# Dissociable mechanisms underpinning effort-cost decision-making across the psychosis spectrum

Alexis E. Whitton, Ph.D.[a,b], Jaisal Merchant, B.A.[c], & Kathryn E. Lewandowski, Ph.D.[b,c]

[a]Black Dog Institute, University of New South Wales, Sydney, NSW, Australia
[b]Department of Psychiatry, Harvard Medical School, Boston, MA, USA
[c]Schizophrenia and Bipolar Disorder Program, McLean Hospital, Belmont, MA, USA

**Corresponding Author:**

Alexis E. Whitton, Ph.D.
Black Dog Institute & University of New South Wales
Hospital Road
Randwick, New South Wales, Australia, 2052
Tel: +61 411 319 386
Fax: +61 2 9351 4195
Email: a.whitton@unsw.edu.au

**Abstract**

Recent theoretical models propose that abnormal effort-cost decision-making (ECDM) likely has divergent underpinnings across mood and psychotic disorders. However, whether this same model applies to individuals across the psychosis spectrum, including individuals with affective psychosis, remains unclear. This study aimed to empirically test whether two component processes – working memory and reward learning – contribute to ECDM impairment across the psychosis spectrum. ECDM was assessed using the Effort Expenditure for Rewards Task in individuals with psychotic disorders (n=190) and healthy controls (n=52). Working memory was assessed using a Digit Sequencing Task and reward learning was assessed using a Probabilistic Reward Task. Relative to the control group, the psychosis group showed reduced willingness to expend effort for higher probability, higher value rewards. This effect was most pronounced in individuals with schizophrenia and schizoaffective disorder relative to individuals with psychotic bipolar disorder. Across the whole sample, better working memory but not reward learning predicted greater willingness to expend effort for higher probability rewards. However, the link between working memory and ECDM differed as a function of patient symptom profile. Specifically, working memory was only predictive of ECDM for individuals with less severe negative symptoms and minimal depressive symptoms. For individuals with more severe negative symptoms, poorer ECDM was instead predicted by deficits in reward learning. Although these findings reiterate the important link between working memory and ECDM in individuals with psychotic disorders, they also show that this link varies in accordance with the presence of prominent negative and depressive symptoms.

**Keywords:** effort, reward learning, working memory, anhedonia, psychosis, depression

## 1.0. Introduction

Motivational deficits have long been documented in schizophrenia (Bleuler, 1950; Kraepelin, 1971), and are associated with poor outcomes and reduced quality of life (Fervaha et al., 2014; Foussias et al., 2011). More recently, studies have also highlighted the prominent role that motivational disturbances play in affective psychoses, such as schizoaffective disorder (Barch et al., 2017) and bipolar disorder (Pizzagalli et al., 2008a; Roiser et al., 2009; Whitton et al., 2015). Yet, despite the wide-reaching impact of motivational symptoms, we lack a comprehensive understanding of the affective and neurocognitive factors that contribute to motivational disturbances across the psychosis spectrum.

Deficits in effort-cost decision making (ECDM) have been identified as a mechanism that may underpin amotivation in psychosis. ECDM encompasses processes that support the ability to compute the value of an outcome against the effort required to obtain it – impairments in which have been well-documented in schizophrenia (for a review, see Culbreth et al., 2018). For example, on paradigms such as the Effort Expenditure for Rewards Task (EEfRT; Treadway et al., 2009), individuals with schizophrenia are less willing to expend physical effort to obtain larger, higher probability rewards relative to healthy controls (Barch et al., 2014; Fervaha et al., 2013; Huang et al., 2016; McCarthy et al., 2016; Treadway et al., 2015). Similar findings have been observed on tasks requiring cognitive effort (Culbreth et al., 2016), and in other psychiatric populations characterized by motivational impairments, such as mood disorders (Hershenberg et al., 2016; Treadway et al., 2012).

However, ECDM is a multicomponent process (Silveira et al., 2018) and an important question is whether shared or distinct mechanisms drive ECDM variation across different psychotic disorder symptom profiles. A recent review highlighted two important processes that

likely contribute to ECDM – reward processing and cognitive control (Culbreth et al., 2018). For example, reduced willingness to expend effort for rewards may arise from an inability to anticipate pleasure, reduced reward responsivity, or failure to use reward feedback to guide behavior. In contrast, reduced effort expenditure may arise due to cognitive deficits, such an inability to retain information about potential future rewards in working memory. Accordingly, examining the degree to which reward processing and working memory contribute to ECDM across different psychosis symptom profiles may highlight distinct factors that contribute to motivational impairment across different patients.

Although deficits in working memory and reward processing have been well-documented in psychosis, evidence suggests that these impairments may differ as a function of clinical presentation. For example, some studies show similar working memory impairments across different psychotic disorders (Gold et al., 2019; Lewandowski et al., 2013; Owoso et al., 2013), whereas others suggest greater deficits in schizophrenia relative to affective psychosis/bipolar disorder (Bora and Pantelis, 2015; Krabbendam et al., 2005; Sperry et al., 2015). In terms of reward processing, studies have shown evidence of intact reward responsivity (Dowd and Barch, 2012) and implicit reinforcement learning (Barch et al., 2017; Heerey et al., 2008) in schizophrenia, which contrasts with studies in individuals with bipolar disorder showing evidence of reduced reward responsivity in some cases (Pizzagalli et al., 2008a) but increased reward responsivity in others (Mason et al., 2016; Nusslock et al., 2012).

Evidence of differences in reward processing and working memory across different psychotic disorder symptom profiles raises the possibility that variation in ECDM may be driven by a number of different mechanisms across the psychosis spectrum. Some support for this hypothesis comes from a recent theoretical model proposed by Culbreth and colleagues (2018),

who suggest that ECDM deficits in primary psychosis (i.e., schizophrenia), may arise due to deficits in cognitive control (including working memory), whereas ECDM deficits in affective disorders (e.g., depression) may be more closely related to deficits in reward processing. Extending this model to different psychotic disorder symptom profiles, it is possible that abnormal ECDM in primary, non-affective psychosis may be driven by working memory deficits, whereas abnormal ECDM in individuals with more prominent co-occurring mood symptoms, may be more closely tied to reward processing deficits. However, no study to date has tested this hypothesis.

Identifying the processes that contribute to reduced effortful behavior is of central importance to understanding and treating motivational deficits in psychosis. Therefore, we aimed to evaluate associations between ECDM, reward processing, and working memory in a transdiagnostic psychosis sample and a sample of healthy controls. We hypothesized that 1) ECDM deficits would be evident in patients with psychosis compared to controls, and 2) ECDM performance would be associated with working memory in patients with more severe psychotic symptoms more generally, but would be associated with reward processing in patients with more prominent co-occurring depressive symptoms.

## 2.0. Materials and methods

### *2.1. Participants*

Patients with psychotic disorders (n=197) and healthy controls (n=53) were recruited from McLean Hospital's Psychotic Disorders Division. Inclusion criteria for patients was a lifetime DSM-IV psychotic disorder diagnosis, determined using medical chart review and the SCID-IV psychosis module (First et al., 2002). All patients were stable outpatients at the time of assessment and were permitted to engage in their regular treatment while taking part in the study.

The control group had no personal or first-degree family history of a psychiatric diagnosis, and no history of psychotropic medication. Exclusion criteria for all participants included head trauma with loss of consciousness and current substance abuse or dependence. All procedures were approved by the McLean Hospital Institutional Review Board.

Of those recruited, N=242 (52 controls, 190 patients) were deemed eligible. The patient group comprised individuals with: bipolar disorder with psychotic features (n=90), schizoaffective disorder (n=49), schizophrenia (n=42), major depressive disorder with psychotic features (n=5), psychotic disorder not otherwise specified (n=3), and schizophreniform disorder (n=1).

### *2.2. Measures*

#### *2.2.1. Clinical symptom scales*

Clinical symptom severity was assessed using the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987), the Young Mania Rating Scale (YMRS; Young et al., 1978), and the Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery and Åsberg, 1979). Community functioning was also assessed using the abbreviated Multnomah Community Ability Scale (MCAS; Barker et al., 1994; Lewandowski et al., 2013). Medication was recorded and chlorpromazine (CPZ) equivalents were calculated (Baldessarini, 2013).

#### *2.2.2. Effort-based decision-making (EEfRT)*

Effort-based decision-making was measured using the EEfRT (Treadway et al., 2009). On each trial of this 20-minute task, participants chose to perform an easy task (use their dominant index finger to press a key 30 times in 10 seconds) or a hard task (use their non-dominant pinky finger to press a key 100 times in 21 seconds). The reward for successful task completion ranged from $1 (easy task) to $1.12 - $4.12 (hard task). Participants were shown the

probability of receiving the money prior to making each choice (12%, 50% or 88%) and were told to try and win as much money as possible. Participants completed as many trials as they could in 20 minutes. They were informed that earnings from two trials would be randomly selected, representing their winnings.

*2.2.3. Reward learning (PRT)*

Reward learning was measured using the Probabilistic Reward Task (PRT; Pizzagalli et al., 2005). On each trial, a fixation cross appeared (500ms), followed by a schematic face with two eyes and a vertical line nose. After 500ms, a short (10mm) or a long (11mm) horizontal line mouth appeared (100ms). Participants indicated whether the short or long mouth was presented. There were 2 blocks of 100 trials, and 40 correct trials per block were rewarded ("Correct!! You won 20 cents"). Long and short mouths were presented at equal frequencies, however, unbeknownst to participants, one mouth (the "rich" stimulus) was rewarded three times more than the other (the "lean" stimulus). This asymmetrical reinforcement schedule induces a behavioral response bias toward the rich stimulus (Pizzagalli et al., 2005), and this bias is blunted in individuals with motivational deficits (Pizzagalli et al., 2008b).

*2.2.4. Working memory (DST)*

Working memory was assessed using the Digit Sequencing Task (DST) subscale of the Brief Assessment of Cognition in Schizophrenia (BACS) (Keefe, 1999). Participants were presented with randomly ordered number series and were required to repeat the numbers back from lowest to highest. Higher scores indicated better working memory.

***2.3. Data reduction and statistical analysis***

*2.3.1. EEfRT data reduction*

We calculated the number of hard choices made at each reward value and probability. Consistent with prior studies (Barch et al., 2014), choice data was entered into a 2 (*Group*: Control, Psychosis) x 3 (*Probability level:* 12%, 50%, 88%) x 4 (*Reward value:* <$1.80; $1.80-$2.55; $2.60-$3.45; $3.50-$4.20) repeated measures ANOVA. To further assess whether effort expenditure differed as a function of diagnosis, we supplemented this main analysis with a 4 (Diagnosis: Control, Bipolar Disorder, Schizoaffective Disorder, Schizophrenia) x 3 (Probability) x 4 (Value) ANOVA.

*2.3.2. PRT data reduction*

PRT data were subject to a quality assessment (Pizzagalli et al., 2008b; Whitton et al., 2016) described in the *Supplement*. Next, signal detection analysis was used to calculate response bias (the tendency to bias responding to the rich stimulus), using the formula:

$$\log b = \frac{1}{2} \log\left(\frac{Rich_{correct} * Lean_{incorrect}}{Rich_{incorrect} * Lean_{correct}}\right)$$

To allow for log-transformation of cells containing a zero, 0.5 was added to every cell of the matrix (Hautus, 1995).

Group differences in response bias were examined using a 2 (*Group*: Control, Psychosis) x 2 (*Block*) repeated measures ANOVA. Consistent with prior studies examining clinical and neurobiological correlates of reward learning (Kaiser et al., 2018; Pizzagalli et al., 2008b; Pizzagalli et al., 2005; Santesso et al., 2008), regression analyses used the reward learning metric, defined as block 2 – block 1 response bias.

*2.3.3. DST data reduction*

Scores were converted to a z-score based on normed scoring procedures and group differences were assessed using an independent samples *t*-test.

*2.3.4. Examining predictors and moderators of ECDM*

We used multiple regression to test whether reward learning and working memory predicted individual differences in EEfRT performance. Then we examined whether positive symptom severity, negative symptom severity, or depressive symptom severity (using PANSS-Positive, PANSS-Negative and MADRS continuous scores), moderated the effect of reward learning or working memory, on EEfRT performance. A significant *Symptom* x *Reward learning* (or *Working memory*) interaction was taken as evidence of moderation.

## 3.0. Results

### *3.1. Sample characteristics*

Sample characteristics are shown in Table 1. The sample was 55.4% male (n=134) and aged 18 to 62 (M±SD=35.12±11.90). The controls had more years of education relative to the patient group, $t(238)=2.54$, $p=0.01$ (16.05±2.24 vs. 15.04±2.63) but did not differ on other demographic characteristics.

----------Insert Table 1 about here----------

### *3.2. Impaired ECDM in individuals with psychosis-spectrum disorders*

Two subjects with psychosis were excluded for failing to complete sufficient EEfRT trials. Performance on the EEfRT across the control and psychosis groups is shown in Figs. 1A and 1B, with further diagnostic breakdowns within the patients shown in Fig. 1C. Overall, there was a trend for the psychosis group to choose the hard option less often than the control group ($p=0.07$). Furthermore, there was a significant *Group* x *Probability* x *Value* interaction, $F(6,1404)=2.25$, $p=0.04$, $\eta_p^2=0.01$. The *Probability* x *Value* interaction was significant within both groups (both $ps<0.001$), therefore, we next considered the *Group* x *Probability* interaction at each level of *Value*. The 2-way interaction was only significant at the highest two reward

values (both *p*s<0.01). Bonferroni-corrected post hoc tests of simple effects showed that for the highest two reward values (i.e., $2.60-$3.45 and >$3.50), the psychosis group chose the hard option significantly less than the control group when the probability of winning was highest (88%; both *p*s<0.01). No group differences were evident at the lower two probabilities (i.e., 12% and 50%, both *p*s>0.10). Finally, we examined the *Group* x *Value* interaction at each level of *Probability*. The 2-way interaction was only significant at the 88% probability (*p*<0.001). Post hoc tests showed that at this probability, individuals in the psychosis group chose the hard option significantly less than the control group at reward values $2.60 or higher (i.e., at the highest two reward values; both *p*s<0.01).

----------Insert Figure 1 about here----------

To further unpack these group differences, we subtracted the average proportion of hard choices made at the lowest probability level (across all reward values) from the average proportion of hard choices made at the highest probability level (a variable we termed "EEfRT-probability"). A higher score on this variable indicated greater modulation of effort as a function of reward probability. We adopted a similar approach for examining the degree to which effort was modulated by reward value, by subtracting the average proportion of hard choices made at the lowest reward value (across all probability levels) from the proportion made at the highest reward value (termed "EEfRT-value"). There were significant group differences for both EEfRT-probability, $t(234)=3.37$, $p=0.001$, Cohen's $d=0.54$, and EEfRT-value, $t(234)=2.66$, $p=0.008$, Cohen's $d=0.42$, with a larger effect size for EEfRT-probability.

In terms of diagnostic differences, there was a significant *Diagnosis* x *Probability* x *Value* interaction, $F(18,1338)=1.73$, $p=0.03$, $\eta_p^2=0.02$. Briefly, post hoc tests of this interaction

(which are described in full in the *Supplement*) showed the greatest deficits in EEfRT performance in those with a diagnosis of schizophrenia.

### *3.3. More severe psychotic symptoms are associated with reduced effort expenditure for high probability, high value rewards*

Poorer EEfRT-probability scores were associated with more severe positive ($r$=-0.28, $p$<0.001) and negative ($r$=-0.28, $p$<0.001) symptoms on the PANSS, as well as poorer community functioning on the MCAS ($r$=0.29, $p$<0.001). Poorer EEfRT-value scores were associated with more severe negative symptoms on the PANSS ($r$=-0.27, $p$<0.001) and poorer community functioning on the MCAS ($r$=0.25, $p$=0.001). Fisher's test of independent correlations showed that EEfRT-probability was more strongly correlated with positive (Z=-7.29, $p$<0.001) and negative symptom severity (Z=-6.68, $p$<0.001), as well as community functioning on the MCAS (Z=-1.97, $p$=0.049), than was EEfRT-value. EEfRT performance was not correlated with general symptoms (both $p$s>0.05; see Table S1).

### *3.4. Intact reward learning in psychosis-spectrum disorders*

PRT data from 34 patients and 3 controls were excluded following quality control assessments. Mean±SEM response bias in each block of the PRT are shown in Fig. 2A, with further breakdowns by diagnosis shown in Fig. 2C. There was a main effect of *Block*, $F(1,203)=9.48$, $p$=0.002, $\eta_p^2$=0.05, where across groups, response bias was higher in block 2 relative to block 1. However, neither the main effect of *Group* ($p$=0.76) nor the *Group* x *Block* interaction ($p$=0.97) were significant. For the *Diagnosis* x *Block* ANOVA, no significant main effects or interactions involving *Diagnosis* emerged (both $p$s>0.10), indicating that reward learning did not differ as a function of psychotic disorder diagnosis.

----------Insert Figure 2 about here----------

### *3.5. Impaired working memory in psychosis-spectrum disorders*

Mean±SEM DST (z-scored) performance in the control and psychosis groups are shown in Fig. 2B, with further diagnostic breakdowns shown in Fig. 2D. Compared to the controls, the psychosis group had poorer working memory, $t(237)=3.93$, $p<0.001$. There was also a significant main effect of *Diagnosis*, $F(3,226)=11.56$, $p<0.001$, $\eta_p^2=0.13$. Bonferroni-corrected post hoc tests showed that working memory was worse in the schizophrenia and schizoaffective disorder groups relative to the controls (both $p$s$<0.01$), and the performance of the schizophrenia group was also significantly worse than that of the bipolar disorder group ($p<0.001$). The bipolar disorder and control groups did not differ ($p=0.35$).

### *3.6. Working memory but not reward learning, predicts ECDM across the entire sample*

When entered into the same regression model, working memory ($\beta=0.29$, $p<0.001$), but not reward learning ($\beta=0.08$, $p=0.27$) predicted effort performance (EEfRT-probability) across the entire sample. Specifically, better working memory predicted greater effort expenditure for higher probability rewards. The same was true when effort performance was defined as EEfRT-value, where better working memory ($\beta=0.23$, $p=0.001$) but not reward learning ($\beta=0.02$, $p=0.81$) predicted greater effort expenditure for higher value rewards. The same pattern of results was evident when examining the patient group alone, indicating that the effects were likely not driven by the control group.

### *3.7. Moderating effects of psychotic symptom severity*

When considering EEfRT-probability, a significant *Negative symptoms* x *Reward learning* interaction emerged, B=0.06, SE=0.03, $t=2.39$, $p=0.02$, indicating that the relationship between reward learning and EEfRT-probability was moderated by negative symptom severity (PANSS-Negative). To unpack this interaction, we used the Johnson-Neyman Technique (JNT)

(Bauer and Curran, 2005; Johnson and Neyman, 1936) to test the conditional effect of reward learning on effort performance as a function of PANSS-Negative scores. This technique identifies the point(s) along a continuous moderator where the relationship between the independent variable and the dependent variable transition from being statistically significant to nonsignificant. Results showed that the relationship between reward learning and EEfRT-probability performance was significant at centered PANSS-Negative values >0.13 (encompassing the top 45% of PANSS-Negative values). For scores in this range, poorer reward learning predicted poorer ability to modulate effort as a function of reward probability (see Fig. 3A). The reverse moderation, which examined whether negative symptom severity moderated the effect of working memory on EEfRT-probability, was not significant ($p$=0.85). Furthermore, neither positive or general psychotic symptom severity moderated the effect of working memory or reward learning, on EEfRT-probability ($p$-values for all interaction terms >0.10), suggesting that this moderation effect was specific to negative symptoms.

### *3.8. Moderating effects of depressive symptom severity*

When considering EEfRT-probability, a significant *Depression* x *Working Memory* interaction emerged, B=-0.01, SE=0.002, $t$=-2.72, $p$=0.007, indicating that the relationship between working memory and EEfRT-probability was moderated by depressive symptom severity (MADRS). Specifically, the relationship between working memory and EEfRT-probability scores were significant at centered MADRS scores <2.30 (the bottom 60% of MADRS scores). For individuals with depression scores in this range (indicative of no or mild depressive symptoms), better working memory predicted better ability to modulate effort as a function of reward probability (Fig. 3B). For individuals with more severe depressive symptoms, working memory did not predict effort performance. We confirmed that the *Depression* x

*Working memory* interaction was specific to working memory and not other components of neurocognition (see *Supplement*). Furthermore, the reverse moderation, which examined whether depression moderated the effect of reward learning (rather than working memory) on EEfRT-probability, was not significant ($p$=0.37).

The final two significant moderation models are shown in Table 2. The moderation effects observed were specific to effort performance defined as EEfRT-probability, as neither interaction term was significant when considering predictors of EEfRT-value (both $p$s>0.1).

----------Insert Table 2 and Figure 3 about here----------

**4.0. Discussion**

Using objective measures of effort, reward learning and working memory in a transdiagnostic sample, we examined predictors of ECDM across the psychosis spectrum, and whether these predictors varied as a function of symptom profile. Replicating prior studies in samples with schizophrenia and/or schizoaffective disorder (Barch et al., 2014; Fervaha et al., 2013; Gold et al., 2013; Huang et al., 2016; McCarthy et al., 2016; Treadway et al., 2015), we found evidence of reduced willingness to expend effort for rewards of medium to high probability and of high value, in individuals with psychosis-spectrum disorders. These deficits were primarily evident in individuals with schizoaffective disorder and schizophrenia, whereas individuals with psychotic mood disorders (bipolar disorder) did not differ significantly from controls. Poorer ECDM was also associated with more severe positive and negative symptoms; however, we found evidence of stronger associations between symptom severity and the ability to modulate effort as a function of reward probability rather than as a function of reward value.

Several other novel findings emerged. Specifically, decreased effort expenditure for higher probability rewards was associated with working memory but not with reward learning

across the entire sample, and diagnosis did not moderate any of these relationships. This suggests that working memory contributes to variability in ECDM across psychotic disorders. However, within the patient group, the strength of the relationship between working memory and ECDM varied as a function of symptom profile. For individuals with more severe depressive symptoms and more severe negative symptoms, working memory was not a significant predictor of ECDM. Moreover, for individuals with more severe negative symptoms, poorer ECDM was instead predicted by impairments in reward learning on the PRT. Taken together, these findings replicate prior evidence of deficits in ECDM in psychosis and reiterate the important contribution of working memory to ECDM disturbances in this population, as outlined in a recent theoretical model (Culbreth et al., 2018). However, they also extend prior work by showing that the degree to which impairments in working memory contribute to ECDM disturbances vary in accordance with clinical symptom type and severity. Specifically, working memory may play less of a role in ECDM deficits in individuals with prominent negative symptoms or depressive symptoms. For the former, ECDM deficits may be more closely tied to disturbances in reward learning.

These findings also hint at the different neural systems that may underpin motivational deficits in psychosis. For example, both prefrontal (e.g., anterior cingulate cortex, medial prefrontal cortex) and subcortical (e.g., striatum) regions have been implicated in ECDM (Filla et al., 2018; Hauser et al., 2017; Hogan et al., 2018) and variable functioning in either system may lead to reduced willingness to work for rewards. Our findings suggest that deficits in working memory processes (i.e., prefrontal systems) may play a particularly prominent role in ECDM in psychosis. Accordingly, interventions targeting prefrontal functioning, such as cognitive remediation therapy, may be efficacious in treating motivational deficits in individuals with psychotic disorders (a hypothesis that has recently received some support, Cella et al., 2017).

However, evidence that reward learning (but not working memory) predicts ECDM in individuals with more severe negative symptoms suggests that alternate treatment targets may be warranted for these individuals. Whether this association is state-dependent (i.e., only arises at times where negative symptoms dominate) or trait-like (i.e., arises in individuals whose psychotic disorder is characterized by marked negative symptoms) is an important question for future research. Indeed, studies in remitted depressed samples have shown that deficits in reward learning on the PRT persist even following symptom remission (Pechtel et al., 2013; Whitton et al., 2016), and may therefore represent trait-like deficits. Longitudinal studies are needed to test whether reward system dysfunction may underpin abnormal ECDM during periods where negative symptoms predominate, or whether there are unique psychotic disorder subtypes that are characterised by marked negative symptoms, for whom ECDM is linked to reward processing impairment.

Some limitations must be kept in mind when interpreting our findings. First, although the DST is a widely-used, well-validated test of working memory, future studies would benefit from including a longer, more dynamic working memory task (such as the N-back task) to ensure comparable sensitivity with behavioral reward learning tasks like the PRT. Second, although the EEfRT and PRT both involve modulating behavior as a function of increasing reward probability, there was no correlation between response bias on the PRT and EEfRT-probability scores. This may be because the EEfRT evaluates an individual's ability to modulate effort using explicit information about reward probability, whereas the PRT evaluates implicit reward learning wherein the differing probability of rewards is not explicitly known. Accordingly, future studies may show a greater association between reward learning and performance on the EEfRT using a measure of explicit, rather than implicit, reward learning. Finally, one alternative

interpretation of our findings is that ECDM abnormalities in patients with poorer working memory may simply reflect a failure of valid testing in this subpopulation due to greater difficulty understanding the EEfRT task requirements. This possibility was highlighted in a recent study using computational modeling of EEfRT data, which found that a subset of individuals with schizophrenia who failed to use reward and probability information to guide effortful behavior also showed the greatest cognitive deficits (Cooper et al., 2019). Although our moderation effects were specific to the working memory subscale of the BACS and did not extend to subscales assessing other aspects of cognition, it will be important for future research to use computational approaches to further disentangle the degree to which ECDM deficits in psychosis can be reliably linked to working memory or reward-based disturbances, as opposed to more general deficits in task performance.

In sum, the current study is the first to show that factors contributing to ECDM impairment in individuals with psychotic disorders vary as a function of clinical symptom profile. This suggests that there may be multiple pathways to ECDM deficits in individuals with psychosis. Identification of specific pathways of impairment in motivation, including associations with trait and state illness characteristics, highlight actionable targets that can be used in the development of effective, individualized interventions for this critical symptom domain.

**References**


Baldessarini, R.J., 2013. Chemotherapy in Psychiatry. Springer Press, New York.

Barch, D.M., Carter, C.S., Gold, J.M., Johnson, S.L., Kring, A.M., MacDonald III, A.W., Pizzagalli, D.A., Ragland, J.D., Silverstein, S.M., Strauss, M.E., 2017. Explicit and implicit reinforcement learning across the psychosis spectrum. J. Abnorm. Psychol. 126 (5) 694-711.

Barch, D.M., Treadway, M.T., Schoen, N., 2014. Effort, anhedonia, and function in schizophrenia: Reduced effort allocation predicts amotivation and functional impairment. J. Abnorm. Psychol. 123 (2) 387-397.

Barker, S., Barron, N., McFarland, B.H., Bigelow, D.A., 1994. A community ability scale for chronically mentally ill consumers: Part I. Reliability and validity. Community Ment. H J. 30 (4) 363-383.

Bauer, D.J., Curran, P.J., 2005. Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. Multivariate Behav. Res. 40 (3) 373-400.

Bleuler, E., 1911/1950. Dementia Praecox or the Group of Schizophrenias. International University Press, New York.

Bora, E., Pantelis, C., 2015. Meta-analysis of cognitive impairment in first-episode bipolar disorder: comparison with first-episode schizophrenia and healthy controls. Schiz. Bull. 41 (5) 1095-1104.

Cella, M., Preti, A., Edwards, C., Dow, T., Wykes, T., 2017. Cognitive remediation for negative symptoms of schizophrenia: a network meta-analysis. Clin. Psychol. Rev. 52 43-51.

Cooper, J., Barch, D., Reddy, L., Horan, W., Green, M., Treadway, M., 2019. Effortful goal-directed behavior in schizophrenia: Computational subtypes and associations with cognition. J. Abnorm. Psychol. 128 (7) 710-722.

Culbreth, A., Moran, E., Barch, D., 2018. Effort-cost decision-making in psychosis and depression: could a similar behavioral deficit arise from disparate psychological and neural mechanisms? Psychol. Med. 48 (6) 889-904.

Culbreth, A., Westbrook, A., Barch, D., 2016. Negative symptoms are associated with an increased subjective cost of cognitive effort. J. Abnorm. Psychol. 125 (4) 528-536.

Dowd, E.C., Barch, D.M., 2012. Pavlovian reward prediction and receipt in schizophrenia: relationship to anhedonia. PloS one 7 (5) e35622.

Fervaha, G., Foussias, G., Agid, O., Remington, G., 2014. Motivational and neurocognitive deficits are central to the prediction of longitudinal functional outcome in schizophrenia. Acta Psychiatr. Scand. 130 (4) 290-299.

Fervaha, G., Graff-Guerrero, A., Zakzanis, K.K., Foussias, G., Agid, O., Remington, G., 2013. Incentive motivation deficits in schizophrenia reflect effort computation impairments during cost-benefit decision-making. J. Psychiatr. Res. 47 (11) 1590-1596.

Filla, I., Bailey, M.R., Schipani, E., Winiger, V., Mezias, C., Balsam, P.D., Simpson, E.H., 2018. Striatal dopamine D2 receptors regulate effort but not value-based decision making and alter the dopaminergic encoding of cost. Neuropsychopharmacology 43 (11) 2180-2189.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B., 2002. Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. Biometrics Research, New York State Psychiatric Hospital, New York.

Foussias, G., Mann, S., Zakzanis, K., van Reekum, R., Agid, O., Remington, G., 2011. Prediction of longitudinal functional outcomes in schizophrenia: the impact of baseline motivational deficits. Schizophr. Res. 132 (1) 24-27.

Gold, J.M., Barch, D.M., Feuerstahler, L.M., Carter, C.S., MacDonald III, A.W., Ragland, J.D., Silverstein, S.M., Strauss, M.E., Luck, S.J., 2019. Working memory impairment across psychotic disorders. Schiz. Bull. 45 (4) 804-812.

Gold, J.M., Strauss, G.P., Waltz, J.A., Robinson, B.M., Brown, J.K., Frank, M.J., 2013. Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. Biol. Psychiatry 74 (2) 130-136.

Hauser, T.U., Eldar, E., Dolan, R.J., 2017. Separate mesocortical and mesolimbic pathways encode effort and reward learning signals. Proc. Natl. Acad. Sci. 114 (35) E7395-E7404.

Hautus, M.J., 1995. Corrections for extreme proportions and their biasing effects on estimated values of d′. Behav. Res. Method. Instrum. Comput. 27 (1) 46-51.

Heerey, E.A., Bell-Warren, K.R., Gold, J.M., 2008. Decision-making impairments in the context of intact reward sensitivity in schizophrenia. Biol. Psychiatry 64 (1) 62-69.

Hershenberg, R., Satterthwaite, T.D., Daldal, A., Katchmar, N., Moore, T.M., Kable, J.W., Wolf, D.H., 2016. Diminished effort on a progressive ratio task in both unipolar and bipolar depression. J. Affect. Disord. 196 97-100.

Hogan, P.S., Galaro, J.K., Chib, V.S., 2018. Roles of ventromedial prefrontal cortex and anterior cingulate in subjective valuation of prospective effort. Cereb. Cortex 29 (10) 4277-4290.

Huang, J., Yang, X.-h., Lan, Y., Zhu, C.-y., Liu, X.-q., Wang, Y.-f., Cheung, E.F., Xie, G.-r., Chan, R.C., 2016. Neural substrates of the impaired effort expenditure decision making in schizophrenia. Neuropsychology 30 (6) 685-696.

Johnson, P.O., Neyman, J., 1936. Tests of certain linear hypotheses and their application to some educational problems. Statistical Research Memoirs 1 57-93.

Kaiser, R.H., Treadway, M.T., Wooten, D.W., Kumar, P., Goer, F., Murray, L., Beltzer, M., Pechtel, P., Whitton, A., Cohen, A.L., 2018. Frontostriatal and dopamine markers of individual differences in reinforcement learning: A multi-modal investigation. Cereb. Cortex 28 (12) 4281-4290.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schiz. Bull. 13 (2) 261-276.

Keefe, R.S., 1999. Brief Assessment of Cognition in Schizophrenia (BACS) Manual - A: Version 2.1. Duke University Medical Center, North Carolina.

Krabbendam, L., Arts, B., van Os, J., Aleman, A., 2005. Cognitive functioning in patients with schizophrenia and bipolar disorder: a quantitative review. Schizophr. Res. 80 (2-3) 137-149.

Kraepelin, E., 1919/1971. Dementia praecox and paraphrenia. Krieger Publishing Company, New York.

Lewandowski, K.E., Cohen, B.M., Keshavan, M.S., Sperry, S.H., Öngür, D., 2013. Neuropsychological functioning predicts community outcomes in affective and non-affective psychoses: a 6-month follow-up. Schizophr. Res. 148 (1-3) 34-37.

Mason, L., Trujillo-Barreto, N.J., Bentall, R.P., El-Deredy, W., 2016. Attentional bias predicts increased reward salience and risk taking in bipolar disorder. Biol. Psychiatry 79 (4) 311-319.

McCarthy, J.M., Treadway, M.T., Bennett, M.E., Blanchard, J.J., 2016. Inefficient effort allocation and negative symptoms in individuals with schizophrenia. Schizophr. Res. 170 (2-3) 278-284.

Montgomery, S.A., Åsberg, M., 1979. A new depression scale designed to be sensitive to change. Br. J. Psychiatry 134 (4) 382-389.

Nusslock, R., Almeida, J.R., Forbes, E.E., Versace, A., Frank, E., LaBarbara, E.J., Klein, C.R., Phillips, M.L., 2012. Waiting to win: elevated striatal and orbitofrontal cortical activity during reward anticipation in euthymic bipolar disorder adults. Bipolar Disord. 14 (3) 249-260.

Owoso, A., Carter, C.S., Gold, J., MacDonald, A., Ragland, J.D., Silverstein, S., Strauss, M., Barch, D., 2013. Cognition in schizophrenia and schizo-affective disorder: impairments that are more similar than different. Psychol. Med. 43 (12) 2535-2545.

Pechtel, P., Dutra, S.J., Goetz, E.L., Pizzagalli, D.A., 2013. Blunted reward responsiveness in remitted depression. J. Psychiatr. Res. 47 (12) 1864-1869.

Pizzagalli, D.A., Goetz, E., Ostacher, M., Iosifescu, D.V., Perlis, R.H., 2008a. Euthymic patients with bipolar disorder show decreased reward learning in a probabilistic reward task. Biol. Psychiatry 64 (2) 162-168.

Pizzagalli, D.A., Iosifescu, D., Hallett, L.A., Ratner, K.G., Fava, M., 2008b. Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. J. Psychiatr. Res. 43 (1) 76-87.

Pizzagalli, D.A., Jahn, A.L., O'Shea, J.P., 2005. Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. Biol. Psychiatry 57 (4) 319-327.

Roiser, J.P., Cannon, D.M., Gandhi, S.K., Tavares, J.T., Erickson, K., Wood, S., Klaver, J.M., Clark, L., Zarate Jr, C.A., Sahakian, B.J., 2009. Hot and cold cognition in unmedicated depressed subjects with bipolar disorder. Bipolar Disord. 11(2) 178-189.

Santesso, D.L., Dillon, D.G., Birk, J.L., Holmes, A.J., Goetz, E., Bogdan, R., Pizzagalli, D.A., 2008. Individual differences in reinforcement learning: behavioral, electrophysiological, and neuroimaging correlates. Neuroimage 42 (2) 807-816.

Silveira, M.M., Tremblay, M., Winstanley, C.A., 2018. Dissociable contributions of dorsal and ventral striatal regions on a rodent cost/benefit decision-making task requiring cognitive effort. Neuropharmacology 137 322-331.

Sperry, S.H., O'Connor, L.K., Öngür, D., Cohen, B.M., Keshavan, M.S., Lewandowski, K.E., 2015. Measuring cognition in bipolar disorder with psychosis using the MATRICS consensus cognitive battery. J. Int. Neuropsychol. Soc. 21 (6) 468-472.

Treadway, M.T., Bossaller, N.A., Shelton, R.C., Zald, D.H., 2012. Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. J. Abnorm. Psychol. 121 (3) 553-558.

Treadway, M.T., Buckholtz, J.W., Schwartzman, A.N., Lambert, W.E., Zald, D.H., 2009. Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. PloS One 4 (8) e6598.

Treadway, M.T., Peterman, J.S., Zald, D.H., Park, S., 2015. Impaired effort allocation in patients with schizophrenia. Schizophr. Res. 161 (2-3) 382-385.

Whitton, A.E., Kakani, P., Foti, D., Van't Veer, A., Haile, A., Crowley, D.J., Pizzagalli, D.A., 2016. Blunted neural responses to reward in remitted major depression: a high-density event-related potential study. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 1 (1) 87-95.

Whitton, A.E., Treadway, M.T., Pizzagalli, D.A., 2015. Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. Curr. Opin. Psychiatry 28 (1) 7-12.

Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A rating scale for mania: reliability, validity and sensitivity. Br. J. Psychiatry 133 (5) 429-435.

**Figure captions**

*Figure 1.* Mean (±SEM) proportion of hard choices made as a function of probability level and reward value across the healthy control (**A**) and psychosis (**B**) groups, as well as within each of three main diagnostic subgroups (**C**). BD=bipolar disorder with psychotic features; SZA=schizoaffective disorder; SZ=schizophrenia.

*Figure 2.* Mean (±SEM) response bias in block 1 and block 2 of the PRT (**A**) and DST (**B**) across the control and psychosis groups. Performance in each of the three main diagnostic subgroups is shown in panels (**C**) and (**D**).

*Figure 3.* Johnson-Neyman plots showing regions of significance (blue/darker regions demarcated by dashed vertical lines) and confidence intervals for the conditional effect of (**A**) reward learning on EEfRT-probability as a function of PANSS-Negative scores. Results show that poorer reward learning predicted poorer effort performance but only for individuals with more severe negative symptoms. The second panel (**B**) shows the conditional effect of working memory on EEfRT-probability as a function of MADRS scores. Results show that poorer working memory predicted poorer effort performance but only for individuals with minimal to no depressive symptoms.

**Table 1.** Sample characteristics

| | Control (n=52) | Psychosis (n=190) | Test Statistic | *p* |
|---|---|---|---|---|
| Age, M (SD) | 37.8 (14.4) | 34.4 (11.0) | $t$=1.86 | 0.06 |
| Yrs. Education, M (SD) | 16.0 (2.2) | 15.0 (2.6) | $t$=2.54 | 0.01 |
| Female, N (%) | 22 (42.3) | 86 (45.3) | $X^2$=0.14 | 0.70 |
| White, N (%) | 35 (67.3) | 123 (64.7) | $X^2$=0.12 | 0.73 |
| Hispanic/Latino, N (%) | 4 (7.7) | 21 (11.1) | $X^2$=0.50 | 0.48 |
| MADRS, M (SD) | -- | 15.9 (10.4) | -- | -- |
| YMRS, M (SD) | -- | 11.1 (8.1) | -- | -- |
| PANSS Pos., M (SD) | -- | 13.9 (5.3) | -- | -- |
| PANSS Neg., M (SD) | -- | 13.1 (4.9) | -- | -- |
| PANSS Gen., M (SD) | -- | 30.8 (7.7) | -- | -- |
| PANSS Tot., M (SD) | -- | 57.8 (14.3) | -- | -- |
| MCAS, M (SD) | -- | 48.0 (5.0) | -- | -- |
| CPZ equiv., M (SD) | -- | 274.0 (325.0) | -- | -- |

*Note.* M=Mean; SD=Standard Deviation; MADRS=Montgomery-Asberg Depression Rating Scale; YMRS=Young Mania Rating Scale; PANSS=Positive and Negative Symptoms Scale; Pos.=positive symptom subscale of the PANSS; Neg.=negative symptom subscale of the PANSS; Gen.=general symptom subscale of the PANSS; Tot.=total symptom subscale of the PANSS; MCAS= Multnomah Community Ability Scale; CPZ equiv.=Chlorpromazine equivalence of antipsychotic medication load.

**Table 2.** Moderation analyses

Model 1: Moderating effect of negative symptom severity on the relationship between reward learning and effort performance

| | B | SE | *t* | *p* | 95% CI | |
|---|---|---|---|---|---|---|
| (constant) | 0.23 | 0.02 | 9.54 | 0.000 | 0.181 | 0.275 |
| Working memory (DST) | 0.04 | 0.02 | 1.88 | 0.062 | -0.002 | 0.080 |
| Reward learning (PRT) | 0.24 | 0.12 | 1.92 | 0.057 | -0.007 | 0.482 |
| Negative symptoms (PANSS Neg) | -0.01 | 0.01 | -2.74 | 0.007 | -0.024 | -0.004 |
| Negative Symptoms x Reward learning | 0.06 | 0.03 | 2.39 | 0.018 | 0.011 | 0.118 |

Model 2: Moderating effect of depression severity on the relationship between working memory and effort performance

| | B | SE | *t* | *p* | 95% CI | |
|---|---|---|---|---|---|---|
| (constant) | 0.27 | 0.02 | 48.41 | 0.000 | 0.159 | 0.256 |
| Reward learning (PRT) | 0.19 | 0.12 | 1.53 | 0.127 | -0.055 | 0.434 |
| Working memory (DST) | 0.05 | 0.02 | 2.67 | 0.009 | 0.014 | 0.095 |
| Depression (MADRS) | 0.00 | 0.00 | 1.16 | 0.249 | -0.002 | 0.008 |
| Depression x Working memory | -0.01 | 0.00 | -2.72 | 0.007 | -0.010 | -0.002 |

*Note.* DST=Digit Sequencing Task subscale from the Brief Assessment of Cognition in Schizophrenia; PRT=Probabilistic Reward Task; MADRS= Montgomery-Asberg Depression Rating Scale. For both models, the dependent variable was effort performance, defined as EEfRT-probability (all interaction terms for models predicting EEfRT-value were non-significant).

# Supplementary Information

## Supplemental Methods

### *PRT quality control*

Trials where the reaction time was <150ms or >2500ms, or where the RT was ±3SD from the mean, were excluded. Subjects with <55% accuracy, with >20% outlier trials, or with a rich: lean reward ratio lower than 2.4:1, were also excluded.

## Supplemental Results

To test whether any of the primary effects of interest differed across the major diagnostic groups included in our psychosis sample, we re-ran all analyses and included diagnosis as both a main effect and moderator. The diagnostic categories included were: bipolar disorder with psychotic features (n=90), schizoaffective disorder (n=49) and schizophrenia (n=42). Given the small number of participants with schizophreniform disorder (n=1), psychotic disorder not otherwise specified (n=3), and major depressive disorder with psychotic features (n=5), these diagnoses were not included in the analysis. For all ANOVAs, *Diagnosis* was coded sequentially as a single categorical variable (control, bipolar disorder, schizoaffective disorder, schizophrenia). For all moderated regression analyses, *Diagnosis* was represented using three dummy coded variables that indicated the presence of bipolar disorder (control=0, bipolar disorder=1, schizoaffective disorder=0, schizophrenia=0), schizoaffective disorder (control=0, bipolar disorder=0, schizoaffective disorder=1, schizophrenia=0) and schizophrenia (control=0, bipolar disorder=0, schizoaffective disorder=0, schizophrenia=1).

***EEfRT performance across categorical diagnoses***

The percentage of hard choices made across each level of probability and reward value as a function of diagnostic group are shown in Figure 1 of the main manuscript. The *Diagnosis* x *Probability* x *Value* interaction was significant, $F(18,1338)=1.73$, $p=0.03$, $\eta_p^2=0.02$. First, this 3-way interaction was followed-up by examining the *Probability* x *Value* interaction separately within each diagnostic category. The *Probability* x *Value* interaction was significant in all four groups (all *p*s<0.01), indicating that the 3-way interaction was not driven by the *Probability* x *Value* interaction differing at the level of *Diagnosis*. Second, we examined the *Diagnosis* x *Probability* interaction separately at each level of *Value*. The interaction was significant at all values (all *p*s<0.005), indicating that the 3-way interaction was not driven by the *Diagnosis* x *Probability* interaction varying at different levels of *Value*. Third, we examined the *Diagnosis* x *Value* interaction separately at each level of *Probability*. This 2-way interaction was not significant at the 12% probability ($p=0.27$), but was significant at the 50% ($p=0.001$) and 88% probabilities ($p<0.001$). This indicates that the *Diagnosis* x *Value* interaction differed across levels of *Probability*.

The two significant 2-way interactions were then followed-up using Bonferroni-corrected post hoc tests of simple effects. For the 50% probability, significant differences between diagnostic categories emerged at the highest reward value ($3.50-$4.20). Specifically, individuals with schizophrenia chose the hard option significantly less compared to the healthy control ($p=0.03$) and bipolar disorder groups ($p=0.049$), but did not differ from the schizoaffective disorder group ($p=1.00$). For the 88% probability, significant differences between diagnostic categories emerged for the top two reward values. For the second-highest reward value ($2.60-$3.45), the percentage of hard choices in the schizophrenia and schizoaffective

disorder groups was significantly lower than in the healthy control (both *p*s<0.01) and the bipolar disorder groups (both *p*s<0.05). For the highest reward value ($3.50-$4.20), the schizophrenia and schizoaffective disorder groups chose the hard option significantly less than the healthy control group (both *p*s<0.001) and the schizophrenia group also chose it less than the bipolar disorder group (*p*<0.001). There was also a trend for the schizoaffective disorder group to choose the hard option less than the bipolar disorder group (*p*=0.06). For all post hoc tests, the bipolar disorder group did not differ from the healthy control group (all *p*s>0.10).

Taken together, these results suggest that when the probability of winning was medium to high (i.e., 50-88%), and when the value of the reward was >$2.60, individuals with schizophrenia or schizoaffective disorder chose the hard option disproportionately less than healthy controls or those with bipolar disorder.

When we examined the effect of *Diagnosis* on the EEfRT-probability and EEfRT-value measures, a main effect of *Diagnosis* emerged for both measures, EEfRT-probability: $F(3,223)=15.30$, $p<0.001$, $\eta_p^2=0.17$; EEfRT-value: $F(3,223)=9.51$, $p<0.001$, $\eta_p^2=0.11$. The pairwise comparisons showed that for EEfRT-probability, the schizophrenia and schizoaffective disorder performed more poorly than both the control and bipolar disorder groups (all *p*s<0.01), however the schizophrenia and schizoaffective disorder groups (*p*=0.19) and the control and bipolar disorder groups (*p*=1.00) did not differ from one another. In terms of effect sizes, the difference in EEfRT-probability scores between the schizophrenia and control groups was Cohen's *d*=1.40, and between the schizoaffective disorder and the control groups was Cohen's *d*=0.76. A similar pattern of findings was observed for EEfRT-value, with the exception that the schizoaffective disorder group did not differ significantly from the bipolar disorder group (*p*=0.08). Effect sizes for the difference in EEfRT-value scores between the schizophrenia and

control groups was Cohen's *d*=1.02, and between the schizoaffective disorder and control groups was Cohen's *d*=0.59.

These results suggest that while both the schizophrenia and schizoaffective disorder groups has greater difficulty modulating effort as a function of both reward probability and reward value, the effect sizes for the impairment in coding reward probability were larger.

***Diagnosis as a moderator of the effects of working memory or reward learning, on EEfRT performance***

*Diagnosis* did not moderate the effect of working memory or reward learning on effort performance (either defined as EEfRT-probability or EEfRT-value; *p*-values for all interaction terms >0.05).

***Testing the specificity of the Depression x Working memory interaction on effort performance***

To test whether the *Depression* x *Working memory* interaction we observed on EEfRT-probability was specific to working memory and not other aspects of cognition, we re-ran the moderation analyses substituting the Digit Sequencing Task subscale for the other subscales of the BACS. These subscales were: Verbal Memory, Motor Speed, Verbal Fluency, Information Processing, Executive Functioning. Results showed evidence of a significant *Depression* x *Executive Functioning* interaction effect, B=-0.004, SE=0.002, *p*=0.04, 95% CI: -0.007 to -0.0001. The results of the Johnson-Neyman Technique analysis showed that the relationship between executive functioning scores and EEfRT-probability scores were significant at centered MADRS scores below 4.29 (the bottom 62.6% of depression scores). For individuals with MADRS scores in this range, poorer executive functioning predicted poorer ability to modulate effort as a function of increasing reward probability. However, when working memory was

entered into the model as a covariate, this interaction became non-significant ($p$=0.27). Furthermore, when executive functioning was entered into the original model that included working memory as the predictor of interest, the *Depression* x *Working memory* interaction remained significant ($p$=0.01). Therefore, although there was evidence that depressive symptom severity also moderated the effect of executive functioning on EEfRT-probability performance, this effect disappeared when controlling for working memory. In contrast, depressive symptom severity moderated the effect of working memory on EEfRT-probability performance, and this effect remained significant when controlling for executive functioning. The interaction terms for the other BACS scales were not significant [*Depression* x *Verbal memory* ($p$=0.48), *Depression* x *Verbal fluency* ($p$=0.29), *Depression* x *Motor speed* ($p$=0.92), *Depression* x *Information processing* ($p$=0.55), *Depression* x *Executive functioning* ($p$=0.12)]. Taken together, these results indicate that the moderation effect observed was specific to working memory and not other aspects of cognition.

**Table S1.** Associations between effort performance, reward learning, working memory and symptom severity

| | EEfRT-prob | EEfRT-value | RL | WM | PANSS Pos. | PANSS Neg. | PANSS Gen. | MADRS | YMRS | MCAS |
|---|---|---|---|---|---|---|---|---|---|---|
| EEfRT-prob | | | | | | | | | | |
| EEfRT-value | 0.45** | | | | | | | | | |
| RL | 0.09 | 0.07 | | | | | | | | |
| WM | 0.23** | 0.21** | 0.02 | | | | | | | |
| PANSS Pos. | -0.28** | -0.06 | 0.04 | -0.27** | | | | | | |
| PANSS Neg. | -0.28** | -0.27** | 0.03 | -0.25** | 0.28** | | | | | |
| PANSS Gen. | -0.14 | -0.08 | -0.09 | -0.19* | 0.51** | 0.50** | | | | |
| MADRS | 0.08 | 0.06 | 0.00 | -0.06 | 0.21** | 0.16* | 0.65** | | | |
| YMRS | -0.09 | 0.00 | 0.08 | -0.14 | 0.73** | 0.07 | 0.41** | 0.30** | | |
| MCAS | 0.29** | 0.25** | -0.05 | 0.24** | -0.44** | -0.66** | -0.52** | -0.22** | -0.24** | |
| CPZ | -0.15* | -0.15* | -0.09 | -0.18* | 0.09 | 0.04 | -0.03 | -0.14 | -0.03 | -0.11 |

*Note.* Correlations reflect associations in patients only. RL=reward learning on the Probabilistic Reward Task; WM=working memory on the Digit Sequencing Task; PANSS=Positive and Negative Symptoms Scale; Pos.=positive symptom subscale of the PANSS; Neg.=negative symptom subscale of the PANSS; Gen.=general symptom subscale of the PANSS; MADRS=Montgomery-Asberg Depression Rating Scale; YMRS=Young Mania Rating Scale; MCAS= Multnomah Community Ability Scale; CPZ equiv.=Chlorpromazine equivalence of antipsychotic medication load.*$p$<0.05; **$p$<0.01