Reduced reward-related probability learning in schizophrenia patients

Alpaslan Yılmaz 1,2
Fatma Simsek 2
Ali Saffet Gonul 2,3
1 Department of Sport and Health, Physical Education and Sports College, Erciyes University, Kayseri, Turkey;
2 Department of Psychiatry, SoCAT Lab, Ege University School of Medicine, Bornova, Izmir, Turkey;
3 Department of Psychiatry and Behavioral Sciences, Mercer University School of Medicine, Macon, GA, USA

Abstract: Although it is known that individuals with schizophrenia demonstrate marked impairment in reinforcement learning, the details of this impairment are not known. The aim of this study was to test the hypothesis that reward-related probability learning is altered in schizophrenia patients. Twenty-five clinically stable schizophrenia patients and 25 age- and gender-matched controls participated in the study. A simple gambling paradigm was used in which five different cues were associated with different reward probabilities (50%, 67%, and 100%). Participants were asked to make their best guess about the reward probability of each cue. Compared with controls, patients had significant impairment in learning contingencies on the basis of reward-related feedback. The correlation analyses revealed that the impairment of patients partially correlated with the severity of negative symptoms as measured on the Positive and Negative Syndrome Scale but that it was not related to antipsychotic dose. In conclusion, the present study showed that the schizophrenia patients had impaired reward-based learning and that this was independent from their medication status.

Keywords: reinforcement learning, reward, punishment, motivation

Introduction

The world that we are living in is constantly changing and survival is based on learning both the reward and the penalty information associated with changing contingencies. During evolution, animals increased their ability to adapt their responses or actions to environmental stimuli depending on trial-to-trial feedback (reward or punishment) after each response or action. This cognitive process is called feedback-driven reinforcement learning, which leads to the development of actions to maximize the reward or minimize the penalty in the long run. Accumulated data suggest that reward-related reinforcement learning takes place in the corticostriatal network, which mainly consists of the basal ganglia, the dorsolateral prefrontal cortex, and the orbitofrontal cortex (reviewed in a recent paper by Fareri et al). Furthermore, reward prediction error, which is the process by which human beings become able to use past and current events to predict future events, is conveyed by midbrain dopamine neurons for new behavioral acquisition associated with contingency changes. In the face of any certain contingency, humans make the best prediction for that contingency based upon previous experiences. However, if the prediction is wrong (prediction error), then the knowledge base is updated so that future predictions are more accurate. Human imaging studies have revealed that the activity in the dopamine-rich ventral striatum and putamen is correlated with the reward prediction errors in classical conditioning tasks.
Schizophrenia is a severe mental disorder that affects nearly 1% of the population. Despite rigorous research efforts, the pathophysiology of the disease is only partially understood. One of the most popular theories on the pathophysiology of schizophrenia is dysfunctional dopamine transmission.\(^1\) Evidence supporting this theory includes the fact that pharmacological applications that intervene in dopamine transmission attenuate the positive symptoms. The caveat of this intervention is the possibility of increasing the negative symptoms and having only a limited benefit for the cognitive symptoms.\(^9\) Other supporting evidence comes from functional neuroimaging studies. Dopamine-based reward prediction error signal has been shown to diminish in schizophrenia patients, suggesting that patients fail to distinguish between relevant and irrelevant feedback signals.\(^10\) Further supporting evidence comes from clinical observations. Two negative symptoms, anhedonia and avolition, are closely related to dysfunctional reward processing, in that the reward is associated with the feeling of pleasure and any decrease in this feeling of pleasure would undermine motivation for goal-directed behavior, as the achievement of behavioral goals would result in an attenuated reward experience.\(^11,12\) Because reward function is a teaching signal about which stimuli and which responses are valued outcomes, a deficit in reward functioning may lead to a decrease in the learning of value of stimuli and action selection in schizophrenia patients. Furthermore, inability to follow the reward-associated contingency changes may contribute to the impairment in the decision-making processes in schizophrenia.\(^13\) It is proposed that feedback-driven reinforcement learning, which is related to the cortical dopamine system, is altered in schizophrenia patients.\(^14\) However, the results of studies are heterogeneous and inconclusive, with some studies showing impairment and the others showing no impairment, as reviewed by Gold et al.\(^15\) One reason for this may be the heterogeneity of patient groups, because it was shown that first-episode and chronic patients performed differently in the cognitive tests.\(^16,17\) The other reason may be the variability of the reinforcement tasks.\(^18\) The contingencies and the reward value of feedback, which might have had different neural correlates, may be a factor that confounds results.

The aim of this study was to test the hypothesis that reward-related probability learning is altered in schizophrenia patients. To test this hypothesis, we used a gambling paradigm as the reward-related probability-learning task, using Turkish lira as the reward currency. Delgado et al\(^19\) employed this same gambling paradigm; it has been shown that the task specifically activates frontostriatal systems where reinforcement-related learning takes place. We preferred to recruit stable outpatients for two reasons: (1) these patients had minimal positive symptoms, which otherwise might have interfered in their performance during the experiment, and (2) stable outpatients are candidates for outpatient treatment programs, which have an objective to improve negative and cognitive symptoms of schizophrenia. Therefore, understanding possible deficits in the reinforcement mechanisms of schizophrenia patients may help to develop better outpatient treatment programs.

**Methods**

**Participants**

Twenty-five schizophrenia patients and 25 healthy control subjects participated in the study. The schizophrenia patients were recruited from the outpatient clinic of the Department of Psychiatry, Ege University School of Medicine, only if they had been followed up for more than 1 year and were clinically stable for the last 6 months. The controls were recruited from local advertisements. To keep the distribution of education level and gender similar across the groups, the controls were selected nonrandomly from the applicants. The clinical diagnosis of schizophrenia was verified using the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (SCID), and a comprehensive review of medical records. Controls were screened with the nonpatient version of the SCID. Exclusion criteria for participants included head trauma with loss of consciousness, any unstable medical or neurological disease like diabetes mellitus, or hypertension. Furthermore, any axis I disorder other than schizophrenia in patients and any axis I disorder in controls were also exclusion criteria. Controls with a history of bipolar disorder or any psychotic disease in first-degree relatives were also excluded from the study.

The Institutional Review Board of Ege University School of Medicine approved the study. Each participant had a clear description of the study before giving his or her consent. During the informed consent procedure, a first-degree relative accompanied each patient. The clinical status of each patient was evaluated with the Positive and Negative Syndrome Scale (PANSS) and the Calgary Depression Scale.\(^20,21\) The extrapyramidal symptoms were assessed with the Extrapyramidal Symptom Rating Scale (ESRS). The experiment protocol began the following day. Because many of the patients had difficulty waking early, the experiment took place in the afternoon between 14:00 and 16:00.
All the patients except one were on atypical antipsychotics (clozapine, quetiapine, olanzapine, and risperidone; zuclopenthixol was the exception as a typical antipsychotic). Mean chlorpromazine-equivalent dosages are given in Table 1.

Before the task began, each participant scored his or her motivation and alertness on a scale from 1 to 10. Those who scored lower than 5 were excluded from the study.

**Task**

The gambling paradigm contained a series of 120 interleaved trials, with each trial lasting 11 seconds. The goal for each participant was to maximize his or her attained rewards by learning probabilities over time. Participants were instructed that during each 11-second trial, they would be shown a card with an unknown value ranging from 1 to 9. The participants were required to guess whether the value of the card was less than or greater than 5 (the card value could not be 5). Correct guesses led to positive feedback (a monetary reward) and incorrect guesses led to negative feedback (a monetary penalty). Each trial consisted of a probabilistic cue period followed by a feedback period. The probabilistic cue period began with one of five different cues (circle, diamond, square, star, and triangle) presented for 1.5 seconds on a 22-inch screen set 2 m away from the participants. Each cue represented the probability that the card to be shown in that particular trial had a value higher or lower than the number 5. The probabilities were 100% (high or low), 67% (high or low), and 50% (random). At the end of the probabilistic cue period, the card to be shown during that particular trial was presented, with a question mark on it, on the screen. Participants had to respond within 2.5 seconds by either a left- or a right-click of a computer mouse to indicate their selection. After the choice was made, the number appeared on the card and positive or negative feedback was given: either a happy cartoon face with an upward green arrow (indicating a gain) or a sad cartoon face with a downward red arrow (indicating a loss) was shown. Each correct guess added TL1.00 to a patient’s account and each incorrect guess lost TL0.50. The account total of each participant was always displayed on the left side of the screen, represented by a bar, with the height of the bar changing with each gain or loss made.

Before the trials, participants were instructed to pay attention to the cues presented at the beginning of the trials and that each of the five cues represented a different probability regarding the card value that was then shown. Participants were not told the contingencies before playing. Each cue was presented 24 times during the experiment for a total of 120 trials in 12 blocks. The cue order was random in each block and cues were not repeated consecutively. Participants were compensated according to their performance, although TL15.00 was guaranteed as a minimum for each participant.

Trials where a response was not made in time carried a monetary penalty of TL1.00. Repeated nonresponse or perseverative mouse click were exclusion criteria (none of the participants met these criteria).

**Statistical analyses**

Clinical, demographic variables were compared by Student’s t test or chi-square test, depending on the type of variable. The total monetary score of each group, which was used as a measure of the performance (total amount earned, taking into account rewards, punishments, and missed trials), was compared by Mann-Whitney U test because the variances between the groups were not similar. Each participant’s accuracy scores were obtained for each cue. The two cues with 100% probability were combined to form the 100% probability condition, and the cues with 67% probability were combined to form the 67% probability condition. A repeated measure of analysis of variance (ANOVA) model was used to compare the accuracy scores and reaction times for each cue. Pearson correlation coefficient analysis was preferred for testing the correlation between amounts of money earned or the accuracy rates and clinical variables. Statistical significance was set at $P = 0.05$.

### Table 1 Clinical and demographic variables of patients and controls

<table>
<thead>
<tr>
<th>Variable</th>
<th>Schizophrenia patients (n = 25)</th>
<th>Controls (n = 25)</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>30.1 ± 8.4</td>
<td>30.8 ± 8.4</td>
<td>t = 0.3; df = 48; $P &gt; 0.05$</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>13/12</td>
<td>14/11</td>
<td>$\chi^2 = 0.1; df = 1; P &gt; 0.05$</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.8 ± 2.8</td>
<td>14.2 ± 5.4</td>
<td>t = 1.2; df = 48; $P &gt; 0.05$</td>
</tr>
<tr>
<td>PANSS score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>12.3 ± 5.9</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Negative</td>
<td>17.7 ± 10</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>General psychopathology</td>
<td>28.7 ± 13.5</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Total</td>
<td>58.4 ± 27.3</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ESRS score</td>
<td>5.9 ± 6</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Calgary Depression Scale score</td>
<td>3.7 ± 4.9</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Antipsychotic dose$^*$</td>
<td>220.3 ± 129.3</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Total performance score$^{**}$</td>
<td>36.6 ± 17.4</td>
<td>53.8 ± 11.7</td>
<td>$t = 4.4; df = 48; P &lt; 0.001$</td>
</tr>
</tbody>
</table>

**Notes:**  
$^*$Chlorpromazine-equivalent dose (mg); $^{**}$amount of money earned in the gambling task (TL).

**Abbreviations:** $\chi^2$, chi-square; df, degree of freedom; ESRS, Extrapyramidal Symptom Rating Scale; NA, not applicable; PANSS, Positive and Negative Syndrome Scale.
Results

As Table 1 shows, patient and control groups were similar in age, gender distribution, and education level. Schizophrenia patients had minor extrapyramidal and depressive symptoms. When the total amount of money earned by the end of the experiment was compared between the two groups, the healthy controls had earned more money than the patients ($U = 132.5; P < 0.001$).

When the accuracy scores were compared according to cues with repeated measure ANOVA, we observed a significant group effect ($F = 16.8; $ df $= 1, 48; P < 0.001$). Furthermore, we observed a significant cue effect ($F = 62.6; $ df $= 2, 96; P < 0.001$). Multivariate ANOVA showed the schizophrenia patients had significantly lower accuracy rates than the healthy controls in cues with 67% probability ($F = 4.22; $ df $= 1, 48; P = 0.04$) and 100% probability ($F = 25; $ df $= 1, 48; P < 0.001$) (Figure 1).

Both groups achieved similar results for cues with 50% probability ($F = 1.34; $ df $= 1, 48; P > 0.05$).

Although the patients’ reaction times were shorter than reaction times of the controls, the difference between groups was not statistically significant ($F = 2.76; $ df $= 1, 48; P > 0.05$). However, there were effects of cue exposure on reaction times ($F = 4.46; $ df $= 1, 47; P = 0.013$) and cue and group interaction ($F = 6.26; $ df $= 1, 47; P = 0.004$). Schizophrenia patients responded more slowly on cues with 50% probability ($F = 4.8; $ df $= 1, 48; P = 0.03$) and showed a tendency to be slower on cues with 100% probability ($F = 3.5; $ df $= 1, 48; P = 0.07$) (Figure 1).

We observed a correlation close to statistical significance between the negative subscale of PANNS scores and the accuracy rate of the 100% probability condition ($r = -0.38; P = 0.06$). The other correlation analyses revealed that clinical symptoms (PANSS and its subscales, Calgary Depression Scale scores, and ESRS scores) and antipsychotic doses had no correlation with the amount of money earned or the accuracy scores (Table 2).

Discussion

The results of the present study showed that stable schizophrenia outpatients had impairment in reward-related probabilistic learning based on affective feedback. The patients had lower scores than the controls in contingencies associated with moderate uncertainty; however, interestingly, their scores were worse in responding following stimulus contingencies that could be fully predicted. Both groups had similar scores in stimulus where learning was impossible (50% probability). Low scores in the Calgary Depression

Table 2 Correlation matrix among clinical parameters with the total performance scores (amount of money earned in the gambling task) and the accuracy scores

<table>
<thead>
<tr>
<th>Clinical parameters</th>
<th>Total performance scores</th>
<th>Accuracy scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>50%</td>
</tr>
<tr>
<td>PANSS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>-0.042</td>
<td>-0.05</td>
</tr>
<tr>
<td>Negative</td>
<td>-0.03</td>
<td>0.17</td>
</tr>
<tr>
<td>General psychopathology</td>
<td>-0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>Total</td>
<td>-0.24</td>
<td>0.13</td>
</tr>
<tr>
<td>ESRS</td>
<td>-0.013</td>
<td>0.3</td>
</tr>
<tr>
<td>Calgary</td>
<td>-0.088</td>
<td>0.25</td>
</tr>
<tr>
<td>Depression Scale</td>
<td>-0.122</td>
<td>-0.13</td>
</tr>
<tr>
<td>Antipsychotic dose**</td>
<td>-0.122**</td>
<td>-0.13</td>
</tr>
</tbody>
</table>

*Notes: $^*P = 0.06; ^{*}Chlorpromazine-equivalent dose.*

Abbreviations: ESRS, Extrapyramidal Symptom Rating Scale; PANSS, Positive and Negative Syndrome Scale.
stimulus-response mapping. However, the results of studies, which require them to use feedback to learn one-to-one correspondence, also showed impairment on conditional associative learning tasks that did not require the integration of trial-by-trial feedback at all—rather, involving incidental learning through repeated stimulus exposures or repeated performance of specific routines—did not show substantial difference between schizophrenia patients and controls.\(^{25,31}\) Collected data from different studies suggest that schizophrenia patients have impaired performance on tests of rule learning that rely on hypothesis learning. One of the most studied examples of this finding is the reduction in the number of categories achieved by patients undergoing the Wisconsin Card Sorting Test.\(^{26}\) Schizophrenia patients have also shown impairment on conditional associative learning tests, which require them to use feedback to learn one-to-one stimulus-response mapping.\(^{27,28}\) However, the results of studies using unsupervised learning tasks that did not require the activity in the caudate nucleus during the trial-and-error learning in a previous study.\(^{19}\) Thus, our finding of reduced reward-related probability learning in schizophrenia patients}\(^{31}\) Reduced reward-related probability learning in schizophrenia patients

However, there were no rewards or penalties associated with feedback. This might have led to a lack of motivation to use the reinforcement to learn and guide future behavior, which makes a great deal of intuitive sense in both schizophrenia patients and controls. One way to overcome this problem is to use an affective feedback in the task such as money, as in the present study and others.\(^{35-37}\) Indeed, Knutson et al\(^{38}\) found that the activity in the caudate nucleus and ventral striatum correlated with the affective value of the anticipated monetary reward if subjects successfully hit a button during the display of a briefly presented target.

Adding feedback with affective value in the task may bring its own penalty in schizophrenia studies. Because it is proposed that anhedonia, as a symptom of the disease, may decrease the feeling of pleasure, then achieving behavioral goals does not result in the experience of pleasure, which provides little reason for vigorous pursuit of the goal. However, the majority of studies using different stimulus types and paradigms showed that schizophrenia patients, compared with healthy controls, report that they experience similar levels of valence and arousal to both negative and positive stimuli.\(^{39-41}\) Therefore, the results obtained from this study and similar studies cannot be explained by diminished reward value of the feedback (see Gold et al\(^{18}\) for further discussion on this).

We observed that the schizophrenia patients had the worst performance on the 100% probability condition (ie, the 100% stimulus-outcome contingency), whereas the difference between the groups in predicting the 67% probability condition (ie, the 67% stimulus-outcome contingency) was moderately significant. The finding of impaired performance of schizophrenia patients in full predictable conditions was in parallel with previous studies.\(^{36,42}\) As the 100% stimulus-outcome contingency is the condition in which feedback is entirely unambiguous, learning is most rapid and easy in healthy people.\(^{19,42}\) Thus, we propose that patients are less sensitive to reward and error information in full predictable conditions than in partly predictable conditions. This proposal is based on Morris et al\(^{42}\) study that examined error-related negativity (ERN), which is related to error monitoring in schizophrenia using a probability-learning task. They found that ERM amplitude reduction was more marked in the 100% probability condition than the 80% probability condition. In the present study, we consider the patients’ slower response to cues with 100% probability was because of an impaired evaluation of error monitoring.

In this study, a gambling paradigm was used that showed activity in the caudate nucleus during the trial-and-error learning in a previous study.\(^{19}\) Thus, our finding of reduced...
reward-related probability learning suggests schizophrenia patients have altered frontostriatal system. In a recent fMRI study, Koch et al\textsuperscript{36} used a task very similar to the one used in the present study and found that schizophrenia patients had significantly impaired reward-related trial-and-error learning associated with altered activation in frontoparietal networks in addition to frontostriatal systems.

In the present study, the finding of impaired implicit learning of repeated cues was contrary to previous studies that showed no impairment in implicit learning of schizophrenia patients.\textsuperscript{29–31} The different results among the studies might have been because of the task used. Adding an affective component to the task might have led to the “overload” of the frontostriatal system and other, compensating brain systems of the patients, resulting in failure of the implicit learning. Similar phenomena were described and shown in a study of the working memory of schizophrenia patients.\textsuperscript{44,45}

As negative symptoms like anhedonia and avolition are related to the reward system, it has been proposed that altered feedback learning may be related to the severity of negative symptoms. Indeed, patients who had prominent, primary negative symptoms had a negative correlation with the reinforcement-based learning.\textsuperscript{16,46} Primary negative symptoms should not be due to depression or anxiety, they should not be an antipsychotic side effect, and they should not be secondary to positive symptoms or psychosocial deprivation.\textsuperscript{47} In the present study, the result of a trend-level significance between negative symptom subscale scores on the PANSS and accuracy rates of cues with 100% probability, in which the patients showed the most impairment, supports the findings of previous studies. As a group, the patients had lower Calgary Depression Scale and ESRS scores and had minimal psychotic symptoms. Therefore, we consider that the PANSS negative symptom scores of the patients mostly reflected the primary negative symptoms of the disease.

One of the main limitations of the study came from a software error that prevented us from presenting the effect of time on learning in each group. Although one study has shown there was no interaction between time and group in PCL in schizophrenia patients and healthy controls,\textsuperscript{34} another study suggested learning rates were different between groups.\textsuperscript{36} Furthermore, it was shown that as the reinforcement learning progresses, neural correlates are changed in healthy controls and the caudate nucleus becomes less activated.\textsuperscript{19} The gambling test used in this study requires working memory for the best performance, and it has been shown that the working memory performance of schizophrenia patients is impaired. To reduce the memory load, the 120 trials were split into twelve blocks where each cue was presented twice. However, we are aware that this manipulation in the task was not enough to fully eliminate the effect of impaired working memory on the performance of the patients. Alternatively, it should be kept in mind that the gambling test is related to implicit learning and depends on frontostriatal rather than frontohippocampal systems.

The other limitation of the study was the medication status of patients. It is well known that antipsychotics reduce dopamine transmission, leading to functional and structural changes.\textsuperscript{24,48,49} In spite of this, we did not observe a negative correlation between antipsychotic dose and reward-related probability learning in the patients. This finding was in line with a previous study that included schizophrenia patients with similar clinical and medication status.\textsuperscript{17} In the present study, most of the patients were receiving second-generation antipsychotics, which have lower dopamine antagonist activity than first-generation antipsychotics. Furthermore, fMRI studies showed that reward-related activity in the ventral striatum is unaffected by atypical antipsychotic treatment.\textsuperscript{25,50} The patient group in the present study was recruited from schizophrenia patients with no other axis I disorder; this was done to eliminate the effects of comorbidity on the results. While applying the results of this study to the schizophrenia population, it should be kept in mind that comorbidity is a remarkably common situation in clinical practice. Although the patient and control groups had similar education status, this does not necessarily mean they had similar intelligence quotient (IQ) scores. Although implicit learning has been shown to be independent from IQ scores,\textsuperscript{51} missing IQ scores in the study population should be recognized as another limitation of this study.

**Conclusion**

The results of this study provide additional evidence for impaired reward-based learning in schizophrenia patients. Furthermore, the results showed the impairment was partially related to negative symptoms but was independent from the mood and medication status of patients.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


