot generalize our findings to people with more recent onsets of schizophrenia, and we cannot separate out the effects of schizophrenia itself from the effects of medication. Our sample size was relatively small and so we did not have the power to assess the relationships with specific symptoms (although, based on previous studies of controlled semantic priming, we did not have specific hypotheses about such relationships, see supplementary material, section 6, for discussion and exploratory analyses).

Second, while we provide evidence that the schizophrenia group was less likely than the control group to use Bayes' Rule to dynamically adapt to the change in predictive validity, our computational model was unable to distinguish between 2 possible reasons for why this was the case. One possibility is that, at the beginning of the higher predictive validity block, patients did not expect the environment to change (holding an overly rigid prior that the environment was stationary). On this account, unexpected inputs, including those arising from a true change in the environment, were inappropriately attributed to noise, and so patients *downweighted* their influence on belief updating, leading to a reduced rate of learning. A second possibility is that patients had little faith in their prior model of the lower predictive validity block, expecting the environment to continuously change (an overly strong prior expectation of environmental volatility). On this account, unexpected inputs, including those arising from the inherent stochasticity of the environment, were inappropriately attributed to true change, leading the patients to *upweight* the influence of unpredicted input during belief updating. Although this would initially drive up the learning rate in the higher predictive validity block, it would explain why patients failed to converge on its correct statistical structure.[67]

To distinguish between these possibilities, it will be important to expand our computational model by incorporating hyperparameters that specify participants'

prior beliefs about environmental instability.[68,69] What the present set of findings do suggest is that patients were unable to distinguish between inputs that were unexpected because of their uncertainty about the statistical structure of the current environment (expected uncertainty) and inputs that were unexpected because of a true change in the environment (unexpected uncertainty).[70] These findings, therefore, extend previous proposals that impairments in inferring the *precision* of prediction error lead to abnormalities in perception and belief in schizophrenia[6,71] by raising the possibility that the same computational impairments might underlie impairments in learning.

The Bayesian model used in the present study was specified at Marr's first level of analysis[72] (see supplementary material, section 4). While this approach has the advantage of being able to explicitly specify the computational principles of the problem to be solved in probabilistic mathematical terms, it does not specify the algorithmic or neural mechanisms used to solve the computational problem. It will, therefore, be important for future studies to specify process-level inference algorithms to explain why patients failed to adapt[73–75] and to link these algorithms with precise neural mechanisms (see Yu and Cohen[76] for an example of work that bridges across Marr's computational, algorithmic, and neural levels of explanation).

In sum, our findings suggest that people with chronic schizophrenia are less likely than healthy participants to engage in *prediction* to facilitate lexico-semantic processing, resulting in reduced modulation of the N400 ERP component and that this may be linked to a failure to adapt to changes in the broader environment. Thus, impairments in prediction might drive impairments in learning, while impairments in learning might drive impairments in prediction, thereby perpetuating the perceptual and cognitive disturbances that characterize schizophrenia.

Supplementary Material

Supplementary material is available at *Schizophrenia Bulletin* online.

Funding

This work was funded by the National Institute of Mental Health (R01MH071635 to G.R.K.) and the National Institute of Child Health and Human Development (R01HD082527 to G.R.K.).

Acknowledgments

We are grateful to the Sidney R. Baer, Jr. Foundation for their support of students, including Sarah Armstrong, Ju Hyung Kim, Emily O'Carroll, and Gianna Wilkie, who contributed to recruitment and data collection. We thank Ellen Lau and Candida Ustine for their central roles in the development of the experimental paradigm and data collection, and Don Goff and the Freedom Trail Clinic for their support in patient recruitment. We also thank Nate Delaney-Busch, Emily Morgan, and Lin Wang for their technical guidance and insights.

References

1. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci.* 2013;36(3):181–204.
2. Griffiths TL, Kemp C, Tenenbaum JB. Bayesian models of cognition. In: Sun R, ed. *The Cambridge Handbook of Computational Psychology.* New York, NY: Cambridge University Press; 2008:59–100.
3. Courville AC, Daw ND, Touretzky DS. Bayesian theories of conditioning in a changing world. *Trends Cogn Sci.* 2006;10(7):294–300.
4. Qian T, Jaeger TF, Aslin RN. Learning to represent a multi-context environment: more than detecting changes. *Front Psychol.* 2012;3:228.
5. Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci.* 2009;10(1):48–58.
6. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. *Front Psychiatry.* 2013;4:47.
7. Sterzer P, Adams RA, Fletcher P, et al. The predictive coding account of psychosis. *Biol Psychiatry.* 2018;84(9):634–643.
8. Ford JM, Mathalon DH. Anticipating the future: automatic prediction failures in schizophrenia. *Int J Psychophysiol.* 2012;83(2):232–239.
9. Rudolph E, Ells E, Campbell D, et al. Finding the missing-stimulus mismatch negativity (MMN) in early psychosis: altered MMN to violations of an auditory gestalt. *Schizophr Res.* 2015;166(1–3):158–163.
10. Barch DM, Smith E. The cognitive neuroscience of working memory: relevance to CNTRICS and schizophrenia. *Biol Psychiatry.* 2008;64(1):11–17.
11. Barch DM, Ceaser A. Cognition in schizophrenia: core psychological and neural mechanisms. *Trends Cogn Sci.* 2012;16(1):27–34.
12. Brown M, Kuperberg GR. A hierarchical generative framework of language processing: linking language perception, interpretation, and production abnormalities in schizophrenia. *Front Hum Neurosci.* 2015;9:643.
13. Rabagliati H, Delaney-Busch N, Snedeker J, Kuperberg GR. Spared bottom-up but impaired top-down interactive effects during naturalistic language processing in schizophrenia: evidence from the visual world paradigm. *Psychol Med.* 2019;49(8):1335–1345.
14. Meyer DE, Schvaneveldt RW. Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *J Exp Psychol.* 1971;90(2):227–234.
15. Neely JH. Semantic priming effects in visual word recognition: a selective review of current findings and theories. In: Besner D, Humphreys GW, eds. *Basic Processes in Reading and Visual Word Recognition.* Hillsdale, NJ: Erlbaum; 1991:264–333.

16. Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu Rev Psychol.* 2011;62:621–647.
17. Bentin S, McCarthy G, Wood CC. Event-related potentials, lexical decision and semantic priming. *Electroencephalogr Clin Neurophysiol.* 1985;60(4):343–355.
18. Rugg MD. The effects of semantic priming and work repetition on event-related potentials. *Psychophysiology* 1985;22(6):642–647.
19. Pomarol-Clotet E, Oh TM, Laws KR, McKenna PJ. Semantic priming in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry.* 2008;192(2):92–97.
20. Kuperberg GR, Kreher DA, Ditman T. What can Event-related Potentials tell us about language, and perhaps even thought, in schizophrenia? *Int J Psychophysiol.* 2010;75(2):66–76.
21. Spitzer M, Braun U, Hermle L, Maier S. Associative semantic network dysfunction in thought-disordered schizophrenic patients: direct evidence from indirect semantic priming. *Biol Psychiatry.* 1993;34(12):864–877.
22. Kreher DA, Goff D, Kuperberg GR. Why all the confusion? Experimental task explains discrepant semantic priming effects in schizophrenia under "automatic" conditions: evidence from Event-Related Potentials. *Schizophr Res.* 2009;111(1-3):174–181.
23. Kuperberg GR, Weber K, Delaney-Busch N, et al. Multimodal neuroimaging evidence for looser lexico-semantic networks in schizophrenia: evidence from masked indirect semantic priming. *Neuropsychologia* 2019;124:337–349.
24. Koyama S, Hokama H, Miyatani M, Ogura C, Nageishi Y, Shimokochi M. ERPs in schizophrenic patients during word recognition task and reaction times. *Electroencephalogr Clin Neurophysiol.* 1994;92(6):546–554.
25. Henik A, Nissimov E, Priel B, Umansky R. Effects of cognitive load on semantic priming in patients with schizophrenia. *J Abnorm Psychol.* 1995;104(4):576–584.
26. Goldberg TE, Dodge M, Aloia M, Egan MF, Weinberger DR. Effects of neuroleptic medications on speech disorganization in schizophrenia: biasing associative networks towards meaning. *Psychol Med.* 2000;30(5):1123–1130.
27. Kiang M, Kutas M, Light GA, Braff DL. An event-related brain potential study of direct and indirect semantic priming in schizophrenia. *Am J Psychiatry.* 2008;165(1):74–81.
28. Condray R, Steinhauer SR, Cohen JD, van Kammen DP, Kasparek A. Modulation of language processing in schizophrenia: effects of context and haloperidol on the event-related potential. *Biol Psychiatry.* 1999;45(10):1336–1355.
29. Neely JH, Keefe DE. Semantic context effects on visual word processing: a hybrid prospective-retrospective processing theory. In: Bower GH, ed. *Psychology of Learning and Motivation: Advances in research and theory.* Vol 24. New York, NY: Academic Press; 1989:207–248.
30. Chwilla DJ, Hagoort P, Brown CM. The mechanism underlying backward priming in a lexical decision task: spreading activation versus semantic matching. *Q J Exp Psychol-A.* 1998;51(3):531–560.
31. Moritz S, Mersmann K, Kloss M, et al. Enhanced semantic priming in thought-disordered schizophrenic patients using a word pronunciation task. *Schizophr Res.* 2001;48(2–3):301–305.
32. Moritz S, Mersmann K, Kloss M, et al. 'Hyper-priming' in thought-disordered schizophrenic patients. *Psychol Med.* 2001;31(2):221–229.
33. Barch DM, Cohen JD, Servan-Schreiber D, et al. Semantic priming in schizophrenia: an examination of spreading activation using word pronunciation and multiple SOAs. *J Abnorm Psychol.* 1996;105(4):592–601.
34. Lau EF, Holcomb PJ, Kuperberg GR. Dissociating N400 effects of prediction from association in single-word contexts. *J Cogn Neurosci.* 2013;25(3):484–502.
35. Lau EF, Weber K, Gramfort A, Hamalainen MS, Kuperberg GR. Spatiotemporal signatures of lexico-semantic prediction. *Cereb Cortex.* 2016;26(4):1377–1387.
36. Weber K, Lau EF, Stillerman B, Kuperberg GR. The Yin and the Yang of prediction: an fMRI Study of semantic predictive processing. *PLoS One.* 2016;11(3):e0148637.
37. Kuperberg GR, Jaeger TF. What do we mean by prediction in language comprehension? *Lang Cogn Neurosci.* 2016;31(1):32–59.
38. Becker CA. Semantic context effects in visual word recognition: an analysis of semantic strategies. *Mem Cognit.* 1980;8(6):493–512.
39. Neely JH, Keefe DE, Ross KL. Semantic priming in the lexical decision task: roles of prospective prime-generated expectancies and retrospective semantic matching. *J Exp Psychol Learn Mem Cogn.* 1989;15(6):1003–1019.
40. Delaney-Busch N, Morgan E, Lau E, Kuperberg GR. Neural evidence for Bayesian trial-by-trial adaptation on the N400 during semantic priming. *Cognition* 2019;187:10–20.
41. Spitzer RL, Williams JB, Gibbon M, First MB. The structured clinical interview for DSM-III-R (SCID). I: history, rationale, and description. *Arch Gen Psychiatry.* 1992;49(8):624–629.
42. Andreasen NC. *Scale for the Assessment of Positive Symptoms (SAPS).* Iowa City, IA: The University of Iowa; 1984.
43. Andreasen NC. *Scale for the Assessment of Negative Symptoms (SANS).* Iowa City, IA: The University of Iowa; 1984.
44. Blair JR, Spreen O. Predicting premorbid IQ: a revision of the National Adult Reading Test. *Clin Neuropsychol.* 1989;3(2):129–136.
45. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9(1):97–113.
46. Hollingshead AB. *Two Factor Index of Social Position.* New Haven, CT: Yale University Press; 1965.
47. Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ. International consensus study of antipsychotic dosing. *Am J Psychiatry.* 2010;167(6):686–693.
48. Lau EF, Gramfort A, Hämäläinen MS, Kuperberg GR. Automatic semantic facilitation in anterior temporal cortex revealed through multimodal neuroimaging. *J Neurosci.* 2013;33(43):17174–17181.
49. Nelson DL, McEvoy CL, Schreiber TA. The University of South Florida free association, rhyme, and word fragment norms. *Behav Res Methods Instrum Comput.* 2004;36(3):402–407.
50. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods.* 2004;134(1):9–21.
51. Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 1995;7(6):1129–1159.
52. R Core Team:*A language and environment for statistical computing [computer program]. Version*. Vienna, Austria: R Foundation for Statistical Computing; 2016.
53. Bates DM, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J StatSoftw.* 2015;67(1):1–48.

54. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw.* 2017;82(13):1–26.
55. Lopez-Calderon J, Luck SJ. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci.* 2014;8:213.
56. Fox J. Effect displays in R for generalised linear models. *J Stat Softw.* 2003;8(15):1–27.
57. Brysbaert M, New B. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav Res Methods.* 2009;41:977–990.
58. Brysbaert M, Warriner AB, Kuperman V. Concreteness ratings for 40 thousand generally known English word lemmas. *Behav Res Methods.* 2014;46(3):904–911.
59. Balota DA, Yap MJ, Cortese MJ, et al. The English Lexicon Project. *Behav Res Methods.* 2007;39(3):445–459.
60. Satterthwaite FE. An approximate distribution of estimates of variance components. *Biometrics* 1946;2(6):110–114.
61. McCarthy G, Nobre AC. Modulation of semantic processing by spatial selective attention. *Electroencephalogr Clin Neurophysiol.* 1993;88(3):210–219.
62. Chwilla DJ, Brown CM, Hagoort P. The N400 as a function of the level of processing. *Psychophysiology* 1995;32(3):274–285.
63. Powers AR, Mathys C, Corlett PR. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* 2017;357(6351):596–600.
64. Kuperberg GR. Separate streams or probabilistic inference? What the N400 can tell us about the comprehension of events. *Lang Cogn Neurosci.* 2016;31(5):602–616.
65. Averbeck BB, Evans S, Chouhan V, Bristow E, Shergill SS. Probabilistic learning and inference in schizophrenia. *Schizophr Res.* 2011;127(1-3):115–122.
66. Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, Gold JM. Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol Psychiatry.* 2011;69(5):424–431.
67. Adams RA, Napier G, Roiser JP, Mathys C, Gilleen J. Attractor-like dynamics in belief updating in schizophrenia. *J Neurosci.* 2018;38(44):9471–9485.
68. Mathys CD, Lomakina EI, Daunizeau J, et al. Uncertainty in perception and the Hierarchical Gaussian Filter. *Front Hum Neurosci.* 2014;8:825.
69. Piray P, Daw ND. A simple model for learning in volatile environments. [published online ahead of print January 16, 2020] *bioRxiv* 2020. doi:10.1101/701466
70. Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. *Neuron* 2005;46(4):681–692.
71. Cassidy CM, Balsam PD, Weinstein JJ, et al. A perceptual inference mechanism for hallucinations linked to striatal dopamine. *Curr Biol.* 2018;28(4):503–514 e504.
72. Marr D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco, CA: W. H. Freeman; 1982.
73. Anderson JR, Schooler LJ. Reflections of the environment in memory. *Psychol Sci.* 1991;2(6):396–408.
74. Sanborn AN, Griffiths TL, Navarro DJ. Rational approximations to rational models: alternative algorithms for category learning. *Psychol Rev.* 2010;117(4):1144–1167.
75. Griffiths TL, Lieder F, Goodman ND. Rational use of cognitive resources: levels of analysis between the computational and the algorithmic. *Top Cogn Sci.* 2015;7(2):217–229.
76. Yu AJ, Cohen JD. *Sequential effects: Superstition or rational behavior? Paper presented at: Advances in Neural Information Processing Systems 21; NIPS 2008;* Vancouver, BC; 2008.