Assessing and remediating altered reinforcement learning in depression

Vanessa Marie Bowen Brown

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Psychology

Pearl Chiu, Chair
Brooks King-Casas
Thomas Ollendick
John Richey

March 16, 2018
Blacksburg, VA

Keywords: Depression, Reinforcement Learning, fMRI, Cognitive Behavioral Therapy, Computational Modeling, Computational Psychiatry

Abstract

Major depressive disorder is a common, impairing disease, but current treatments are only moderately effective. Understanding how processes such as reward and punishment learning are disrupted in depression and how these disruptions are remediated through treatment is vital to improving outcomes for people with this disorder. In the present set of studies, computational reinforcement learning models and neuroimaging were used to understand how symptom clusters of depression (anhedonia and negative affect) were related to neural and behavioral measures of learning (Study 1, in Paper 1), how these alterations changed with improvement in symptoms after cognitive behavioral therapy (Study 2, in Paper 1), and how learning parameters could be directly altered in a learning retraining paradigm (Study 3, in Paper 2). Results showed that anhedonia and negative affect were uniquely related to changes in learning and that improvement in these symptoms correlated with changes in learning parameters; these parameters could also be changed through targeted queries based on reinforcement learning theory. These findings add important information to how learning is disrupted in depression and how current and novel treatments can remediate learning and improve symptoms.

General Audience Abstract

Major depression is very common and current treatments are sometimes helpful and sometimes not. In order to create more effective treatments, we need to better understand what exactly goes wrong when people are depressed. The present set of studies uses computational modeling and imaging of brain function to gain a clearer understanding of how people with depression learn from rewarding and punishing events differently, how these differences in learning improve with symptom improvement after receiving treatment for depression, and how learning differences can be directly targeted by teaching people to learn differently. I found that a reduced ability to experience pleasure, or anhedonia, in depression was related to differences learning from good outcomes while low mood was related to perceiving bad outcomes as worse. Both of these differences improved with successful treatment, and asking people questions related to learning also changed the way people learned in a way that may be useful for improving treatments.

Acknowledgements

First of all, I'm grateful to my advisor, Pearl Chiu, for her guidance, support, and encouragement throughout this process. My growth as a scientist the past six years is due in large part to her advice and help along the way. I'm also thankful for Brooks King-Casas for his scientific insight and sarcasm, and for John Richey and Tom Ollendick for their perspectives and leadership throughout my dissertation progress. Thanks as well to George Clum and Russell Jones for conversations that guided my development and thinking.

I'm also grateful for the rest of the three musketeers, Katie and Lara, for our 'one for all and all for one' journey through grad school. Thanks to Lusha for teaching more more than I can remember, to John for his own brand of encouragement, to Dongil for his patience and attempts to laugh at my jokes, and to the rest of the CKC lab (actual and honorary) for being a great scientific community to develop in. Thanks to Merage, Andrew, Haley, Rachel, and Robert, and to Erin, Irene, Jessie, and Rachel, for being wonderful grad school and internship cohorts. And, of course, thanks to Ben for his unwavering support (including emergency ramen) all along the way.

Table of Contents

Abstract ii
General Audience Abstract iii
Acknowledgements iv
Table of Contents v
Introduction 1
Disrupted reward and punishment processing in depression 2
Computational models of learning 3
Computational model-based learning investigations in depression 5
Learning processes and treatment response in depression 7
Reinforcement learning and learning retraining 9
Present studies and hypotheses 11
Paper 1 (Studies 1 and 2): Neurobehavioral Evidence for Symptom- and Valence-Specific Reinforcement Learning Alterations in Depression and the Effects of CBT 13
Abstract 14
Introduction 15
Results 18
Discussion 25
Methods 30
References 45
Figures 59
Tables 74
Paper 2 (Study 3): Towards Learning Retraining: Creating Targeted Changes in Reinforcement Learning 81
Introduction 83
Results 85
Discussion 90
Methods 94
References 100
Tables 105
Figures 107
Discussion 116
Summary of studies 116
General discussion 118
Limitations and future directions 120
Conclusion 122
References 123

Introduction

Major depressive disorder (MDD) is a prevalent, disabling psychiatric disorder that currently is the highest cause of disability worldwide (World Health Organization, 2017). MDD is a highly heterogeneous disease (Rush, 2007) characterized by a constellation of symptoms related to low mood and anhedonia. Current treatments are only moderately effective, with response rates of 40-50% in clinical trials (Hollon, Stewart, & Strunk, 2006) and lower effectiveness in the community (K. A. Collins, Westra, Dozois, & Burns, 2004). Recent improvements in CBT and pharmacological treatments have improved dissemination or reduced side effects, but have not led to greater treatment response (Johnsen & Friborg, 2015). Differential treatment prediction and refinement of current treatments are hampered by the trial-and-error method by which current treatments have been discovered and refined. To better connect mechanisms of treatment to treatment approaches and to improve outcomes, a better understanding of disrupted processes underlying aspects of MDD and their relationship to treatment is needed. In the current work, learning from reward and punishment is examined as a potential process that is disrupted in depression and that can be remediated by current or novel treatments. To understand how disruptions in specific components of learning are related to symptom improvement, computational models of reinforcement learning are used to understand precise aspects of learning that are related to different symptom clusters in depression, how these disrupted learning components change with successful treatment with CBT, and how these components could be changed

Disrupted reward and punishment processing in depression

Potential candidate mechanisms in depression include disrupted cognitive and affective processing, both of which are affected by MDD. On 'cold' (i.e. lacking in affective content) tasks assessing executive function, people with depression show impairments in most domains (Snyder, 2013). However, some of these impairments in executive function are a similar magnitude in comparison with non-executive function tasks (e.g. Stroop neutral versus interference performance), suggesting that some of these impairments can be attributed to aspects of the disorder such as psychomotor slowing or other non-executive function impairments rather than pure executive function deficits. Tasks with affective content are more clearly related to impairments in task performance in people with MDD. People with depression show reduced responsiveness to rewarding stimuli (Henriques & Davidson, 2000; McFarland & Klein, 2009; Sloan, Strauss, & Wisner, 2001); Henriques and Davidson found that participants without depression changed their behavior on rewarding versus non-rewarding conditions of a verbal memory task to earn more reward while participants with depression did not, and McFarland & Klein and Sloan et al. found less emotional reactivity to pleasant stimuli in people with depression. People with MDD also show over responsiveness to punishment such that their performance falls drastically after errors (Elliott, Sahakian, Herrod, Robbins, & Paykel, 1997), a behavior termed catastrophic response to failure. These behavioral impairments are accompanied by altered neural processing in striatum, ventromedial prefrontal cortex, insula, and lateral cortical areas (Diener et al., 2012; Siegle, Thompson, Carter, Steinhauer, & Thase, 2007; Smoski et al., 2009; although these findings are not consistent across studies, Müller et al., 2016), areas involved in incorporating affective information during cognitive processing.

These altered cognitive and affective processes in depression likely lead to changes in expectations of, and experiences of, reward and punishment in daily life. In turn, these changes in expectations and experiences of affective feedback may explain the negative mood and anhedonia experienced by people with depression. Similar to work reviewed above that found difficulties in modulating behavior under rewarding conditions, studies of reward processing and anhedonia in depression have found an inability to change responses based on feedback to maximize reward (Pizzagalli, Jahn, & O'Shea, 2005). Although studies of responses after negative feedback found disrupted performance in depression, learning studies with punishment have often found facilitation of performance with increasing depressive symptoms (Beevers et al., 2013; Cavanagh, Bismark, Frank, & Allen, 2011; Maddox, Gorlick, Worthy, & Beevers, 2012). This discrepancy suggests that assessing overall performance through measures such as total number of correct responses may miss distinctions in behavior that correspond to affected domains in depression. Even when results are more consistent, changes in performance in people with MDD could be due to a variety of factors in how rewards and punishments are perceived and processed. This ambiguity necessitates further investigation into what aspects of reward and punishment processing are disrupted in depression and how these disruptions relate to symptoms.

Computational models of learning

Computational models of learning and decision making can provide insight into what aspects of these processes are disrupted in psychological disorders, an approach falling under the umbrella of computational psychiatry (Montague, Dolan, Friston, & Dayan,

2012). In particular, computational models of reinforcement learning can take advantage of the precision afforded by algorithmic formalization of learning processes along with advances in modern learning theory in animal research. Reinforcement learning models have been used to study a wide range of psychological phenomena and disruptions in psychiatric disorders (Chen, Takahashi, Nakagawa, Inoue, & Kusumi, 2015; Maia & Frank, 2011). The central idea of reinforcement learning is that values associated with stimuli (or actions) are dynamically updated based on the discrepancy between observed and predicted outcomes. In this way, then, learning about what is rewarding and punishing only occurs when outcomes deviate from what was expected. Learning therefore scales with the extent that predictions are violated; an expected positive or negative outcome causes no learning.

Reinforcement learning has its early roots in psychological research by Bush and Mosteller (1953), later updated by Rescorla and Wagner (1972) and others, as well as simultaneous work in computer science by Bellman (1952) and others, but this area has been rejuvenated in recent years by discoveries that reinforcement learning not only explains behavior in learning tasks but also describes neural aspects of learning well (Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997), suggesting that this framework represents both outward manifestations of behavior and the neural calculations that give rise to this behavior. Because of this tight connection between behavioral and neural levels, reinforcement learning can combine behavioral (Yechiam, Busemeyer, & Stout, 2004), large-scale neural activity as measured by fMRI (O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Pagnoni, Zink, Montague, & Berns, 2002), pharmacological (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006), and genetic (Doll,

Hutchison, & Frank, 2011; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007) approaches to understanding how people learn about value. Recent work has used reinforcement learning as a framework to extend to other psychological phenomena, such as self-control (Berkman, Hutcherson, Livingston, Kahn, & Inzlicht, 2017) and emotional regulation (Etkin, Büchel, & Gross, 2015), and to bring in complementary processes, such as attention (Niv et al., 2015) and working memory (A. G. E. Collins, Brown, Gold, Waltz, & Frank, 2014). Reinforcement learning's ability to account for behavior through the impact of positive and negative outcomes and through unifying multiple levels of analysis makes it a particularly useful tool for understanding how processing positive and negative events differently in psychopathology leads to differences in behavior.

Computational model-based learning investigations in depression

To study reinforcement learning alterations in MDD, specific learning model parameters are estimated from behavior and compared between groups or along a symptom dimension (although variability in which learning model is used can also be used as the dependent variable, e.g. Rigoux et al., 2013, and neural measures can also be used to estimate models alongside behavioral measures, e.g. Turner et al., 2013). These behavioral parameter differences are assumed to correspond to neural differences in processing expected value, prediction error, or the value of outcomes, among other measures, as measured by functional MRI (O'Doherty, Hampton, & Kim, 2007). Early computational work in reward learning disruptions in participants with a range of anhedonia or depressive symptoms distinguished between alterations in updating expectations of value (indexed by a learning rate parameter) and valuation of rewards (indexed by outcome

sensitivity or inverse temperature parameters) and found that valuation was reduced with increased depression severity (Kunisato et al., 2012) or anhedonia (Huys, Pizzagalli, Bogdan, & Dayan, 2013). Other contemporary findings in MDD found reduced prediction error in the striatum (Gradin et al., 2011; Kumar et al., 2008), which is also suggestive of reduced valuation during learning. However, behavioral differences in reinforcement learning parameters have not been found in all studies (Gradin et al., 2011; Rothkirch, Tonn, Kohler, & Sterzer, 2017) and neural results in larger clinical samples have found intact striatal prediction errors, unaffected by depression or anhedonia severity (Greenberg et al., 2015; Rothkirch et al., 2017; Rutledge et al., 2017). Some of these studies have found altered prediction error signals in areas outside the striatum, including orbitofrontal cortex (Rothkirch et al., 2017) and anterior cingulate cortex (Steele, Meyer, & Ebmeier, 2004), and studies with sufficient power to detect statistical moderations have found that anhedonia in depression may be related to a disconnect in the relationship between neural representations of expected reward and prediction error (Chase et al., 2013; Greenberg et al., 2015; see also Sherdell, Waugh, & Gotlib, 2012 for a similar behavioral finding). Therefore, a possible explanation for altered reward learning in depression may be that carrying signals of experienced reward forward to future encounters with similar rewards (perhaps due to increased noise, Robinson & Chase, 2017) is disrupted in anhedonic depression, but the specific disruptions in anhedonia and depression with reward learning are yet unclear.

The above work has primarily focused on reward processing, reflecting the general focus on reward learning in modern reinforcement learning studies; fewer studies have assessed punishment contexts in reinforcement learning in depression. Behavioral studies

of punishment learning, introduced above, have sometimes found improved performance in depression, reflected in parameter differences indicating reduced noise or greater exploitation (Beevers et al., 2013); however, other reinforcement learning studies have found no differences during punishment learning (Rothkirch et al., 2017).

In addition to examining both reward and punishment, focusing on specific symptom clusters of depression (e.g. anhedonia in reward learning and negative affect in punishment learning) may better correspond to learning differences than the presence of a depression diagnosis. Several of the studies reviewed above (Greenberg et al., 2015; Huys et al., 2013; Rothkirch et al., 2017) found effects specific to anhedonia symptoms, rather than overall depression diagnosis or severity, in reward learning. Related studies have found similar anhedonia-specific reward processing problems with computational studies of learning or reward processing (Harlé, Guo, Zhang, Paulus, & Yu, 2017; Luking, Pagliaccio, Luby, & Barch, 2015; Young et al., 2016), some of which have also found learning-related problems in negative contexts that correspond better to negative affect symptom severity (Luking et al., 2015).

Learning processes and treatment response in depression

If depression is mechanistically related to problems learning from and processing rewards and punishments, these differences should predict or correlate with treatment success. Cognitive-behavioral therapy (CBT), in particular, may relate to altered learning processes due to its roots in learning theory and its focus on new learning (Disner, Beevers, Haigh, & Beck, 2011; Jacobson, Martell, & Dimidjian, 2006). Computational work in reinforcement learning has not yet investigated the effects of treatment in depression, but

related work suggests potential relationships. Studies of changes in reward processing with CBT have found changes in neural processing of rewards (Dichter et al., 2009); improvement in positive affect after treatment with selective serotonin reuptake inhibitors also correlates with ability to sustain frontostriatal reward connectivity (Heller et al., 2013). However, these are preliminary studies with small sample sizes, so the effects of treatment on reward function are not yet well characterized. Reward processing has also been found to predict response to treatment, although the direction and nature of effects differ by study. Better ability to sustain neural response to rewards (Carl et al., 2016) and better behavioral reward learning (Vrieze et al., 2013) predict better response to treatment, while reduced neural reactivity to rewards (Burkhouse et al., 2016), lower reward circuitry connectivity (Downar et al., 2013), and disrupted Pavlovian withdrawal (Huys et al., 2016) have been found to predict worse outcomes. Taken together, these studies suggest that altered reward processing may predict and index symptom improvement with treatment, but it is unclear precisely how they relate; the specificity of computational approaches may clarify some of these findings.

The relationship of punishment learning to treatment has been little studied, but studies of processing of other negative stimuli implicated in depression and their relationship to treatment outcomes may inform treatment studies using reinforcement learning models. However, these findings are also somewhat contradictory. Several studies examined neural reactivity to emotional stimuli (words or pictures) prior to treatment as a potential predictor of treatment response. Reduced responsivity to negative words in cognitive control networks predicts better response to antidepressants (Miller et al., 2013) and, similarly, increased responsivity in similar areas predicts worse response (Delaveau et

al., 2015) to antidepressants; however, other studies found greater responsivity in cognitive control areas (anterior cingulate cortex) and nearby valence-related areas (ventromedial prefrontal cortex) to predict more improvement (Davidson, Irwin, Anderle, & Kalin, 2003; Ritchey, Dolcos, Eddington, Strauman, & Cabeza, 2011). Therefore, this lack of clarity in predicting or correlating with treatment is not specific to reward learning studies; studies of other cognitive processes have similarly mixed findings and point to the need for more mechanistically specified studies informed by computational process models.

Reinforcement learning and learning retraining

In addition to unraveling how computational models of learning can inform current treatment approaches, generative models such as reinforcement learning can provide insights into how mechanisms of disorders can be directly targeted. Such retraining approaches are similar to cognitive training approaches in areas such as recovery from stroke and in schizophrenia (Keshavan, Vinogradov, Rumsey, Sherrill, & Wagner, 2014), which remediate certain cognitive functions as a way to improve symptoms. Retraining approaches also share features with strongly behavioral approaches in CBT such as exposure for anxiety and posttraumatic stress disorder (Asnaani, McLean, & Foa, 2016) and applied behavioral analysis for autism (Foxx, 2008), which seek to target specific aspects of these disorders in order to create overall improvement in the disorder. A number of retraining interventions have been piloted in depression, with the most commonly studied retraining paradigms including training attention away from negative material (Cooper et al., 2014), increasing the concreteness of thinking (Watkins, Baeyens, & Read, 2009), and

training more positive predictions of outcomes (Collier & Siegle, 2015). However, meta-analyses and large-scale RCTs find little to no effect of these paradigms (Blackwell et al., 2015; Cristea, Kok, & Cuijpers, 2015; Hallion & Ruscio, 2011) on depressive symptoms, despite the success of retraining paradigms in other disorders (Hallion & Ruscio, 2011; McGurk, Twamley, Sitzer, McHugo, & Mueser, 2007). However, other retraining approaches show some effectiveness, with cognitive retraining focused on non-affective processing showing significant effects on attentional control, working memory, overall functioning, and depressive symptoms (Motter et al., 2016). However, this sort of executive function training may not affect problems in affective processing that appear to be more prevalent and problematic in depression. Understanding the processes being targeted by retraining approaches is necessary to ensure retraining effectively remediates disrupted processes and leads to symptom improvement.

Retraining approaches to changing altered learning processes in disorders like MDD can benefit from reinforcement learning's ability to represent processes underlying learning; as a result, using reinforcement learning to develop learning retraining interventions may lead to more robust and precise changes than other approaches. This approach requires understanding which aspects of reinforcement learning are disrupted in depression. With this knowledge, disrupted parameters or other model components can be targeted by learning-related approaches to determine which approaches best change disrupted learning, with the intention of then deploying this intervention at a clinically meaningful intensity. Ensuring learning generalizes outside of the training environment is also a vital component of retraining approaches (Keshavan et al., 2014; Swan, Carper, & Kendall, 2016), and so retraining approaches using reinforcement learning should also

ensure participants are able to generalize the intervention to apply to learning situations in real life.

Present studies and hypotheses

The present work intended to investigate how computational models of reinforcement learning can inform knowledge of the mechanisms of depression and its treatment. Study 1 (Paper 1) aimed to illustrate altered learning from rewards and punishments on behavioral and neural levels in a large, well-characterized sample of adults with and without depression. Study 2 (Paper 1) then examined how these alterations correlated with and predicted response to cognitive behavioral therapy for depression. Study 3 (Paper 2) used a reinforcement learning framework to see if behavioral parameters of a reinforcement learning model could be altered in a retraining approach through targeted queries about participants' learning.

Study 1 used a reward and loss learning task intended to measure reinforcement learning (after Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006) in participants with and without depression and used behavioral and neural approaches to characterize altered learning with different symptom clusters of anhedonia and negative affect. We hypothesized that 1) depression, particularly symptoms of anhedonia and negative affect, would correspond to altered behavioral patterns as indexed by reinforcement learning parameters and 2) that these altered behavioral patterns would be accompanied by related neural alterations in processing value and value update signals from the reinforcement learning model.

Study 2 examined learning on the same reward and loss learning task after a subset of participants with depression from Study 1 completed a standard, manualized course of cognitive behavioral therapy in order to measure predictors and correlates of treatment response. We hypothesized that the learning alterations found at baseline in Study 1 would predict and correlate with symptom improvement after CBT.

Study 3 used a similar reward learning task with the addition of queries on specific learning components to test if querying participants would lead to changes in learning. Participants in this study were recruited from a large online study with a range of depressive symptoms to determine effects of these queries in general and to ensure similar effects were found in participants high in depressive symptoms. We hypothesized that querying about task components related to aspects of learning would change these facets of behavior while leaving other learning processes intact. Specifically, we hypothesized that querying about prediction error or components of prediction error would increase learning rate and that querying about the probability or value of outcomes would increase outcome sensitivity while querying about the differences in value between options would decrease outcome sensitivity.

Paper 1 (Studies 1 and 2): Neurobehavioral Evidence for Symptom- and Valence-Specific Reinforcement Learning Alterations in Depression and the Effects of CBT

Vanessa M Brown, Lusha Zhu, Alec Solway, John M Wang, Katherine L McCurry, Brooks King-Casas, & Pearl H Chiu

Abstract

Major depressive disorder (MDD) is prevalent and impairing, but a limited neuromechanistic understanding of the disorder hinders therapeutic progress. In participants with and without depression, we demonstrate that a computational psychiatry approach may provide an increased mechanistic understanding of MDD and identify novel intervention targets. We first show that the cardinal symptoms of anhedonia and negative affect uniquely relate to distinct features of reward and loss learning, respectively. In depression, greater anhedonia, and not negative affect, was related to computational model-derived learning parameters and neural learning signals in the ventral striatum; these patterns were observed only during reward learning. Only during loss learning, increased negative affect, and not anhedonia, was related to learning parameters and disrupted subgenual anterior cingulate cortex encoding of learning signals. Second, the translational potential of these learning substrates was assessed in a subset of participants who received cognitive-behavioral therapy (CBT). Symptom improvement following CBT was related to normalization of learning parameters. These results suggest the utility of neurocomputational psychiatry for revealing covert mechanistic features and new therapeutic targets of depression.

Introduction

Major depressive disorder (MDD) is a prevalent, disabling disease affecting approximately 7% of people in the US each year (Kessler et al., 2005) and is the highest cause of disability in the world (World Health Organization, 2017). Barriers to characterizing and treating MDD include challenges in connecting symptom heterogeneity to distinct neural and behavioral substrates (Stephan et al., 2015). Through algorithms linking behavioral and neural levels of dysfunctional processes, neurocomputational approaches to psychiatry have the potential to illuminate novel mechanisms, connect these mechanisms to symptoms and subtypes within disorders, and test and refine treatments (Montague et al., 2012; Wang & Krystal, 2014; Wiecki et al., 2015). For depression, recent advances in understanding the basic neurocomputational substrates of reinforcement learning (Eshel & Roiser, 2010; Schultz et al., 1997; Sutton & Barto, 1998) provide a framework for linking the central impairments of anhedonia and negative affect (Clark & Watson, 1991) with mechanistically-defined learning parameters to both characterize depression and test the translational potential of computational psychiatry for depression. To that end, we sought to delineate the relationships among mechanistic components of learning and canonical depression symptoms, and to test the translational potential of these relationships through the responsiveness of learning components to cognitive-behavioral therapy, an efficacious, learning-based psychotherapy for depression.

Computational formalizations of reinforcement learning characterize a process whereby expectations about outcomes are updated based on prediction errors (i.e., differences between expected and actual outcomes) which provide a causal connection between neural

and behavioral substrates of value-based learning (Pessiglione et al., 2006; Schultz et al., 1997; Steinberg et al., 2013). In this framework, disrupted learning in depression can be separated into computationally-derived, behaviorally- and neurobiologically- distinct, components (e.g., outcome valuation vs. updating and their respective neural substrates; Chen et al., 2015; Huys et al., 2013; Robinson & Chase, 2017). For example, disrupted value-based learning may result from impaired updating of expected outcomes (quantified in RL models as learning rates), an inability to distinguish among levels of outcomes (quantified in RL models as outcome sensitivities), or shifts in how outcomes are valued (quantified in RL models as outcome shifts); such components are differentiable and quantified as learning parameters which can then be related to symptoms at the individual level. In a parallel literature, relationships between depression and abnormal responses to rewards and losses have been consistently reported (Eshel & Roiser, 2010) and appear to differ by valence such that individuals with depression show neural and behavioral overreactions to failure (Chiu & Deldin, 2007; Elliott et al., 1997) coupled with diminished modulation of behavior in response to reward (Pizzagalli et al., 2008; Robinson & Cools, 2012) but intact representation of reward independent of learning (Rutledge et al., 2017). Accumulating evidence suggests that symptoms of anhedonia and negative affect better correspond to these valence-specific neurobehavioral disruptions in reward and loss learning than do overall depression severity or the presence of a depression diagnosis (Disner et al., 2011; Harlé et al., 2017; Luking et al., 2015; Rothkirch et al., 2017; Young et al., 2016). Together these data argue for a context-dependent role of learning disruptions in depression. However, reports of overall learning differences in depression do not distinguish which aspect of learning is disrupted; hypotheses which can be quantitatively disambiguated via

computational models of learning. In turn, a computationally specific description of disrupted learning in depression would allow precise mapping across levels (similar to Anticevic et al., 2015) and enable more targeted treatments.

To date, some computational approaches applied to learning in depression have incorporated differences in learning from rewards versus punishments (Beevers et al., 2013; Cavanagh et al., 2011; Kunisato et al., 2012; Maddox et al., 2012; Robinson & Cools, 2012; Rothkirch et al., 2017), while other studies suggest that anhedonia symptoms better relate to neural and behavioral learning differences than depression as a whole (Chase et al., 2010; Gradin et al., 2011; Greenberg et al., 2015; Huys et al., 2013; Kumar et al., 2008; Pizzagalli et al., 2008). These studies provide a foundation for the present study that encompasses valence-specific effects of reward and loss learning, symptom-specific effects of anhedonia and negative affect, and individual estimates of behavior and neural activity, to fully explain how disrupted learning processes in MDD.

Beyond the ability to precisely define value-related learning disruptions and their connections to symptoms of depression, reinforcement learning models have the translational potential to assess mechanisms of change with treatment and thus identify potential new targets of therapy. As potential endophenotypes of depression (Pizzagalli, 2014), reward and loss learning alterations connect to neurobiological and behavioral targets of interventions and may index changes in treatment targets with greater precision. Cognitive-behavioral therapy, which targets altered valuation and updating (Beck, 2005) and is based on related learning theories that gave rise to computational formulations of reinforcement learning (Lewinsohn, 1974; Rescorla & Wagner, 1972), offers a particularly

clear connection between changes in learning and improvements in symptoms (see also Burkhouse et al., 2016; Dichter et al., 2009; Heller et al., 2013; McCabe et al., 2010; Pizzagalli, 2014; Vrieze et al., 2013 for initial studies examining responsivity to rewards and losses and treatment effects).

In the current study, we examined clinically depressed and non-depressed participants with a range of symptoms, and also a subset of depressed participants who were tested before and after treatment with CBT, on a reward and loss learning task while undergoing functional magnetic resonance imaging (fMRI) scanning. We hypothesized that distinct learning patterns in reward and loss contexts, captured by computational model-derived parameters measuring aspects of updating and valuation and their corresponding neural signals, would be related to symptoms of anhedonia and negative affect, respectively, on behavioral and neural levels. Additionally, we hypothesized that these learning alternations in reward and loss domains would predict and correlate with improvement in related symptoms following cognitive-behavioral therapy.

Results

Participant characteristics and model-free learning at baseline

One hundred and one participants (69 participants with depression and 32 without) were included in baseline analyses; clinical and demographic data are reported in Table 1.1. Depressed participants comprised 63 with current MDD, 3 with current dysthymia, and 3 with both MDD and dysthymia. As expected, participants with depression had higher symptom scores on the Beck Depression Inventory (BDI) as well as Mood and Anxiety

Symptom Questionnaire (MASQ) scales of anhedonia, negative affect, and anxious arousal; the groups did not differ on estimated IQ, age, or gender.

Participants completed a probabilistic learning task with alternating blocks of trials with rewarding and losing outcomes, respectively, while undergoing fMRI scanning (Figure 1.1a); behavior was analyzed separately for reward and loss trials to allow separate examination of reward and loss learning. Participants showed intact learning on the task; participants with depression had somewhat reduced reward learning (average [SE] performance on reward trials: controls 74.5% (2.0%) and MDD 69.0% (1.6%), $t_{100} = 1.98$, $p = .05$; Figure 1.1b) but showed similar loss learning (average [SE] performance on loss trials: controls 70.8% (1.9%) and MDD 68.9% (1.4%), $t_{100} = 0.80$, $p > .1$).

Reinforcement learning model fits participants' behavior well

To ensure the reinforcement learning model described participants' behavior accurately, the proposed model was compared against possible alternative models and assessed for its suitability in measuring distinct aspects of behavior. Participants' behavioral choices on the probabilistic learning task were fit to a computational reinforcement learning model that differentiated components of learning potentially disrupted in depression (i.e., updating and valuation). A model with a learning rate parameter, indexing speed of updating based on prediction error, and two valuation-related parameters of outcome shift (linearly shifting all outcome values, resulting in an overall positive or negative valuation bias) and outcome sensitivity (multiplicatively scaling more extreme outcome values, resulting in differential scaling of values) fit participants' behavior better than plausible alternative models (Daw, 2011; Huys et al., 2013; Figure 2a). Reflecting the separation of reward and

loss learning in the task, a model with unique parameters in reward and loss conditions improved model fit over combining some or all parameters across conditions (Figure 1.2c). Model validation methods including parameter recovery (Figure 1.2b) and inspection of posterior probability distributions (Figure 1.3) indicated that model parameters were recoverable and uniquely identifiable.

To examine the relationships between learning parameters and model-agnostic measures of behavior, individual estimates for values of each learning parameter, separated by reward and loss learning conditions, were extracted and compared to summary behavioral measures. Plotting of individuals' parameter values against the total proportions of correct choices and switches confirmed expected relationships between these model-based and model-agnostic measures (Figure 1.4). Learning rate was moderately related to performance, with low learning rates, reflecting a slower acquisition of contingencies, related to lower proportion of correct choices (gain: $r^2 = .097$; loss: $r^2 = .079$) and more switches (gain: $r^2 = .122$; loss: $r^2 = .056$). Outcome sensitivity, which indexes the scaling between large and small outcomes, was related to the ability to pick the stimulus more likely to lead to a better outcome (i.e. proportion correct choices; gain: $r^2 = .613$; loss: $r^2 = .690$) and to a reduced tendency to switch options (gain: $r^2 = .591$; loss: $r^2 = .404$), as would be expected with increased ability to differentiate among outcomes. Meanwhile, outcome shift, which indexes an overall shift in outcome values, was less related to the proportion of correct choices (gain: $r^2 = .0001$; loss: $r^2 = .117$) but was related to a reduced tendency to switch with higher valuation of outcomes (gain: $r^2 = .169$; loss: $r^2 = .560$), reflecting its effect on the tendency to switch away from all choices when the associated outcomes are perceived as more negative and to stay with all choices when outcomes are viewed as

positive. We note each parameter's unique relationship with traditional summary statistics; these relationships illustrate how differences in learning strategies, indexed by combinations of increases and decreases in model-derived parameter values, may be obscured when viewing summary statistics only.

Anhedonia symptoms are related to parameters of reward learning in depression

Having established the appropriateness of the model for describing participants' behavior based on model fit and comparisons to model-agnostic summaries of behavior, we then tested the relationships among the model-derived learning parameters (i.e., learning rate, outcome sensitivity, and outcome shift) and symptom clusters of anhedonia, negative affect, and overall depression during reward and loss learning, respectively. During reward learning, in participants with depression, greater anhedonia was related to reduced learning rate (mean transformed/untransformed effect of anhedonia = -0.136/-0.739; 95% credible interval of -0.214 to -0.023/-1.39 to -0.109) and greater outcome sensitivity (mean effect of anhedonia = 0.181; 95% credible interval of 0.016 to 0.377; Figure 1.5a); these effects were unique to anhedonia and absent with negative affect or overall depression severity. These results indicate that, with increasing anhedonia, people with depression are slower to update expectations of reward based on prediction errors, but accentuate the rewarding effect of large rewards.

Neurally, previous work with anhedonia in depression has suggested a relationship between lower learning rate and reduced correlation between signaling of prediction error (when receiving information about outcomes) and expected value (when shown options to select; Greenberg et al., 2015); we found similarly intact signaling of both prediction error

and expected value in participants with depression (all ps <.05, corrected; Figure 1.6 and Tables 1.2 and 1.3; see Tables 1.4 and 1.5 for prediction error and expected value signaling in controls), but a disrupted correlation between the two signals with increasing anhedonia (t_{97} = -2.10, p < .05; Figure 1.5b).

Negative affect symptoms are related to parameters of loss learning

During loss learning, a distinctly different pattern of relationships emerged such that severity of negative affect was related to more negative outcome shift, indicating more negative valuation of all outcomes (mean effect of negative affect = -0.107; 95% credible interval of -0.205 to -0.013; Figure 1.5c). This relationship was specific to negative affect and not observed with anhedonia or overall depression severity.

Neurally, prediction error activity in a subgenual anterior cingulate cortex region of interest showed a negative relationship with negative affect (r = -.277, p = .005; Figure 1.5d). As prefrontal cortical signals are linked more to value representation than prediction error itself (Gläscher et al., 2009; Niv, 2009), the value-related components of prediction error (i.e., 'expected' value and 'actual' value received) were investigated to determine if either explained the relationship between reduced prediction error signaling in subgenual anterior cingulate and negative affect. This analysis revealed greater negative affect was related to more negative signaling (less modulation) of 'actual' outcome value in ventromedial prefrontal cortex and precuneus (p < .05 corrected; Figure 1.7a and Table 1.6), with no significant relationship with 'expected value' either at time of cue onset or outcome. This result suggests that participants greater in negative affect differ in neural processing of actual outcome values, with subsequent effects on prediction error, rather

than in processing expected value. Further inspection of the neural activity modulated by outcome value in participants with high negative affect (based on a median split) revealed significant neural processing of outcome value, with activation negatively related to the level of outcome value in dorsomedial prefrontal cortex and insula, but no positive relationship between outcome value and vmPFC signal ($p < .05$ corrected; Figure 1.7b and Tables 1.7 and 1.8), indicating these participants engaged a different, more negatively-valenced network of brain areas rather than the positively-valenced regions of vmPFC and striatum engaged by low negative affect participants.

Reinforcement learning parameters show selective remediation with symptom changes after treatment and predict changes in anhedonia

The specificity of reward and loss learning differences to specific symptom clusters suggested the translational potential of neurocomputational approaches beyond descriptive explanation for depression. In this case, indices of altered learning should be sensitive to treatment. To test this possibility, we investigated whether changes in measures of reinforcement learning were related to symptom changes with cognitive-behavioral therapy and if pre-treatment measures predicted symptom changes after CBT. As CBT targets reduced engagement in positive activities and biased interpretations of negative outcomes through changes in cognition and behavior, this treatment approach is well positioned to investigate changes in learning positive and negative information. Participants completed a standard, manualized (Munoz & Miranda, 1996) course of weekly CBT. After treatment, participants showed large average decreases in all symptom measures (overall depression severity: average decrease of 57%, $t_{27} = 8.74$, $p<.001$;

anhedonia: average decrease of 24%, $t_{27} = 8.49$, $p<.001$; negative affect: average decrease of 29%; $t_{27} = 7.31$, $p<.001$), indicating that as expected, CBT was effective. Consistent with the literature (Dobson et al., 2008), participants also showed considerable variation in treatment response; this heterogeneity enabled an investigation of individual differences in the degree of symptom change (Figure 1.8a).

Correlates of reward learning changes. In depressed participants who underwent CBT, reward learning rate, which had been negatively correlated with anhedonia at baseline, significantly increased with improvement in anhedonia during treatment (mean transformed/untransformed change with percent change in anhedonia: 0.255/1.16; 95% credible interval of 0.138 to 0.344/0.577 to 1.79). Meanwhile, reward outcome sensitivity, which had been positively correlated with anhedonia at baseline, significantly decreased with anhedonia improvement (mean change with percent change in anhedonia: -0.512; 95% credible interval of -0.771 to -0.281; Figure 1.8b). Similar effects were also seen with changes in overall depression as well as negative affect; note that within-participant correlations in decreases in symptom clusters during treatment were high (correlation between percent change in anhedonia and negative affect: $r = .651$, $p < .001$, and between anhedonia and overall depression severity: $r = .634$, $p < .001$).

Neurally, treatment was related to a significant change in the correlation between ventral striatum signaling to prediction error and expected value in participants with high anhedonia pre-treatment (Fisher's r to z = 1.65, $p < .05$ one-tailed; Figure 1.9a). Examining pre-treatment neural signaling in this same ventral striatum ROI to expected value and prediction error showed that both neural signals correlated with improvement in

anhedonia symptoms, such that more negative expected value ($r = -.358$, $p < .05$) and more positive prediction error ($r = .348$, $p < .05$) were related to greater improvement in anhedonia (Figures 1.9c and 1.9d).

Correlates of loss learning changes. In loss, the behavioral measure of outcome shift, which had been negatively related with negative affect at baseline, showed significant increases with improvement in negative affect (mean change with percent change in negative affect: 0.393; 95% credible interval of 0.244 to 0.564), while outcome sensitivity showed significant decreases (mean change with percent change in negative affect: -1.17; 95% credible interval of -1.73 to -0.652; Figure 1.8c; note this parameter had a trend-level positive relationship with negative affect at baseline). However, pre-treatment prediction error signaling in subgenual anterior cingulate cortex during loss learning did not predict change in negative affect (Figure 1.9b), nor did this neural signal significantly change pre- to post-treatment (all ps > .05).

Lack of practice effects in control participants. To ensure changes with treatment were not related to practice effects, we estimated changes in control participants' behavior at both time points. Control participants did not show changes in any behavioral parameter between time points (all 95% credible intervals encompassing 0; Figure 1.10), indicating a lack of systematic practice effects in the absence of clinical changes.

Discussion

A large body of literature has suggested relationships among learning about rewards and losses and symptom dimensions of anhedonia and negative affect, but the precise neurobehavioral relationships among these components have been unclear and difficult to

relate to translational targets. In the current study, we used a computational model of reinforcement learning to distinguish among learning processes and showed, across neural and behavioral levels, the specific relationships of i) anhedonia to reduced updating of rewards and ii) negative affect to more negative valuation of losses. Following treatment with cognitive behavioral therapy, symptom change was related to normalization of these learning differences; additionally, neural measures of reward learning predicted change in anhedonia with treatment.

These results illustrate how a neurocomputational approach to psychiatric disorders can provide precision and specificity in elucidating disease mechanisms, indexing treatment response, and suggesting novel therapeutic targets. Differences in valuation provide one possible explanation for learning differences in depression, with depression related to a tendency to perceive all outcomes more negatively than they actually are (Elliott et al., 1997). We find support for this explanation, but only during loss learning and specifically related to symptoms of negative affect; participants higher in negative affect showed more negative valuation of outcomes behaviorally, a reduced prediction error signal in subgenual anterior cingulate, and a shift in processing the values of outcomes from positive value-related brain areas, including ventromedial prefrontal cortex and precuneus, to brain regions involved in negatively-valenced processing, including insula and dorsomedial prefrontal cortex. After CBT, improvements in negative affect correlated with increased loss valuation, suggesting that reductions in negative affect are associated with viewing negative outcomes more positively. Of interest, the network of brain areas modulated by negative affect during processing of outcome values found here (subgenual ACC, insula, and dorsomedial prefrontal cortex) shows extensive overlap with areas previously identified

with negative self-focus, rumination, and cognitive disruptions in depression, primary targets of CBT (Disner et al., 2011; Mayberg, 2003; McTeague et al., 2017; Sheline et al., 2010). The possible connection between these cognitive aspects of depression and overly negative valuation during loss learning deserves further study. In contrast, during reward learning, behavioral valuation of outcomes was intact and neural signals of expected value and prediction error were unaffected. This specificity to losses is in line with maladaptive responses to negative, and not positive, feedback previously observed in depression (Chiu & Deldin, 2007; Elliott et al., 1997). Our results extend this previous work to show that this exaggerated response to punishment in depression is not due to direct differences in over-adjustment after negative feedback *per se*, as would be shown by alterations in learning rate, but rather due to valuing all negative feedback more strongly. In turn, this enhanced valuation of negative information would lead to exaggerated behavioral adjustments to avoid these negative outcomes.

Meanwhile, under-adjustment following outcomes, another possible explanation for altered learning in depression, was supported during reward learning and was shown to be specific to anhedonic symptoms. The relationship between anhedonia and reward learning has been intensely studied in recent years, with equivocal findings regarding the extent of behavioral and neural alterations during reward learning in depression (Chase et al., 2010; Chen et al., 2015; Gradin et al., 2011; Rothkirch et al., 2017). As suggested by recent large studies in depression (Greenberg et al., 2015; Rutledge et al., 2017), the effect of anhedonia on reward processing may be more complex than a simple deficit in value representation. In the present comparably large sample of participants with depression, we also found no support for reduced valuation of rewards either behaviorally or neurally, but rather a

moderation of expected value – prediction error correlations by anhedonia accompanied by a behavioral reduction in learning rate. Interestingly, this reduced learning rate was only present within participants with depression, suggesting that reward learning deficits are specific to clinically impairing levels of anhedonia. In depressed participants, then, our behavioral and neural findings show that the effect of anhedonia is not a straightforward dysfunction in representing the actual or expected values of behaviors, but rather in adaptively updating these values to adjust behavior to maximize reward. This inability to effectively use reward-related information likely underlies reported difficulties in seeking pleasurable activities as well as modulating activities to boost internal mood states, e.g. Taquet et al., 2016.

Symptom change with CBT correlated with successful remediation of alterations in both reward and loss learning. Specifically, greater improvement in anhedonia was accompanied by increased learning rate, reduced outcome sensitivity, and a normalized relationship between neural signals of expected value and prediction error, all during reward learning; greater improvement in negative affect was accompanied by more positive outcome shift and decreased scaling of outcomes, during loss learning. The two primary components of cognitive behavioral therapy involve challenging and reappraising negative evaluations and engaging in and reflecting on pleasurable activities (Beck, 2008; Munoz & Miranda, 1996). These treatment targets have clear connections to the learning changes seen here with CBT: more positive valuation of negative outcomes matches well with a greater ability to reappraise negative thoughts, and greater modulation of behavior to achieve rewards with an emphasis on reflecting on rewarding experiences. An improved mechanistic understanding has the potential to enable tailored novel treatment approaches based on

individualized patterns of learning alterations (e.g., focusing on updating reward expectations in patients high in anhedonia versus more positive valuation of negative outcomes in patients high in negative affect), targeting learning disruptions more directly through retraining paradigms (Keshavan et al., 2014), and in refining techniques in CBT and other treatments to more precisely change learning. Our findings also add to a growing body of literature on the ability of neural mechanisms of reward learning to predict response to CBT in depression (Burkhouse et al., 2016; Carl et al., 2016), indicating that a certain pattern of reward learning pre-treatment, characterized by strong neural responses to expected value and prediction error, may enable anhedonic patients to receive the most benefit from therapy. Although inferring a causal relationship between CBT strategies and the learning changes seen here requires further investigation, these results are an encouraging step toward establishing a computationally-informed mechanism of this learning-based treatment (Paulus et al., 2016).

In addition to enabling new treatment approaches and targets, the present results provide support for possible etiologies of depression. The experience of stress, a major risk factor for depression, selectively impairs reward learning (Berghorst et al., 2013; Bogdan & Pizzagalli, 2006; Hanson et al., 2015), disrupts the relationship between anticipatory and consummatory reward responses in ventral striatum (Kumar et al., 2014), and disrupts dopaminergic pathways in ventral striatum (Francis & Lobo, 2016), suggesting that response to stress may underlie anhedonia and altered reward learning in depression (Pizzagalli, 2014). Meanwhile, serotonin function modulates learning from losses (Cools et al., 2008; Faulkner & Deakin, 2014), particularly in dorsomedial prefrontal cortex (Robinson et al., 2013), an effect that is exaggerated in depression but eliminated with

successful SSRI treatment (Herzallah et al., 2013), indicating that altered serotonergic function may relate to altered processing of outcomes with negative affect.

In depression, the relationship between anhedonia and altered reward processing on one hand, and between symptoms relating to negative mood and differences in processing losses on the other, have long been hypothesized but difficult to dissociate. By computationally assessing reinforcement learning in both reward and loss contexts in a large, well-characterized sample using both neural and state-of-the-art behavioral analyses, the current findings show that these relationships are indeed present and both context- and symptom- specific. Remediation of these differences along with symptom change after treatment provides further support that these learning differences index state-dependent characteristics of the disorder. These results provide a computationally formalized framework to understand altered processing of positive and negative information in depression and suggest novel pathways for understanding and treating this common and debilitating disorder.

Methods

Study design & participants

The current study used a neurocomputational psychiatry approach to: (i) characterize neural and behavioral differences in reward and loss learning across depression symptom clusters in participants with and without depression, (ii) test whether learning differences changed with changes in symptoms after treatment, and (iii) test if aspects of learning prior to treatment predicted changes in symptoms after treatment. To achieve these goals, we fit behavioral and neural data to formal reinforcement learning models and tested for (i)

differences in aspects of reinforcement learning across symptom clusters, (ii) changes in aspects of reinforcement learning correlated with changes in symptoms with treatment, and (iii) correspondence between pre-treatment aspects of reinforcement learning and changes in symptoms with treatment. The sample size was selected based on (and generally exceeded) previous work with similar paradigms and populations (Chiu et al., 2008; Fu et al., 2004; Greenberg et al., 2015; King-Casas et al., 2008; McGrath et al., 2013).

Participants were recruited via community advertisements from the southwest Virginia and Houston, Texas areas. Participants were initially screened via phone to determine eligibility and, if eligible, completed study procedures in person. Study procedures were approved by the Institutional Review Boards of Baylor College of Medicine and Virginia Tech and all participants provided informed consent. Participants were required to have a primary DSM-IV diagnosis of Major Depressive Disorder or Dysthymia (diagnosed using the Structured Clinical Interview for DSM-IV, SCID; First et al., 1996), for participants with depression, or for control participants, no history of depression or other psychiatric disorders. Inclusion criteria for all participants included: age 18 to 64, English speaking, normal or corrected to normal vision, verbal IQ greater than 80, no contraindications to MRI scanning, no loss of consciousness greater than 30 minutes, no hormonal disorders, no behaviors meeting criteria for substance abuse or dependence (excluding nicotine dependence) in the past 30 days, and no current or past psychotic or bipolar disorders. Clinical and demographic data for participants are reported in Table 1.1. To be included in the present analyses, participants were required to demonstrate engagement on the behavioral task and successfully complete fMRI scanning (see Procedures below for further description of exclusion criteria). Sixty-nine participants with depression and 32 controls

were included in baseline (i.e., pre-treatment) analyses (total N = 101). Excluded participants did not differ from those included at baseline by age; gender; depression diagnosis; severity of depression, anhedonia, or negative affect; education or income level; ethnicity; marital status; or use of psychotropic medications; but had lower estimated IQ.

Self-report measures

In addition to the SCID, participants also completed self-report measures including the Beck Depression Inventory-II (BDI; Steer et al., 1999) to assess overall depression severity, the Mood and Anxiety Symptom Questionnaire (MASQ; Watson et al., 1995) to assess severity of symptom clusters of anhedonia (anhedonic depression subscale) and negative affect (general distress subscale), the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001) to estimate verbal IQ, and a demographics questionnaire. Participants with a SCID diagnosis of depression were required to have a BDI score greater than 12 on the day of the baseline scan, while controls were required to have a BDI score less than 13 (Dozois et al., 1998). Consistent with previous reports (Buckby et al., 2007; Watson et al., 1995), symptom severity measures were modestly related within participants with depression (R^2 values of .13 to .32) and more strongly related across all participants (R^2 values of .65 to .81), indicating that these measures mapped onto distinct constructs, particularly in participants with clinically elevated symptoms.

Reinforcement learning task

Participants completed a probabilistic reward and loss learning task (Figure 1.1a; similar to Pessiglione et al., 2006) while undergoing functional MRI scanning. The task was presented

in pseudo-randomized blocks of trials consisting of all reward outcomes or all loss outcomes. On each trial, participants were presented with two abstract stimuli. After choosing a stimulus, the chosen option was highlighted for a brief period and then an outcome (monetary reward or loss) was shown. For reward blocks, worse outcomes ranged from +$0.20 to +$0.30 and better outcomes ranged from +$0.70 to +0.80; for loss blocks, the worse outcomes were -$0.70 to -$0.80 and the better outcomes were -$0.20 to -$0.30. Participants were only instructed that 'one picture is always better than the other'; unknown to them, the structure of the task was that one stimulus had a 75% chance of leading to the better outcome and a 25% chance of leading to the worse outcome, with opposite probabilities for the other stimulus. An adaptive design titrated task difficulty, such that a block ended when 7 of the last 10 choices were the correct stimulus (with correct defined as the stimulus more likely to lead to the better outcome); additionally, the first block within each condition was required to be at least 15 trials long. The same stimuli were used for all trials within each block, and new stimuli, requiring new learning, were used for each new block. The task ended when participants completed at least 50 trials total and at least 25 correct trials (average [SE] number of trials for depressed and control groups were reward: depressed 53.1 (0.68) and control 50.9 (0.44), $t_{100} = 2.12$, $p < .05$; loss: depressed 52.6 (0.66) and control 52.2 (1.10), $t_{100} = 0.35$, $p > .1$; average [SE] number of blocks for depressed and control groups were reward: depressed 3.84 (0.14) and control 4.28 (0.21), $t_{100} = -1.80$, $p = .08$; loss: depressed 3.94 (0.13) and control 4.16 (0.19), $t_{100} = -0.93$, $p > .1$). Participants completed a practice round prior to entering the scanner. Participants were given an initial endowment of $10. To ensure participants were attending to the task and had suitable behavior for model fitting, participants who switched

options in either reward or loss blocks less than 5% of the time were excluded from analyses (similar to Klein-Flügge et al., 2015; Sokol-Hessner et al., 2009).

Neuroimaging data collection and preprocessing

Participants were scanned on a 3T Siemens Tim Trio MR scanner. Echoplanar images were collected in 34 4-mm slices at a 30° hyperangulation from the anterior-posterior commisure (AC-PC) line (TR = 2000 ms, TE = 30 ms, flip angle = 90°, matrix = 64 x 64, voxel size = 3.4 x 3.4 x 4.0 mm^3). A high resolution (1 mm^3) anatomical Magnetization Prepared Rapid Gradient Echo (MPRAGE) T1 image (TR = 1200 ms, TE = 2.66 ms, flip angle = 12°) was collected to aid in registration.

Preprocessing and all further imaging analyses were conducted using SPM8 for fMRI (Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) and consisted of slice timing correction, realignment to the first functional image, coregistration to the participant's high-resolution structural image, normalization to the MNI template, and smoothing to ensure Gaussianity (6mm FWHM). Participants with motion greater than 3 mm or 0.05 radians in any direction or who had incomplete scanning data were excluded.

Cognitive-behavioral therapy

After completing baseline study procedures, participants with depression were offered 12 weeks of cognitive-behavioral therapy. The naturalistic design of our study meant that participants with depression were free to enroll in the treatment phase of the study or to decline treatment. Patients receiving treatment were treated by doctoral-level clinical

psychologists using the manual by Munoz & Miranda (Munoz & Miranda, 1996). After completing therapy, participants completed all study procedures again; control participants also participated in follow-up analyses at similar time points to patients in treatment (mean [SE] number of days between first and second time point: patients 115 [2.8], controls 111 [2.3], $t_{57} = 0.930$, $p > .1$). Thirty-seven treatment completers had clinical data at both time points and were used for analyses of baseline learning measures predicting symptoms at follow-up; nine of these completers lacked post-treatment task data, resulting in 28 treatment completers (and 20 controls) with suitable data for analyses of pre- to post-treatment behavioral and neural changes (see Figure 1.11 for a TREND diagram of participant flow through treatment). Similar to baseline analyses, treatment completers did not differ from patients who did not complete treatment or who were excluded from analyses on any clinical or demographic measures except estimated IQ.

Data analyses

Model-free analysis of behavior at baseline

Model-free analyses assessed the proportion of correct choices and proportion of switches per condition and participant and compared between groups using t-tests and by symptom severity using correlations.

Reinforcement learning model specification and estimation

Model-based analyses were conducted using reinforcement learning models (Rescorla & Wagner, 1972; Sutton & Barto, 1998). The primary reinforcement learning model included three free parameters per condition, and was tested against alternative models with

different effects on values and choices as well as against models combining some or all parameters across reward and loss conditions (see *Model validation* for more information). The primary model included the parameters of learning rate α, which indexed the speed of updating based on prediction error, outcome sensitivity ρ, which multiplicatively scaled more extreme outcome values, resulting in differential scaling of values, and outcome shift τ, which linearly shifted all outcome values, resulting in an overall positive or negative valuation bias. Outcome shift was added to all outcome values *R*, while outcome sensitivity was multiplied on the outcome if it was the outcome farther from 0 (+/-$0.70 to $0.80) to create a modified outcome value *R*':

$$R_t' = \rho * R_t + \tau \quad if\ |R_t| > .5$$
$$else \quad R_t' = R_t + \tau$$

To update the expected value *Q* for the next trial, the expected value for the current trial was subtracted from this modified outcome value to create the prediction error, which was scaled by the learning rate and added to the expected value for the current trial:

$$Q_{t+1} = Q_t + \alpha * (R_t' - Q_t)$$

Expected values were initialized at 0 at the beginning of each block and were updated separately for each stimulus. Expected values were transformed into choice probabilities *P* using a softmax function (here, the probability of choice A relative to choice B is shown):

$$P(A)_t = \frac{e^{\beta * Q(A)_t}}{(e^{\beta * Q(A)_t} + e^{\beta * Q(B)_t})}$$

Inverse temperature β was set at a value estimated without outcome sensitivity due to collinearity (β ≈ 5); because of this collinearity, in models with outcome sensitivity, inverse temperature was fixed at this value rather than estimated as a free parameter. In models without outcome sensitivity, inverse temperature was estimated as a free parameter.

Participants' choices were fit to models using hierarchical Bayesian estimation, which estimated the distribution of each free parameter over the group of participants and for each participant individually (Daw, 2011; Wiecki et al., 2015). By allowing group and individual level distributions over parameters and allowing each level to inform estimates at other levels, hierarchical Bayesian estimation more accurately recovers true parameters, especially parameters that are somewhat correlated as is often the case in reinforcement learning models (Ahn et al., 2011; Gershman, 2016; Huys et al., 2013). Posterior distributions were estimated using Hamiltonian Monte Carlo as implemented in Stan via its RStan interface (Carpenter et al., 2016, version 2.11). Group level parameters were specified as normally distributed, with a lower bound of 0 on outcome sensitivity (or inverse temperature, in models with this parameter) to constrain the parameter to be positive. Parameters were given a non-centered parameterization to aid in estimation by specifying mean, scale, and error distributions for each parameter (Betancourt & Girolami, 2015). Similar to (Gillan et al., 2015; Otto et al., 2013), mean distributions, estimated at the group level, were specified as normally distributed with priors of mean = 0 and standard deviation = 10, for parameters that were not logit transformed, or with standard deviation = 2.5 for parameters that were logit transformed. Scale distributions, estimated at the group level, were given a half-Cauchy prior (Gelman et al., 2014; bounded to be greater than 0) with values of 0 and 2.5 (0 and 2 for parameters that were logit transformed). Error

distributions, which were estimated for each subject, were given a normal prior with mean = 0 and standard deviation = 1. Similarly, effects of covariates, and for treatment analyses, effects of time and the interaction of covariate and time, were given a normal prior with mean = 0 and standard deviation = 1. For learning rate, these parameter values were then run through a logistic transformation to bound values between 0 and 1. Therefore, each subject's parameter (for example, learning rate) consisted of a group estimated mean value plus the combined value of the group estimated scale value multiplied by the individually estimated error value. The effects of covariates and of time were assumed to adjust the mean of each parameter (due to homogeneity of variance), and so acted on the mean value of the parameter per subject. Models were estimated separately for reward and loss conditions. Four chains were run for each condition, with 4000 samples per chain (2000 after discarding warm-up samples). Chains were visually inspected for convergence and showed good mixing, with all values of the potential scale reduction factor (Gelman & Rubin, 1992) less than 1.1.

Model validation

To test if the primary RL model explained participants' behavior well, alternate plausible models that tested a basic RL model, a model adding only outcome sensitivity and not outcome shift (per Huys et al., 2013), or a model adding perseverative effects on choices independent of value (per Daw, 2011) were tested and compared against the primary model. These alternative models consisted of (1) a basic reinforcement learning model without outcome shift or outcome sensitivity, and with inverse temperature as a free parameter (model $\alpha + \beta$; 2 free parameters per reward and loss condition); (2) a model

with outcome sensitivity and learning rate and no outcome shift (model α + ρ; 2 free parameters per condition); (3) a model accounting for value-independent choice effects that included learning rate, inverse temperature, and a perseveration term ω:

$$P(A)_t = \frac{e^{\beta * Q(A)_t + choice_{t-1} * \omega}}{(e^{\beta * Q(A)_t + choice_{t-1} * \omega} + e^{\beta * Q(B)_t + |(1 - choice_{t-1})| * \omega})}$$

where $choice_{t-1}$ is 1 if the stimulus (A in this example) was chosen on the previous trial and 0 if it was not (model α + β + ω; 3 free parameters per condition).

We also assumed a priori that learning patterns differed between reward and loss conditions. To test whether participants' behavior was better accounted by combining parameters across reward and loss conditions, models with (1) all parameters combined across reward and loss (total number of parameters = 3), (2) each parameter split between conditions in turn (e.g., learning rate split between conditions while outcome sensitivity and outcome shift were combined, repeated in turn for outcome sensitivity and outcome shift; total number of parameters per model = 4), and (3) all parameters split between conditions (total number of parameters = 6). The integrated BIC (iBIC) was computed from the likelihood over the posterior distribution, penalizing for the number of parameters (Guitart-Masip et al., 2012). To check if the best fitting model was the same across groups, iBIC was calculated across all participants as well as within control and depressed groups separately.

To ensure model parameters accurately reflected participants' behavior and were independently estimable, several model validation steps were conducted. First, individually estimated parameters were compared against model-free summary statistics of

performance and switching to determine how parameters related to model-free behavior. Second, simulated data for 100 subjects was created with three different levels of the mean value of each parameter (.25, .50, and .75 for reward and loss learning rate; .50, 1.0, and 1.5 for reward and loss outcome sensitivity; -.50, -.25, and 0 for reward outcome shift; and 0, .50, and 1.0 for loss outcome shift); values were chosen based on the range of parameters in real participants' behavior. Model parameters were estimated for this simulated data and recovered parameter values were plotted against simulated parameter values to verify parameters could be separably estimated at different values. Lastly, to ensure parameters were separably identifiable, the samples from the posterior distributions for each parameter were plotted against the other parameters in each condition to allow visual inspection of any correlations or trade-offs in the value of each parameter across its posterior distribution.

Assessment of model based behavioral differences at baseline

Estimation of parameters using both group and individual level information introduces dependencies among the individual level estimates, such that using participants' parameters on an individual basis to compare against outside measures (e.g., symptom severity, diagnosis) can be biased (Efron & Morris, 1977; Gelman et al., 2014). Therefore, to examine relationships between parameters and variables of interest, the effects were estimated within the model by introducing another parameter to index the effect of the variable of interest. To do so, covariates were z-scored or, in the case of binary variables such as diagnosis, dummy coded, and entered into a regression to predict the mean of a

parameter. For example, the following analysis determines the effect of anhedonia on learning rate:

$$\alpha_{total} = \alpha_{intercept} + anhedonia*\alpha_{anhedonia}$$

In this manner, $\alpha_{anhedonia}$ represents the effect of (standardized) anhedonia severity on learning rate. To determine significance, the 95% credible intervals of these parameters were required to not include 0 (i.e., to be entirely above or below 0). To account for potential nonlinearities in the relationship between symptoms and behavior across clinical and non-clinical levels of depression, models were also run within the depressed participants only as well as across all participants. Note that as a scaling parameter, learning rate lies only in the range of 0 to 1, but to aid in estimation, this parameter was estimated as a continuous variable and then logistically transformed to be bounded between 0 and 1. Therefore, the effects of symptoms on learning rate were reported both for untransformed (continuous) and transformed (range of 0 to 1) values for ease of interpretation (figures show transformed distributions for ease of interpretation). Analyses were also run with estimated IQ and presence of psychotropic medication as additional covariates, but inclusion of these covariates did not meaningfully change any results.

Baseline imaging analyses

First level imaging analyses used parametric regressors of prediction error δ or outcome value R_t at the time of outcome and expected value Q_t of the chosen option at the time of onset. Prediction error and expected value were calculated based on participants' individually estimated parameters from the reinforcement learning model and were z-

transformed prior to entering in the imaging model (Lebreton & Palminteri, 2016). Regressors were separated by condition (reward or loss) and all regressors were modeled as stick functions. Additional regressors of no interest were included for button presses, block number, and six motion parameters. Data were high pass filtered with a cutoff of 128 seconds.

Primary group level analyses focused on ventral striatum and ventromedial prefrontal cortex, brain areas known to be central to reinforcement learning (O'Doherty et al., 2007; Rangel et al., 2008). Regions of interest were defined from a recent meta-analysis of prediction error and expected value BOLD response in reinforcement learning tasks (Chase et al., 2015); specifically, a 6 mm sphere was drawn around the peak coordinate from right and left striatum from prediction error-related activation and subgenual anterior cingulate cortex from expected value-related activation. The first eigenvariate of the beta values in each ROI from prediction error, outcome value, and expected value activations was extracted for each participant and regressed against measures of interest. Additional whole brain analyses were run to examine contributions of expected value and outcome value signals at the outcome time point.

To relate neural activation to symptom measures, BOLD activity (ROI values and whole-brain activation) were correlated with symptom measures. The relationship between expected value and prediction error related activity and symptom measures was further tested by testing the interaction of prediction error neural signal and symptom measures on expected value neural signal in striatal ROIs (as in Greenberg et al., 2015). Analyses

were also run with estimated IQ and presence of psychotropic medication as additional covariates; inclusion of these covariates did not meaningfully change any results.

Treatment-related behavioral analyses

To investigate the relationship between changes in reinforcement learning parameter values and changes in symptoms with the depressed participants, changes in symptoms were defined as the percent change from pre- to post-treatment, such that higher values indicated more improvement. Similar to baseline analyses, the mean of each parameter was estimated as a regression of the intercept of the parameter, the within-subject effect of time (dummy coded for 0 = first session and 1 = second session to represent the change in the mean value from pre- to post-treatment independent of changes in symptoms), the effect of changes in symptoms, and the interaction of time and changes of symptoms (Gelman & Hill, 2006):

$$\alpha_{total} = \alpha_{intercept} + time*\alpha_{time} + \Delta anhedonia*\alpha_{anhedonia} + time*\Delta anhedonia*\alpha_{interaction}$$

In this analysis, the $\alpha_{interaction}$ parameter assesses the change in learning rate from the first to second session that correlates with percent change in anhedonia. Analyses were also run with estimated IQ and presence of psychotropic medication as additional time-independent covariates; inclusion of these covariates did not meaningfully change any results.

To assess practice effects on the task, a similar analysis was carried out in control participants, but omitting the effects of changes with symptoms. Therefore, this analysis examined overall changes in parameters from the first to second visit in participants whose behavior was not affected by treatment or natural fluctuations in symptoms.

Treatment imaging analyses

To assess changes with treatment that correlate with percent change in symptoms, changes in neural activity within ROIs of striatum and subgenual anterior cingulate cortex were correlated with percent change in symptoms. Additionally, to test neural predictors of treatment, pre-treatment activity within ROIs was correlated with percent change in symptoms.

Statistical analysis

Linear regressions were run on behavioral data (within model estimation, see above for details) and first and second level imaging data. For Bayesian analyses, 95% credible intervals were used to determine significance, and for frequentist analyses, alpha < .05. Whole brain imaging analyses used an initial cluster defining threshold of $p<.001$ and a cluster-level FDR significance of $p<.05$ (Woo et al., 2014).

References

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 95–110. doi:10.1037/a0020684.A

Anticevic, A., Murray, J. D., & Barch, D. M. (2015). Bridging levels of understanding in schizophrenia through computational modeling. *Clinical Psychological Science*. doi:10.1177/2167702614562041

Beck, A. T. (2005). The current state of cognitive therapy. *Archives of General Psychiatry*, *62*(9), 953–959. doi:10.1001/archpsyc.62.9.953

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *The American Journal of Psychiatry*, *165*(8), 969–77. doi:10.1176/appi.ajp.2008.08050721

Beevers, C. G., Worthy, D. A., Gorlick, M. a, Nix, B., Chotibut, T., & Maddox, W. T. (2013). Influence of depression symptoms on history-independent reward and punishment processing. *Psychiatry Research*, *207*(1–2), 53–60. doi:10.1016/j.psychres.2012.09.054.Influence

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Frontiers in Human Neuroscience*, *7*(April), 133. doi:10.3389/fnhum.2013.00133

Betancourt, M., & Girolami, M. (2015). Hamiltonian Monte Carlo for Hierarchical Models. In *Current Trends in Bayesian Methodology with Applications* (pp. 79–101).

Bogdan, R., & Pizzagalli, D. A. (2006). Acute stress reduces reward responsiveness: Implications for depression. *Biological Psychiatry*, *60*(10), 1147–1154. doi:10.1016/j.biopsych.2006.03.037

Buckby, J. A., Yung, A. R., Cosgrave, E. M., & Killackey, E. J. (2007). Clinical utility of the Mood and Anxiety Symptom Questionnaire (MASQ) in a sample of young help-seekers. *BMC Psychiatry*, *7*(1), 50. doi:10.1186/1471-244X-7-50

Burkhouse, K. L., Kujawa, A., Kennedy, A. E., Shankman, S. A., Langenecker, S. A., Phan, K. L., & Klumpp, H. (2016). Neural reactivity to reward as a predictor of cognitive behavioral therapy response in anxiety and depression. *Depression and Anxiety*, *33*(4), 281–288. doi:10.1002/da.22482

Carl, H., Walsh, E., Eisenlohr-Moul, T., Minkel, J., Crowther, A., Moore, T., … Smoski, M. J. (2016). Sustained anterior cingulate cortex activation during reward processing predicts response to psychotherapy in major depressive disorder. *Journal of Affective Disorders*, *203*, 204–212. doi:10.1016/j.jad.2016.06.005

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2016). Journal of Statistical Software Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, *VV*(Ii).

Cavanagh, J. F., Bismark, A. J., Frank, M. J., & Allen, J. J. B. (2011). Larger error signals in major depression are associated with better avoidance learning. *Frontiers in Psychology*, *2*(November), 331. doi:10.3389/fpsyg.2011.00331

Chase, H. W., Frank, M. J., Michael, A., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2010). Approach and avoidance learning in patients with major depression and

healthy controls: Relation to anhedonia. *Psychological Medicine*, *40*(3), 433–40. doi:10.1017/S0033291709990468

Chase, H. W., Kumar, P., Eickhoff, S. B., & Dombrovski, A. Y. (2015). Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*. doi:10.3758/s13415-015-0338-7

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience & Biobehavioral Reviews*, *55*, 247–267. doi:10.1016/j.neubiorev.2015.05.005

Chiu, P. H., & Deldin, P. J. (2007). Neural evidence for enhanced error detection in major depressive disorder. *The American Journal of Psychiatry*, *164*(4), 608–16. doi:10.1176/appi.ajp.164.4.608

Chiu, P. H., Lohrenz, T. M., & Montague, P. R. (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nature Neuroscience*, *11*(4), 514–20. doi:10.1038/nn2067

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*(3), 316–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1918611

Cools, R., Robinson, O. J., & Sahakian, B. J. (2008). Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology*, *33*(9), 2291–9. doi:10.1038/sj.npp.1301598

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning:*

Attention and performance XXIII (pp. 3–38).

Dichter, G. S., Felder, J. N., Petty, C. M., Bizzell, J., Ernst, M., & Smoski, M. J. (2009). The effects of psychotherapy on neural responses to rewards in major depression. *Biological Psychiatry*, *66*(9), 886–97. doi:10.1016/j.biopsych.2009.06.021

Disner, S. G., Beevers, C. G., Haigh, E. A. P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467–77. doi:10.1038/nrn3027

Dobson, K. S., Hollon, S. D., Dimidjian, S., Schmaling, K. B., Kohlenberg, R. J., Gallop, R. J., … Jacobson, N. S. (2008). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *Journal of Consulting and Clinical Psychology*, *76*(3), 468–77. doi:10.1037/0022-006X.76.3.468

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment*, *10*(2), 83–89. doi:10.1037/1040-3590.10.2.83

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. *Scientific American*, *236*, 119–127.

Elliott, R., Sahakian, B. J., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1997). Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *Journal of Neurology, Neurosurgery, and Psychiatry*, *63*(1), 74–82. doi:10.1136/jnnp.63.1.74

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression.

Biological Psychiatry, *68*(2), 118–24. doi:10.1016/j.biopsych.2010.01.027

Faulkner, P., & Deakin, J. F. W. (2014). The role of serotonin in reward, punishment and behavioural inhibition in humans: Insights from studies with acute tryptophan depletion. *Neuroscience & Biobehavioral Reviews*, *46*, 365–378. doi:10.1016/j.neubiorev.2014.07.024

First, M., Spitzer, R., Gibbon, M., & Wiliams, J. (1996). *User's guide for the structured interview for DSM-IV axis I disorders—research version (SCID-I)*. New York: Biometrics Research.

Francis, T. C., & Lobo, M. K. (2016). Emerging role for nucleus accumbens medium spiny neuron subtypes in depression. *Biological Psychiatry*, 645–653. doi:10.1016/j.biopsych.2016.09.007

Fu, C. H. Y., Williams, S. C. R., Cleare, A. J., Brammer, M. J., Walsh, N. D., Kim, J., … Mitterschiffthaler, M. T. (2004). Attenuation of the neural response to sad faces in Major Depression by antidepressant treatment. *Archives of General Psychiatry*, *61*, 877–889. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Attenuation+of+the+Neural+Response+to+Sad+Faces+in+Major+Depression+by+Antidepressant+Treatment#1

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis* (Third.). Boca Raton, FL: CRC Press.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–511. doi:10.1214/ss/1177011136

Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, *71*, 1–6. doi:10.1016/j.jmp.2016.01.006

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective, & Behavioral Neuroscience*, (March), 523–536. doi:10.3758/s13415-015-0347-6

Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex*, *19*(2), 483–495. doi:10.1093/cercor/bhn098

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., … Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, *134*, 1751–64. doi:10.1093/brain/awr059

Greenberg, T., Chase, H. W., Almeida, J. R., Stiffler, R., Zevallos, C. R., Aslam, H. A., … Phillips, M. L. (2015). Moderation of the relationship between reward expectancy and prediction error-related ventral striatal reactivity by anhedonia in unmedicated major depressive disorder: Findings from the EMBARC study. *American Journal of Psychiatry*, *172*(9), 881–891. doi:10.1176/appi.ajp.2015.14050594

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, *62*(1), 154–166. doi:10.1016/j.neuroimage.2012.04.024

Hanson, J. L., Albert, D., Iselin, A.-M., Carre, J. M., Dodge, K. A., & Hariri, A. R. (2015).

Cumulative stress in childhood is associated with blunted reward-related brain activity in adulthood. *Social, Cognitive and Affective Neuroscience*, (October), 405–412. doi:10.1093/scan/nsv124

Harlé, K. M., Guo, D., Zhang, S., Paulus, M. P., & Yu, A. J. (2017). Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. *Plos One*, *12*(10), e0186473. doi:10.1371/journal.pone.0186473

Heller, A. S., Johnstone, T., Light, S. N., Peterson, M. J., Kolden, G. G., Kalin, N. H., & Davidson, R. J. (2013). Relationships between changes in sustained fronto-striatal connectivity and positive affect in major depression resulting from antidepressant treatment. *The American Journal of Psychiatry*, *170*(2), 197–206. doi:10.1176/appi.ajp.2012.12010014

Herzallah, M. M., Moustafa, A. A., Natsheh, J. Y., Abdellatif, S. M., Taha, M. B., Tayem, Y. I., … Gluck, M. A. (2013). Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. *Frontiers in Integrative Neuroscience*, *7*(September), 67. doi:10.3389/fnint.2013.00067

Huys, Q. J. M., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood and Anxiety Disorders*, *20*, 1–29. Retrieved from http://www.biolmoodanxietydisord.com/content/3/1/12/abstract

Keshavan, M. S., Vinogradov, S., Rumsey, J., Sherrill, J., & Wagner, A. (2014). Cognitive training in mental disorders: update and future directions. *The American Journal of Psychiatry*, *171*(May), 510–522. doi:10.1176/appi.ajp.2013.13081075

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(June). Retrieved from http://archpsyc.ama-assn.org/cgi/reprint/62/6/617.pdf

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T. M., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science*, *321*, 1–6. Retrieved from http://www.sciencemag.org/content/321/5890/806.short

Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLOS Computational Biology*, *11*(3), e1004116. doi:10.1371/journal.pcbi.1004116

Kumar, P., Berghorst, L. H., Nickerson, L. D., Dutra, S. J., Goer, F. K., Greve, D. N., & Pizzagalli, D. A. (2014). Differential effects of acute stress on anticipatory and consummatory phases of reward processing. *Neuroscience*, *266*, 1–12. doi:10.1016/j.neuroscience.2014.01.058

Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., & Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain*, *131*(Pt 8), 2084–93. doi:10.1093/brain/awn136

Kunisato, Y., Okamoto, Y., Ueda, K., Onoda, K., Okada, G., Yoshimura, S., … Yamawaki, S. (2012). Effects of depression on reward-based decision making and variability of action in probabilistic learning. *Journal of Behavior Therapy and Experimental*

Psychiatry, *43*(4), 1088–94. doi:10.1016/j.jbtep.2012.05.007

Lebreton, M., & Palminteri, S. (2016). Assessing inter-individual variability in brain-behavior relationship with functional neuroimaging. *bioRxiv*, 1–14.

Lewinsohn, P. M. (1974). A behavioral approach to depression. In *Essential papers on depression* (pp. 150–172). New York: NYU Press.

Luking, K. R., Pagliaccio, D., Luby, J. L., & Barch, D. M. (2015). Child gain approach and loss avoidance behavior: Relationships with depression risk, negative mood, and anhedonia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(8), 643–651. doi:10.1016/j.jaac.2015.05.010

Maddox, W. T., Gorlick, M. A., Worthy, D. A., & Beevers, C. G. (2012). Depressive symptoms enhance loss-minimization, but attenuate gain-maximization in history-dependent decision-making. *Cognition*, *125*(1), 118–24. doi:10.1016/j.cognition.2012.06.011

Mayberg, H. S. (2003). Modulating dysfunctional limbic-cortical circuits in depression: Towards development of brain-based algorithms for diagnosis and optimised treatment. *British Medical Bulletin*, *65*, 193–207. doi:10.1093/bmb/65.1.193

McCabe, C., Mishor, Z., Cowen, P. J., & Harmer, C. J. (2010). Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. *Biological Psychiatry*, *67*(5), 439–45. doi:10.1016/j.biopsych.2009.11.001

McGrath, C. L., Kelley, M. E., Holtzheimer, P. E., Dunlop, B. W., Craighead, W. E., Franco, A. R., … Mayberg, H. S. (2013). Toward a neuroimaging treatment selection biomarker for major depressive disorder. *JAMA Psychiatry*, *70*(8), 821–9. doi:10.1001/jamapsychiatry.2013.143

McTeague, L., Huemer, J., Carreon, D., Jiang, Y., Eickhoff, S. B., & Etkin, A. (2017). Identification of a common neural circuit disruption in executive function across psychiatric disorders. *American Journal of Psychiatry*, *39*(9), S522–S522. doi:10.1176/appi.ajp.2017.16040400

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80. doi:10.1016/j.tics.2011.11.018.Computational

Munoz, R. F., & Miranda, J. (1996). *Individual therapy manual for cognitive-behavioral treatment for depression*. Santa Monica, CA: RAND.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. doi:10.1016/j.jmp.2008.12.005

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, *1104*, 35–53. doi:10.1196/annals.1390.022

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 20941–20946. doi:10.1073/pnas.1312011110

Paulus, M. P., Huys, Q. J. M., & Maia, T. V. (2016). A Roadmap for the Development of Applied Computational Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. doi:10.1016/j.bpsc.2016.05.001

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-

dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–5. doi:10.1038/nature05051

Pizzagalli, D. A. (2014). Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annual Review of Clinical Psychology*, *10*, 393–423. doi:10.1146/annurev-clinpsy-050212-185606

Pizzagalli, D. A., Iosifescu, D. V, Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87. doi:10.1016/j.jpsychires.2008.03.001

Rangel, A., Camerer, C. F., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–56. doi:10.1038/nrn2357

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: Current research and theory* (pp. 64–99).

Robinson, O. J., & Chase, H. W. (2017). Learning and choice in mood disorders: Searching for the computational parameters of anhedonia. *Computational Psychiatry*, 1–26.

Robinson, O. J., & Cools, R. (2012). Ventral striatum response during reward and punishment reversal learning in unmedicated major depressive disorder. *American Journal of Psychiatry*, *169*(2), 152–159. Retrieved from http://neuro.psychiatryonline.org/article.aspx?articleid=483659&RelatedWidgetArticles=true

Robinson, O. J., Overstreet, C., Allen, P. S., Letkiewicz, A., Vytal, K., Pine, D. S., & Grillon, C. (2013). The role of serotonin in the neurocircuitry of negative affective bias: Serotonergic modulation of the dorsal medial prefrontal-amygdala "aversive amplification" circuit. *NeuroImage*, *78*, 217–223. doi:10.1016/j.neuroimage.2013.03.075

Rothkirch, M., Tonn, J., Kohler, S., & Sterzer, P. (2017). Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Cerebral Cortex*, 1–11. doi:10.1093/cercor/bhw393

Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hrynkiewicz, L., … Dolan, R. J. (2017). Association of Neural and Emotional Impacts of Reward Prediction Errors With Major Depression. *JAMA Psychiatry*, 10–12. doi:10.1001/jamapsychiatry.2017.1713

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9054347

Sheline, Y. I., Price, J. L., Yan, Z., & Mintun, M. A. (2010). Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(24), 11020–5. doi:10.1073/pnas.1000446107

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences of the United States of America*,

106(13), 5035–40. doi:10.1073/pnas.0806761106

Steer, R. A., Ball, R., Ranieri, W. F., & Beck, A. T. (1999). Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. *Journal of Clinical Psychology*, *55*(1), 117–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10100838

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

Stephan, K. E., Bach, D. R., Fletcher, P. C., Flint, J., Frank, M. J., Friston, K. J., … Breakspear, M. J. (2015). Charting the landscape of priority problems in psychiatry, part 1: classification and diagnosis. *The Lancet Psychiatry*, *366*(15), 1–7. doi:10.1016/S2215-0366(15)00361-2

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT Press.

Taquet, M., Quoidbach, J., de Montjoye, Y.-A., Desseilles, M., & Gross, J. J. (2016). Hedonism and the choice of everyday activities. *Proceedings of the National Academy of Sciences*, 201519998. doi:10.1073/pnas.1519998113

Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., … Claes, S. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biological Psychiatry*, *73*(7), 639–45. doi:10.1016/j.biopsych.2012.10.014

Wang, X.-J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, *84*(3), 638–654. doi:10.1016/j.neuron.2014.10.018

Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, *104*(1), 3–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7897050

Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. Psychological Corporation.

Wiecki, T. V, Poland, J., & Frank, M. J. (2015). Model-based cognitive neuroscience approaches to computational psychiatry: Clustering and classification. *Clinical Psychological Science*, *3*(3), 378–399.

Woo, C. W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, *91*, 412–419. doi:10.1016/j.neuroimage.2013.12.058

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Geneva. doi:CC BY-NC-SA 3.0 IGO

Young, C. B., Chen, T., Nusslock, R., Keller, J., Schatzberg, A. F., & Menon, V. (2016). Anhedonia and general distress show dissociable ventromedial prefrontal cortex connectivity in major depressive disorder. *Translational Psychiatry*, *6*(5), e810. doi:10.1038/tp.2016.80

Figures

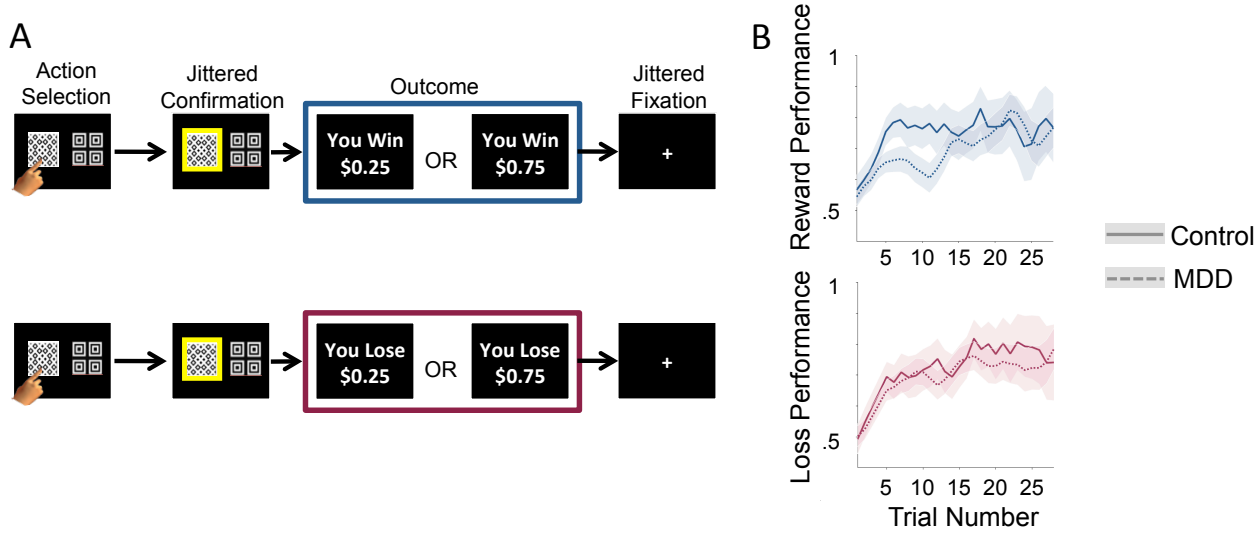

Figure 1.1. Task schematic and overall learning curves. **A)** Schematic description of reinforcement learning task. Participants choose between two stimuli, view the monetary outcome of the choice, and learn over time which is the 'better' option. The task involved blocks consisting of trials with all reward (top) and all loss (bottom) outcomes. **B)** Reward and loss learning performance. Performance was quantified as proportion of choices that were the 'better' option. Over time, participants show learning (running average over three trials; averaged over all blocks in condition; mean ± SE). Top panel (navy) comprises reward learning blocks while bottom panel (maroon) comprises loss learning blocks. Behavior is separated by diagnostic group, with control participants' behavior marked by a solid line and the behavior of participants with depression marked by a dotted line in each condition.

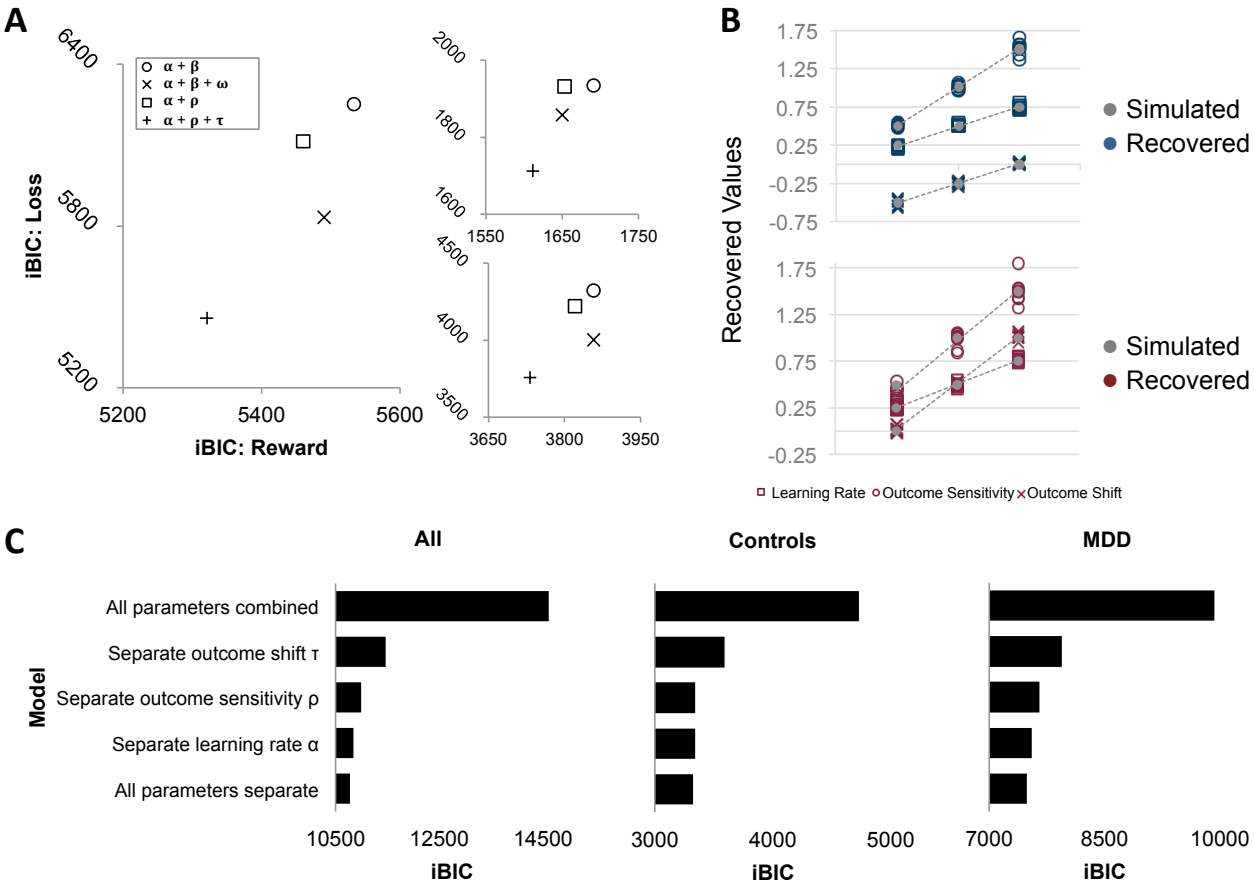

Figure 1.2. Model fit and parameter recovery. **A)** The model with learning rate α, outcome sensitivity ρ, and outcome shift τ fits better across all participants (large left panel), control participants (top right panel), and participants with depression (bottom right panel) across reward (x axis) and loss (y axis) conditions, relative to plausible alternative models of: learning rate α and inverse temperature β; learning rate α, inverse temperature β, and value-independent perseveration ω; and learning rate α and outcome sensitivity ρ. Values shown are integrated BIC (iBIC) values (Guitart-Masip et al., 2012). **B)** Parameter values can be independently recovered from simulated data; 100 participants with mean parameter values at three different levels, determined based on the range of real participants' values, were simulated and recovered. Top panel shows reward

parameters and bottom panel shows loss parameters, with simulated values indicated by gray dots connected by dotted gray lines, and recovered values indicated by navy (reward) or maroon (loss) symbols. Squares indicate recovered learning rate values, circles indicate recovered outcome sensitivity values, and crosses indicate recovered outcome shift values. **C)** Separating all parameters by condition (reward and loss) fits better across all participants, within control participants only, and within participants with depression only, relative to models combining one or all parameters across conditions. Values shown are integrated BIC (iBIC) values.

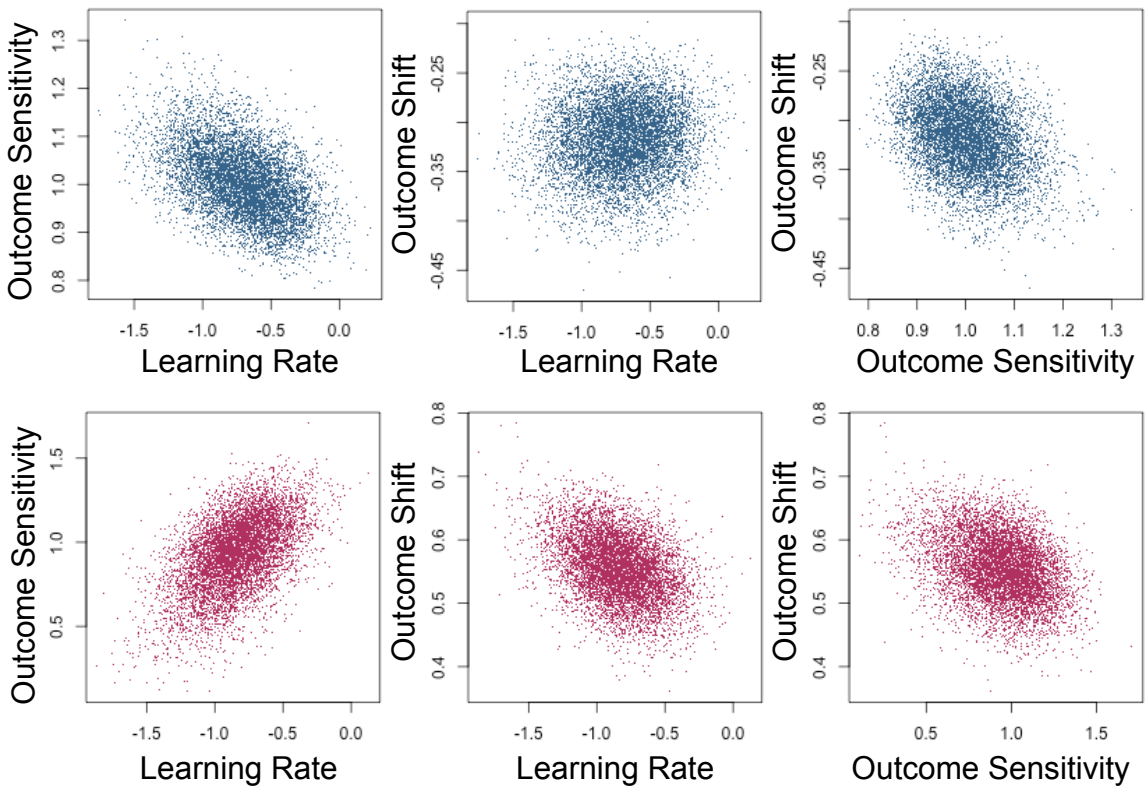

Figure 1.3. Posterior distributions of parameters. Top panels (navy) indicate relationships among posterior distributions of reward parameters and bottom panels (maroon) of loss parameters. Each dot represents a sample from the posterior distribution of all participants included in baseline analyses during MCMC sampling (after discarding warm-up samples).

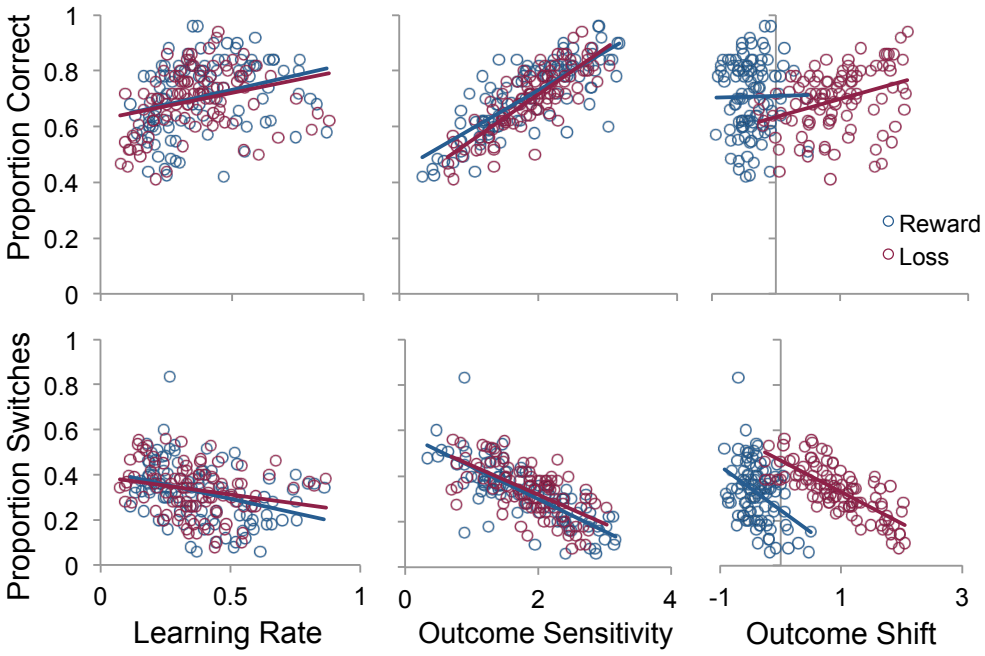

Figure 1.4. Relationship of model parameters with model-free summaries of behavior. Parameters of learning rate (left column), outcome sensitivity (middle column), and outcome shift (right column) show differentiable relationships with overall proportion of correct choices (top row) and overall proportion of switches (bottom row). Navy circles denote values from the reward learning condition and maroon circles denote values from the loss learning condition. Independent of condition, learning rate influences trial-by-trial changes in behavior but shows small relationships with model-free summaries. Higher outcome sensitivity leads to better discrimination of outcome values and better performance and fewer switches, while higher outcome shift does not lead to large changes in performance but reduces the tendency to switch.

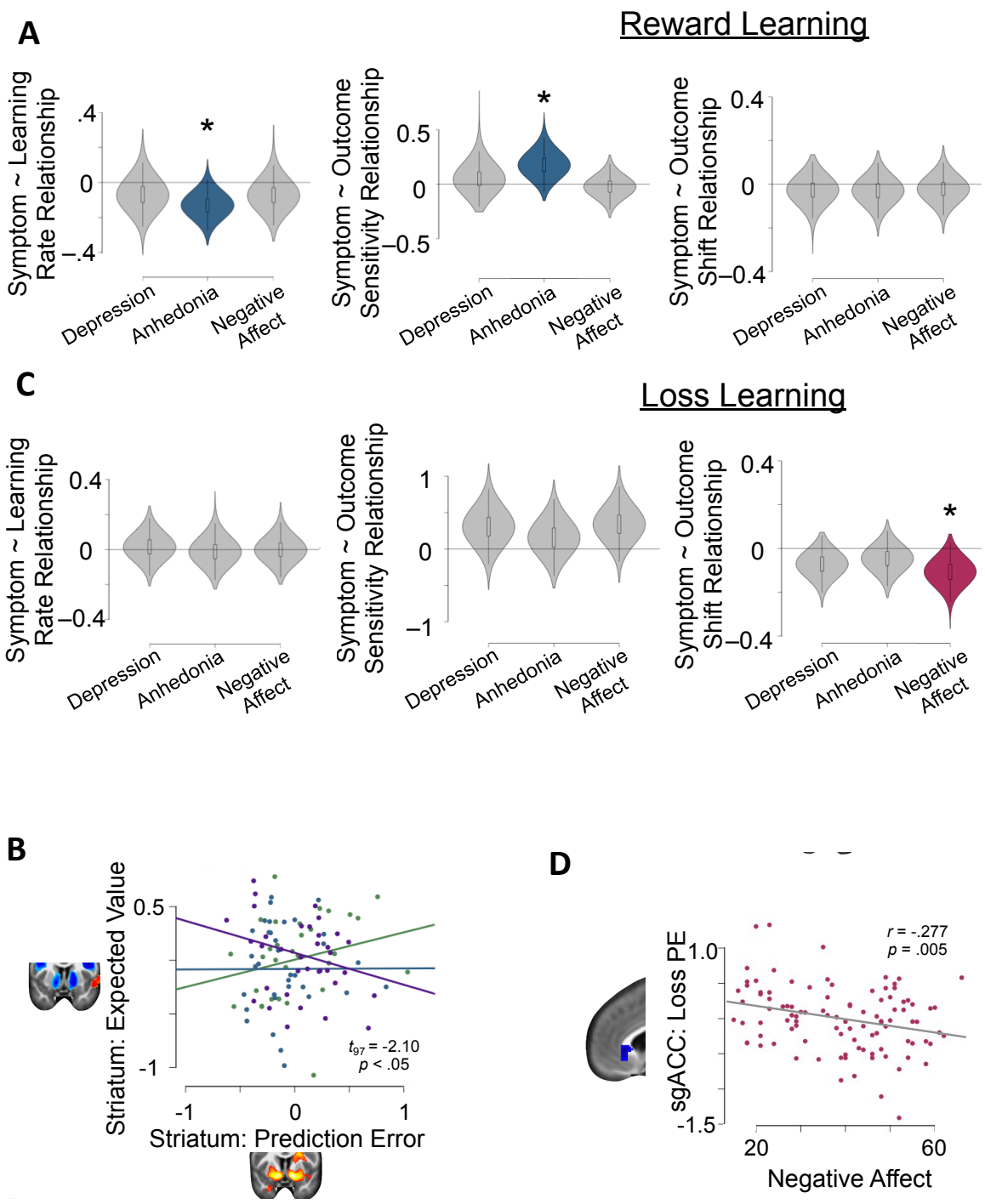

Figure 1.5. Relationship of behavioral and neural indicators of reinforcement learning and symptoms of depression. A) Behaviorally, in reward learning in

participants with depression, greater anhedonia is related to lower learning rate and higher outcome sensitivity. Violin plots indicate posterior distribution of relationship between behavioral learning parameter and symptom measure; navy violin plots with asterisks indicate significant relationships (95% credible interval does not include 0). **B)** Neurally, anhedonia moderates the relationship between striatal signaling to prediction error and expected value. Green dots indicate controls, blue dots indicate low anhedonic participants with depression (based on median split), and purple dots indicate highly anhedonic participants with depression for the right striatum region of interest activation to prediction error (x axis) and expected value (y axis). Respectively colored lines indicate regression lines for each group showing moderation of the prediction error – expected value relationship by anhedonia. **C)** Behaviorally, in loss learning across all participants, greater negative affect is related to more negative outcome shift. Violin plots indicate posterior distribution of relationship between behavioral learning parameter and symptom measure; maroon violin plots with asterisks indicate significant relationships (95% credible interval does not include 0). **D)** Neurally, negative affect is negatively related to subgenual anterior cingulate cortex (sgACC) signaling of prediction error at the time of outcome receipt. Maroon dots indicate individual participants' negative affect severity versus sgACC region of interest activation to prediction error; regression line (gray) indicates overall negative relationship.

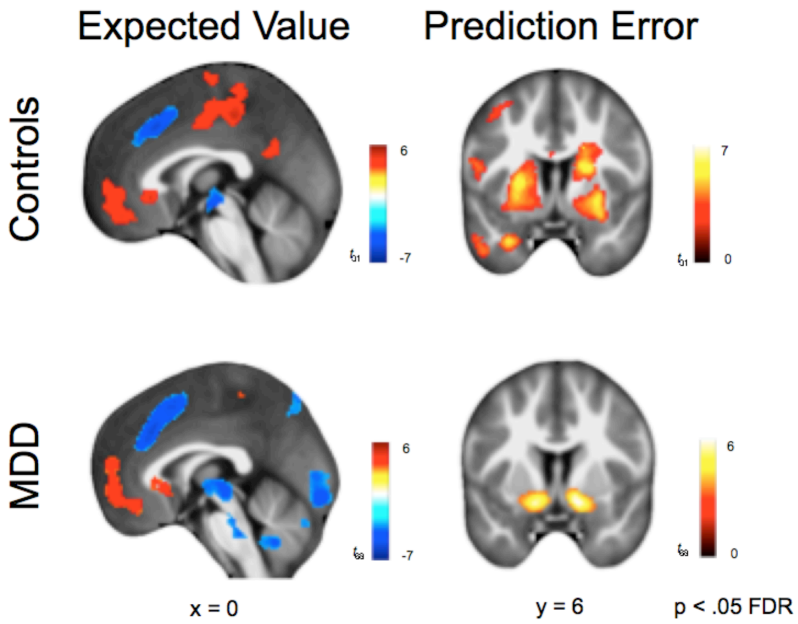

Figure 1.6. Lack of differences in expected value and prediction error reward signals in depression. Left column is modulation of brain activity by the parametric modulator of expected value at time of stimulus onset and right column is modulation of brain activity by the parametric modulator of prediction error at the time of outcome; top row is control participants and bottom row is participants with depression. Values are shown p<.05 whole brain FDR corrected (p<.001 cluster forming threshold).

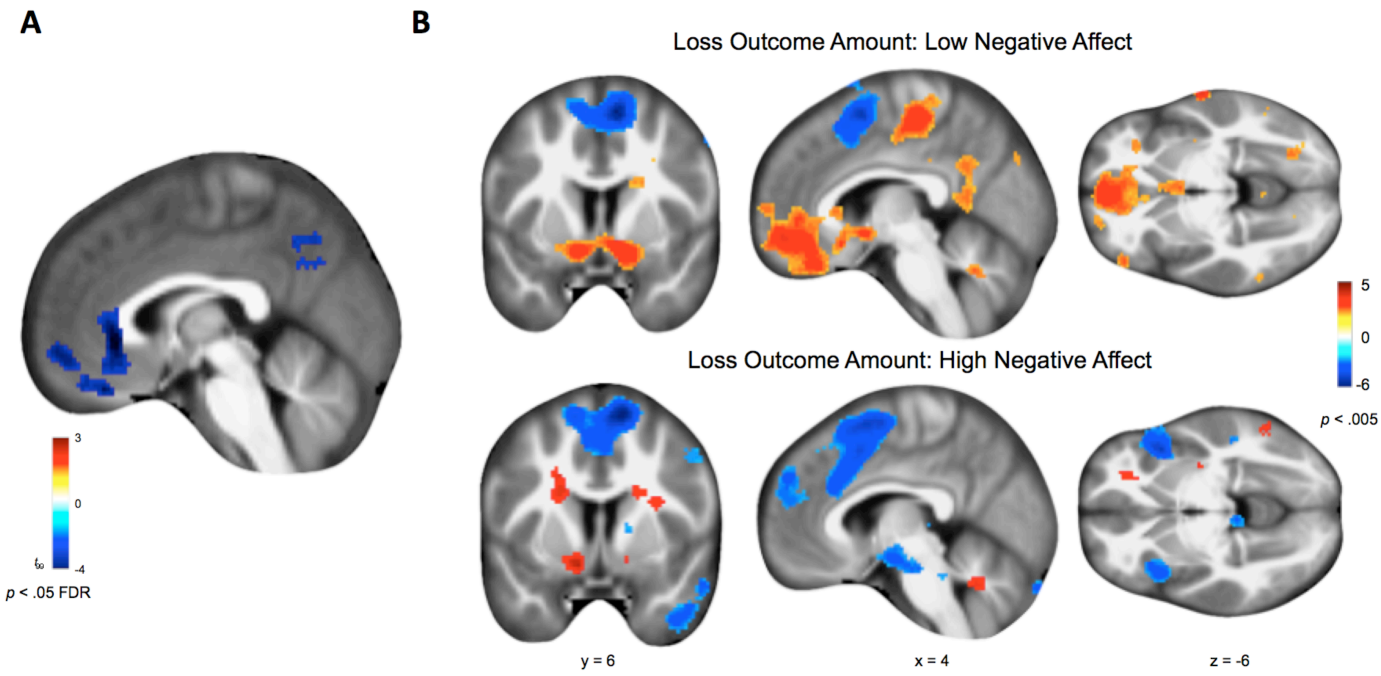

Figure 1.7. Differences in processing of loss outcomes by negative affect. A) Significant whole-brain corrected differences by negative affect in modulation of brain activity with level of outcome (FDR p<.05, cluster forming threshold of p<.001). Group-level covariate of negative affect on a parametric modulator of outcome value at the time of outcome receipt. **B)** Processing of outcome value separated into low (top) and high (bottom) negative affect participants. Note that both groups show robust responses that are modulated by outcome value (activation is significant p<.05 corrected & displayed at p<.005 uncorrected), but, reflecting the differences in signal by level of negative affect shown in A), these responses show a different spatial pattern and direction of activation in low and high negative affect participants.

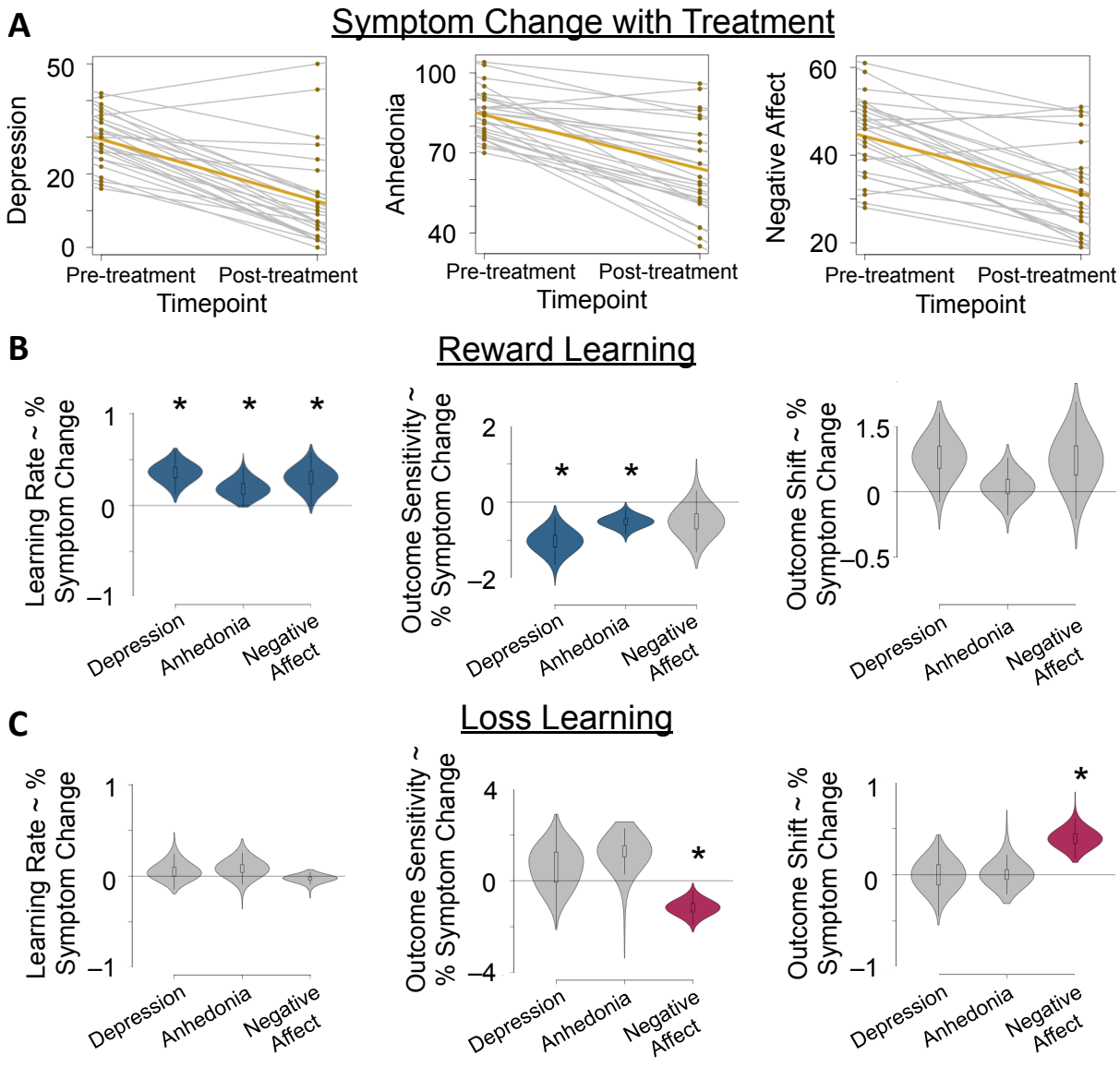

Figure 1.8. Changes in depression symptoms and behavioral reinforcement learning parameters with treatment. A) Changes in symptoms from pre-to post-treatment for individual patients (gray lines) and on average (gold line). Overall depression symptoms, anhedonia, and negative affect all decreased on average from pre- to post-treatment with cognitive behavioral therapy (CBT), but with large heterogeneity in treatment response across participants. **B)** Relationship between changes in reinforcement learning

parameters for reward learning and percent decrease in symptom severity with treatment. Increases in learning rate and decreases in outcome sensitivity correlated with improvements in symptoms, particularly anhedonia, during reward learning. Violin plots indicate posterior distribution of relationship between changes in behavioral learning parameter and percent change in symptom measure; navy violin plots with asterisks indicate significant relationships (95% credible interval does not include 0). **C)** Relationship between changes in reinforcement learning parameters for loss learning and percent decrease in symptom severity with treatment. Increases in outcome shift and decreases in outcome sensitivity correlated with improvements in negative affect during loss learning. Violin plots indicate posterior distribution of relationship between changes in behavioral learning parameter and percent change in symptom measure; maroon violin plots with asterisks indicate significant relationships (95% credible interval does not include 0).

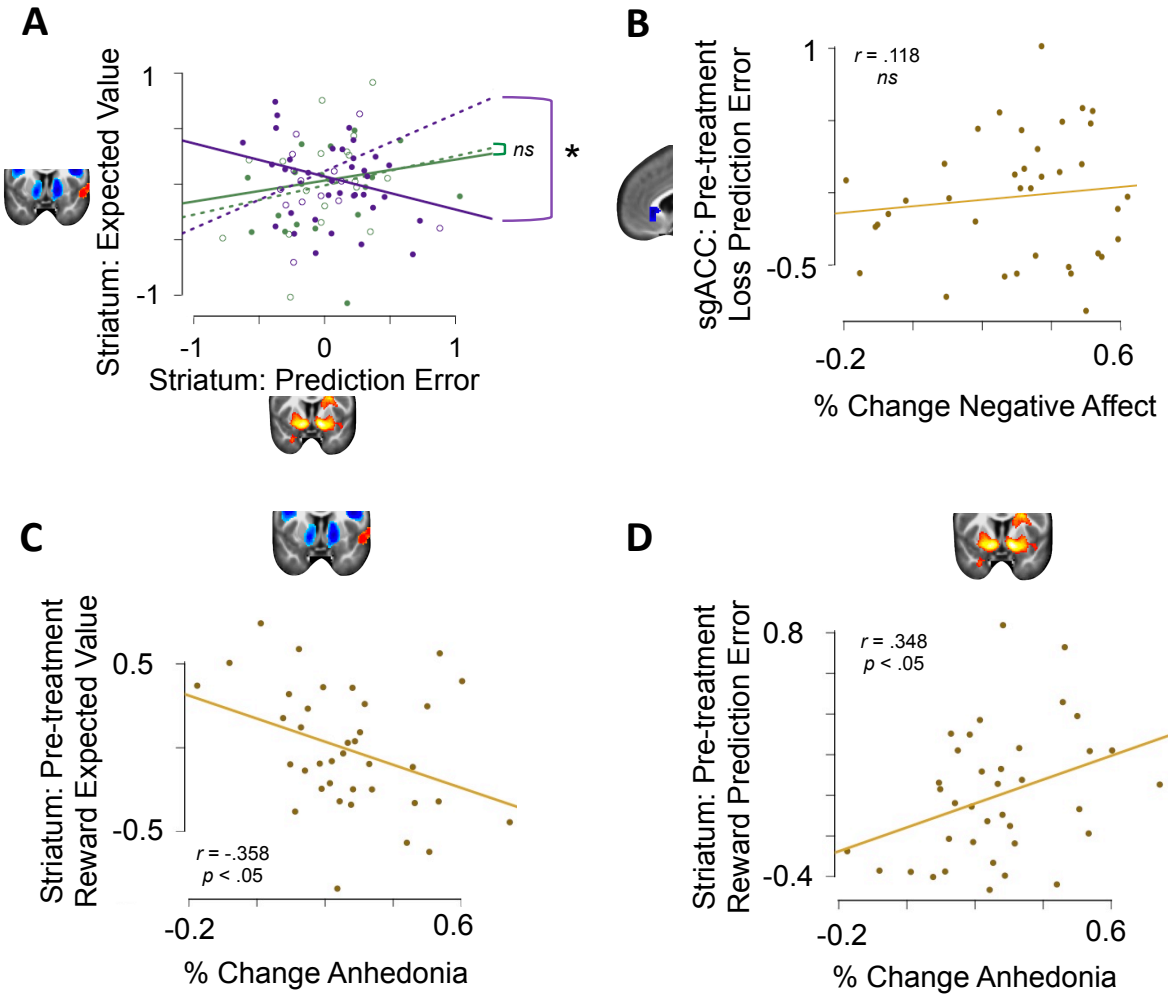

Figure 1.9. Neural predictors of symptom improvement and changes with treatment. **A)** Moderation of striatal expected value-prediction error relationship by pre-treatment anhedonia is significantly reduced post-treatment. Filled green dots indicate pre-treatment values for controls, green open circles indicate post-treatment values for controls, filled purple dots indicate pre-treatment values for highly anhedonic participants with depression, and open purple circles indicate post-treatment values for participants with high anhedonia pre-treatment. Values are right striatum region of interest activation to prediction error (x axis) and expected value (y axis). Solid lines represent regression lines

pre-treatment and dotted lines represent regression lines post-treatment, with a significant change in slope for the high anhedonia group and no change for controls. **B)** Pre-treatment sgACC signal does not correlate with changes in negative affect. Gold dots indicate patients' percent change in negative affect versus pre-treatment sgACC region of interest activation to prediction error. Line is regression line showing lack of relationship. **C)** Pre-treatment striatal expected value signal correlates negatively with changes in anhedonia. Gold dots indicate patients' percent change in anhedonia versus pre-treatment right striatum region of interest activation to expected value. Line is regression line showing significant negative relationship. **D)** Pre-treatment striatal prediction error signal correlates positively with changes in anhedonia. Gold dots indicate patients' percent change in anhedonia versus pre-treatment right striatum region of interest activation to prediction error. Line is regression line showing significant positive relationship.

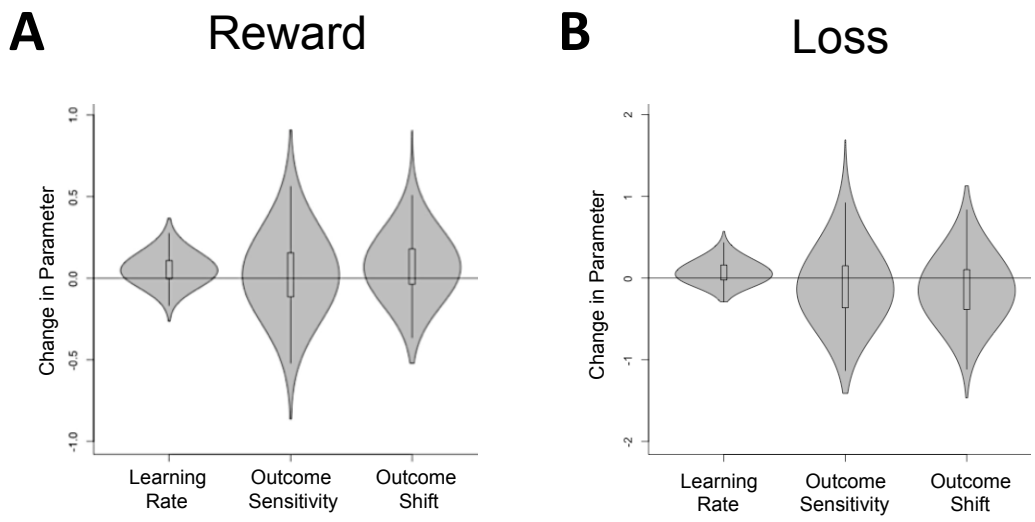

Figure 1.10. Stability of parameter estimates over time for control participants. Left panel is reward parameters and right panel is loss parameters; violin plots indicate posterior densities of the within-subject change in each parameter from the first timepoint to the second, with a value of 0 representing no change in learning parameters.

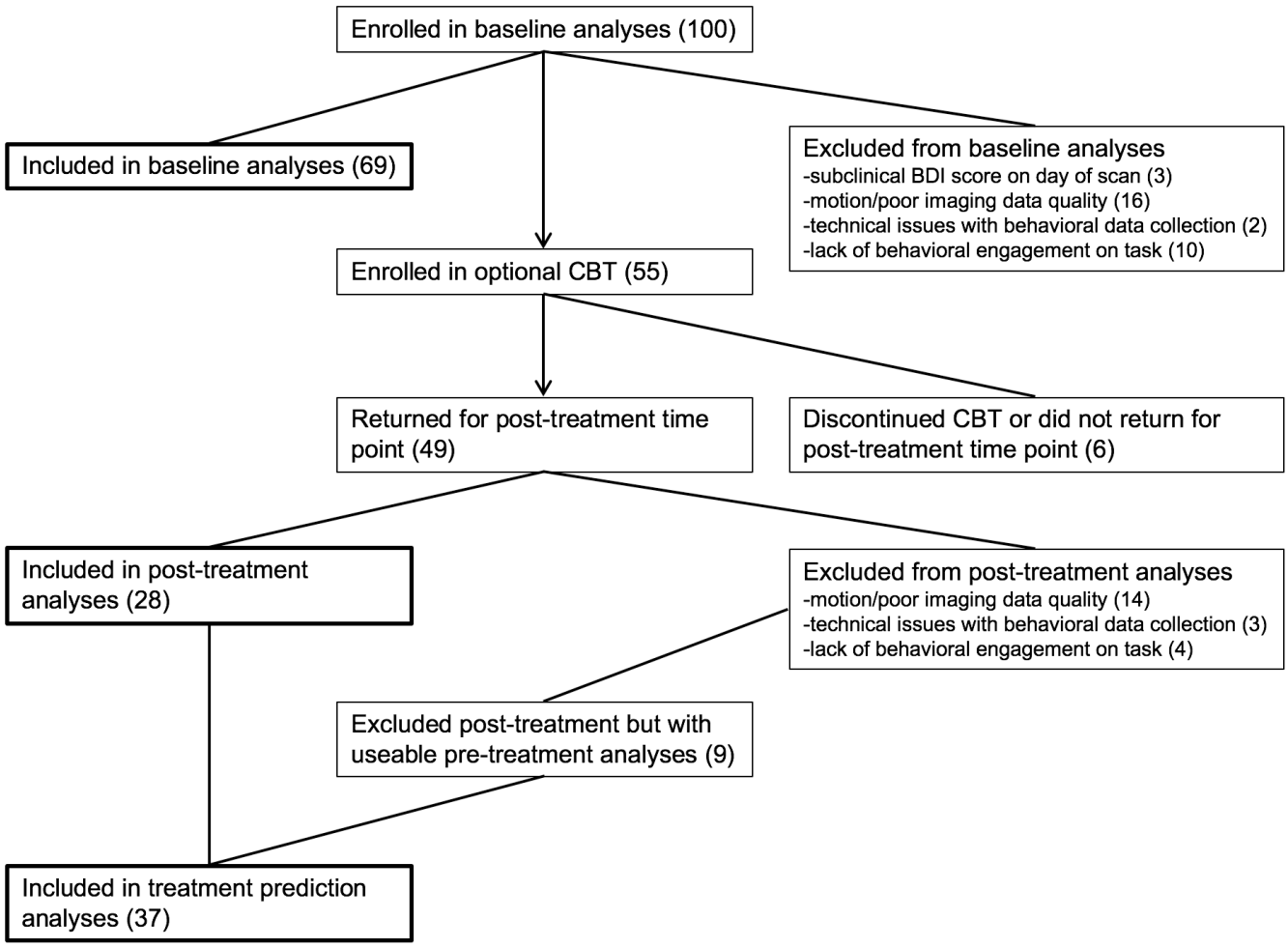

Figure 1.11. TREND diagram of flow of participants with depression through study, including optional CBT portion. Bolded boxes on left side of diagram indicate final numbers for baseline analyses (top), analyses of correlations with treatment (middle), and predictors of treatment outcome (bottom).

Tables

Table 1.1: Baseline clinical and demographic data

	Controls	**Depression**	**Group Comparison**
Age	32.38 (1.90)	35.42 (1.38)	$t_{100} = -1.36$, $p > .1$
Gender (# [%])	20 (62.5)	49 (71.0)	$\chi^2_1 = 0.392$, $p > .1$
Estimated IQ	107.3 (2.30)	107.6 (1.30)	$t_{100} = -0.098$, $p > .1$
Depression severity[a]	2.0 (0.465)	31.16 (0.963)	$t_{100} = -20.2$, $p < .001$
Anhedonia[b]	44.75 (1.72)	83.83 (1.16)	$t_{100} = -19.1$, $p < .001$
Negative affect[c]	23.06 (1.05)	45.99 (1.09)	$t_{100} = -13.2$, $p < .001$

All values are mean (SE) unless otherwise noted.

[a]Beck Depression Inventory total; [b] MASQ anhedonic depression subscale; [c]MASQ general distress subscale

Table 1.2. Reward prediction error, MDD group (n = 69)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Right ventral striatum *Ventromedial prefrontal cortex*	14 *-12*	6 *38*	-12 *-10*	6.04 *4.74*	2001
2	Left ventral striatum	-16	6	-12	5.53	812
3	Left precuneus	-4	-50	32	5.3	848
4	Left superior parietal lobule	-32	-78	46	4.43	323
5	Left cerebellum	-14	-82	-24	3.94	213

Table 1.3: Reward expected value, MDD group (n = 69)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Right fusiform gyrus	32	-48	-18	-7.34	20497
	Left inferior parietal lobule	*-34*	*-62*	*44*	*-7*	
	Right occipital lobe	*36*	*-78*	*26*	*-6.87*	
2	Right middle frontal gyrus	48	48	18	-6.56	7339
	Right middle cingulate gyrus	*6*	*28*	*36*	*-6.5*	
3	Right calcarine sulcus	26	-38	18	6.33	608
4	Left calcarine sulcus	-26	-42	10	6.2	621
5	Right striatum	16	4	-4	-6.17	3403
	Left thalamus	*-6*	*-10*	*-4*	*-5.45*	
6	Left middle frontal gyrus	-52	24	34	-5.74	1335
7	Left insula	-30	20	4	-5.59	304
8	Right superior temporal gyrus	58	-32	14	5.34	1990
9	Left medial frontal gyrus	-8	60	16	5.17	1235
10	Left superior temporal gyrus	-54	-32	14	4.8	683
11	Right postcentral gyrus	16	-46	68	4.6	514
12	Right subgenual cingulate	6	24	-4	4.58	239
13	Left postcentral gyrus	-18	-48	68	4.38	160

Table 1.4. Reward prediction error, nondepressed controls (n = 32)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Left striatum	-24	0	-4	7.65	21994
	Right caudate	*20*	*4*	*16*	*6.57*	
2	Right cerebellum	20	-80	-28	7.31	8522
	Left middle temporal gyrus	*-60*	*-44*	*-10*	*6.94*	
	Left cerebellum	*-36*	*-72*	*-46*	*6.57*	
3	Left angular gyrus	-50	-68	26	5.68	5614
	Posterior cingulate gyrus	-2	-36	38	5.2	
4	Right middle frontal gyrus	28	38	46	5.19	893
5	Right middle temporal gyrus	58	-38	-12	4.26	213

Table 1.5. Reward expected value, nondepressed controls (n = 32)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Right inferior parietal lobule	42	-62	42	-7.44	2077
2	Right middle frontal gyrus *Right middle cingulate gyrus*	46 8	10 18	34 44	-6.94 -6.58	4764
3	Midbrain	-8	-14	-12	-6.9	424
4	Left inferior frontal gyrus	-42	46	12	-6.29	745
5	Right precentral gyrus	20	-20	78	6.21	4684
6	Right insula	32	22	-8	-5.98	437
7	Right thalamus	16	-32	20	5.94	265
8	Left inferior parietal lobule	-36	-60	42	-5.62	1321
9	Left cerebellum	-38	-62	-50	-5.55	1909
10	Left thalamus	-18	-36	14	5.39	284
11	Right superior temporal gyrus	54	-10	-2	5.25	517
12	Right fusiform gyrus	32	-60	-10	-4.93	1154
13	Left inferior frontal gyrus	-46	2	30	-4.84	183
14	Subgenual anterior cingulate cortex	6	28	-4	4.78	170
15	Left insula	-32	18	-8	-4.55	284
16	Ventromedial prefrontal cortex	0	48	-20	4.46	582
17	Right striatum	16	4	-2	-4.4	151
18	Left precuneus	-8	-54	26	4.19	269

Table 1.6: Loss outcome value correlated with negative affect (MASQ Mixed Distress subscale; n = 101)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Left precuneus	-6	-64	44	-4.46	277
2	Subgenual anterior cingulate	4	32	-2	-4.39	976

Table 1.7: Loss outcome value, low negative affect participants (n = 52)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Left middle frontal gyrus	-20	32	44	6.5	848
2	Right superior frontal gyrus	8	8	60	-5.87	869
3	Ventromedial prefrontal cortex	4	54	-10	4.86	1287
4	Right supplementary motor area	4	-26	58	4.63	1357
5	Right ventral striatum	12	4	-14	4.61	190
6	Left precuneus	-12	-46	40	4.44	578
7	Right cerebellum	52	-72	-42	4.39	200

Table 1.8: Loss outcome value, high negative affect participants (n = 49)

Cluster Number	Region	Peak MNI Coordinate			Peak T Value	Cluster Size
1	Right supplementary motor area	12	6	68	-6.09	2417
2	Right insula	42	16	-6	-5.23	430
3	Right precentral gyrus	24	-10	36	5	618
4	Right inferior parietal lobe	40	-40	26	4.68	269
5	Left insula	-40	24	0	-4.49	283
6	Left postcentral gyrus	-34	-28	32	4.27	381

Paper 2 (Study 3): Towards Learning Retraining: Creating Targeted Changes in Reinforcement Learning

Vanessa Brown, Jacob Lee, Brooks King-Casas, & Pearl H Chiu

Abstract

Disruptions in learning and valuation are central features of many forms of psychopathology and may be responsive to existing treatments. However, these disruptions may be more directly targeted through retraining techniques. In the current study, a novel approach to learning retraining was tested in a large online sample in participants with a range of depressive symptoms. Participants were asked targeted questions based on reinforcement learning theory and parameters of a reinforcement learning model were compared between conditions with different queries versus a no-query control condition. Different queries significantly changed different parameters in different directions, an effect that was not moderated by depression severity. These findings point toward the potential to use this approach to change disrupted learning parameters in disorders like depression.

Introduction

Disruptions in learning and valuation are central features of many forms of psychopathology, including reduced processing of positive information in anhedonia and altered learning about negative events in anxiety and heightened negative affect (Bouton, Mineka, & Barlow, 2001; Eshel & Roiser, 2010). Recent work in computational modeling of disrupted learning processes suggests that alterations in algorithmically defined learning parameters can predict and correlate with symptom improvement from existing treatments (Huys et al., 2016; Vrieze et al., 2013). Directly targeting disrupted learning processes has the potential to more precisely and effectively treat psychopathology; however, approaches that lead to robust, generalizable changes in learning are lacking. In the current study, we set out to develop a learning retraining paradigm to test the feasibility of altering learning in a systematic, enduring fashion.

Experimental work aimed at understanding which aspects of learning and decision making are disrupted in psychiatric disorders has taken advantage of the precision afforded by computational models, particularly reinforcement learning, that describe these processes. Reinforcement learning is a mathematically formalized, neurobiologically grounded framework to describe value-related learning processes (Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997; Sutton & Barto, 1998) and provides a method to specify and understand learning across levels of analysis (Niv & Langdon, 2016). Reinforcement learning models formalize learning as the process of updating values associated stimuli based on the discrepancy between the experienced and expected outcomes (known as prediction error). Free parameters in this model, which can be

estimated based on learning behavior or used in simulations to understand varied influences on learning, index how much prediction error is used to update values on each trial (learning rate) and how much the values or probabilities of different outcomes are dissociated (outcome sensitivity or inverse temperature), among others (Daw, 2011).

If reinforcement learning models can represent learning and decision making dysfunctions that characterize psychiatric disorders, understanding how to change aspects of reinforcement learning would lead to new insights in treating these disorders; however, how to create persistent changes in learning behavior is unclear. Previous work has shown that people can change learning behavior within a task under certain conditions. Instructed knowledge about tasks, such as giving information about the probabilities of outcomes during learning, changes participants' behavior such that they incorporate this explicit knowledge instead of relying solely on experienced prediction errors for learning (Atlas, Doll, Li, Daw, & Phelps, 2016; Li, Delgado, & Phelps, 2011). Meanwhile, implicitly altering aspects of the task environment, such as modulating the volatility of outcomes (Behrens, Woolrich, Walton, & Rushworth, 2007; Pulcu & Browning, 2017), or re-presenting past choices (Bornstein, Khaw, Shohamy, & Daw, 2017), changes behavior in accordance with these environmental manipulations. These findings provide support that learning can be systematically changed; however, changes in learning from these explicit or implicit changes is unlikely to persist in the absence of instructions (in the case of explicit changes) or environmental alterations (for implicit changes).

A middle ground of learning training would provide people with guidance about changing learning, similar to instructed learning, while allowing them to formulate independent adjustments to meet their goals, similar to responding to changing

environments. Existing approaches to behavior change in psychopathology, such as cognitive behavioral therapy, effect change through raising awareness of relevant aspects of behavior through targeted questions while allowing patients to generate and practice new behaviors to use in a variety of situations (Beck, 2011). In these approaches, interventions that are explicit to the learner and can be used in other situations are an important component of generalization and eventual treatment success (Swan, Carper, & Kendall, 2016). Therefore, we set out to test whether a similar approach would be effective in changing reinforcement learning.

In the current study, we aimed to establish whether altering reinforcement learning through explicit reminders is a feasible approach to targeting changes in learning. We tested a suite of specific approaches to determine which were powerful enough to change learning in a single session, with the eventual goal of extending this paradigm to target psychopathology-related learning dysfunctions over multiple sessions. We hypothesized that querying about task components related to aspects of learning would change these facets of behavior while leaving other learning processes intact. Specifically, we hypothesized that querying about prediction error or components of prediction error would increase learning rate and that querying about the probability or value of outcomes would increase outcome sensitivity while querying about the differences in value between options would decrease outcome sensitivity.

Results

Participants completed fifty trials of a common task used to measure reinforcement learning (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). To create an explicit

learning manipulation, participants were queried every three trials about a reinforcement learning-related aspect of the task (Figure 2.1a). Participants were assigned to one treatment upon starting the task and each treatment consisted of one type of query only. Queries were designed to test effects on components of reinforcement learning, including value estimation, probability estimation, updating values of chosen and unchosen options, comparison of values, and prediction error (see Table 2.1 for the full list of treatments). Some queries were positively valenced (asking about better outcomes, high value options, etc.), while others were negatively valenced (queried worse outcomes, low value options, etc.). Participants responded to queries by moving a slider bar to their answer, but did not receive any feedback on their answers to queries. Participants assigned to the active control treatment were asked to move the slider bar to a certain point rather than answer a learning-related query.

Participants were recruited through Amazon's Mechanical Turk platform. After basic data cleaning and manipulation checks (excluding participants with poor performance or incomplete learning data, see Methods for further details), 1,299 participants (approximately 100 per query; 200 in the active control condition) were included in analyses. Simulation-based power analyses based on parameters previously estimated from Study 1 indicated the study was sufficiently powered to detect a small to medium effect size of changes in learning parameters with queries. Participants reported a range of depression, anxiety, and stress symptoms, including a high proportion of participants reporting clinically elevated symptoms (22-32% of sample, consistent with previous reports in online crowdsourced populations [Arditte, Demet, Shaw, & Timpano, 2016; Chandler & Shapiro, 2016]; see Table 2.2 for additional demographic information).

We first assessed the acceptability and feasibility of our retraining task with participants. Overall, participants rated the task as engaging (mean = 6.5/10, mode = 10), interesting (mean = 6.3, mode = 10) and not difficult (mean = 3.0, mode = 0). Participants also rated their positive and negative affect before and after completing the test, and participants who rated the task as more engaging and interesting also reported increases in positive affect (*r*s = .40 and .44, ps < .001) and decreases in negative affect (*r*s = -.15 and -13, ps < .001), while those who rated the task as not difficult reported a decrease in negative affect (*r* = .08, p < .01; Figure 2.1b). Ratings of engagement, interest, and difficulty did not differ by treatment (ANOVA of effect of treatment on ratings: all Fs < 1.7, ps > .05). Participants displayed intact learning across all treatments (Figure 2.2), showing initial performance near chance, gradual improvement over time, and a plateau in performance near the matching ratio of 75% (Herrnstein, 1974) for most treatments. Accuracy on responses to queries was also high (Figure 2.2). To assess whether participants were focusing on answering queries accurately at the expense of performing the learning task well, we examined the relationship between choice accuracy on the learning task and query accuracy. Overall, participants who performed better at the learning task had more accurate responses to queries (*r* = .138, p < .001; Figure 2.5a), suggesting that accurate performance on the learning task and attending to queries were not at odds. Together, these data confirm that participants found the task acceptable and displayed both good performance on the learning task and accurate responses to queries, supporting the use of the task in retraining approaches.

Examination of learning curves also showed differences in learning patterns by treatment; to quantify these learning differences, we assessed the effect of queries on

reinforcement learning parameters. We estimated learning parameters of learning rate (measuring the speed of value updating based on prediction error) and outcome sensitivity (measuring the relative valuation of high versus low outcomes), based on a RL model previously validated in this learning task and sensitive to depression-related alterations in reward learning (see Paper 1). Parameters were estimated using hierarchical Bayesian estimation within a model contrasting parameters for each treatment group against the active control group, covarying for depression symptoms. Posterior probability distributions of the effects of each treatment on parameters are plotted in Figure 2.3, with significant effects plotted with opaque colors. The treatment with queries assessing value estimation of the most recently chosen option resulted in a higher learning rate (transformed mean effect on learning rate = 1.05, 95% credible interval [CI] .182 to 1.95), while queries assessing the value of the unchosen option resulted in a lower learning rate (transformed mean effect = -1.01, 95% CI -2.00 to -0.021). For outcome sensitivity, queries asking about the probability of both high (mean effect = 0.225, 95% CI 0.071 to 0.419) and low outcomes (mean effect = 0.174, 95% CI 0.015 to 0.415), the least ever received for an option (mean effect = 0.214, 95% CI 0.066 to 0.412), and the value of the high value option (mean effect = 0.165, 95% CI 0.054 to 0.283) all increased outcome sensitivity. Next, to determine the effects of the overall valence of the query, queries were grouped by valence and the effects of all positively and all negatively valenced queries were compared to the active control (Figure 2.4). Both groups of valenced queries increased outcome sensitivity (positively valenced queries: mean effect = 0.123, 95% CI 0.041 to 0.205; negatively valenced queries: mean effect = 0.114, 95% CI 0.019 to 0.209), while negatively valenced queries selectively reduced learning rate (positively valenced queries: mean effect = -0.102,

95% CI -0.839 to 0.629; negatively valenced queries: mean effect = -0.861, 95% CI -1.60 to -0.117).

To determine if depressive symptoms affected the efficacy of the learning manipulation, we then examined the effect of depression as a covariate. Depression was not related to treatment effects on parameters for any treatment (all 95% credible intervals encompassing 0), suggesting that the treatment effects were consistent across levels of depression. This effect was similar when the depression covariate was binarized to categorize reported severity as above versus below the clinical cutoff (Lovibond & Lovibond, 1995). We additionally tested the effects of depression on query accuracy; depression was related to slightly worse query accuracy overall ($F_{1,1232} = 5.50$, $p < .05$) and showed a trending interaction with treatment type ($F_{11,1232} = 1.67$, $p < .1$; Figure 2.5b). Participants reporting higher depression symptoms were more accurate when queried about the chance of getting the larger outcome ($t_{1232} = 2.01$, $p = .04$) but less accurate when queried about the value of the last unchosen option ($t_{1232} = -2.09$, $p = .04$). For choice accuracy on the learning task, there was a main effect of depression ($F_{1,1232} = 7.414$, $p < .01$; Figure 2.5c), but no interaction with treatment type ($F_{11,1232} = 0.414$, $p > .1$), indicating that participants with higher depression were less accurate on the learning task regardless of the queries asked.

To test the effects of query dosage on changes in learning, we enrolled an additional set of participants who completed the same learning task with queries, but with queries either every trial or every three trials throughout the task. There was no main effect of query frequency on choice accuracy on the learning task (Figures 2.6a and 2.6b; $F_{1,520} = 2.62$, $p > .1$), but there was an interaction of query frequency and treatment on choice

accuracy ($F_{11,520} = 2.04$, $p < .05$), due to slightly worsened accuracy when queried every versus every third trial about the chance of getting the high value outcome ($t_{520} = -1.78$, $p < .1$), and improved accuracy when queried about the value of the most recently chosen option ($t_{520} = 2.46$, $p < .05$). There was both a main effect ($F_{1,520} = 6.61$, $p < .05$) of query frequency and interaction with treatment ($F_{11,520} = 2.11$, $p < .05$) on proportion switching (Figures 2.6a & 2.6b), with more switching overall and particularly when queried every versus every third trial on the chance of getting the low value outcome ($t_{520} = 3.03$, $p < .005$) and on the value of the last unchosen option ($t_{520} = 2.38$, $p < .01$). Querying every trial versus every three trials was related to an increased outcome sensitivity for the treatment querying the value of the last chosen outcome (Figure 2.6c; mean interaction effect = 0.373, 95% CI -0.742 to -0.027), but otherwise did not lead to significant changes in reinforcement learning parameters. Therefore, the overall effect of querying more frequently was to increase switching, particularly for treatments querying about low value or unchosen options; conversely, query frequency had a specific effect within the treatment querying about the value of the most recently chosen option such that greater query frequency increased choice accuracy and outcome sensitivity on the learning task.

Discussion

Understanding how and under what circumstances reinforcement learning parameters can be changed is vital to better comprehension of learning-related changes and deployment of learning retraining approaches to treat disorders characterized by learning dysfunctions. In the present study, we found that querying participants about reinforcement learning-related task components, in the absence of feedback on these queries, robustly altered learning in directions that were specific to the type and valence of

query. These effects on learning parameters were unaffected by level of depression, suggesting generalizability to samples with psychopathology, and participants' behavior and reports showed the task to be feasible and tolerable.

Queries had specific, separable effects on reinforcement learning parameters. Learning rate increased relative to the control group in the treatment querying participants about the expected value of their most recent choice, while this parameter was lower when participants were queried about the value of the option they had not chosen. Surprisingly, queries more directly assessing prediction error or recent chosen outcomes did not affect learning rate. This pattern of results suggests that some components of prediction error, such as past expected value, may be incorporated more into learning when participants are queried about these components, whereas others, such as outcome value, may be less affected by reminding via queries. Supporting this, querying participants about the average value of the option just chosen did not increase learning rate, suggesting this information may already be used during learning; however, querying about the average value of the option not chosen did decrease learning rate, suggesting that recalling the values of other options did disrupt updating of the chosen option. Similarly, when collapsing across queries by valence, negatively valenced queries related to low value options, probability of receiving low values, the value of unchosen options, and similar queries decreased learning rate, lending support to the finding that asking about negative outcomes or unchosen options disrupts value updating.

For outcome sensitivity, querying about the probability of outcomes increased this parameter; in addition, asking about the value of the high value option as well as the least ever received when choosing an option increased outcome sensitivity, while no queries

decreased outcome sensitivity. Another treatment querying about the most recently chosen option's value also showed increased outcome sensitivity when queried every trial versus every three trials. Additionally, when collapsing across queries by valence, both positively and negatively valenced queries increased outcome sensitivity relative to the active control condition. Queries focused on probability or extremes of value may increase outcome sensitivity more than other queries, but this pattern of results also suggests that most learning-related queries increase outcome sensitivity to some degree. The effect of queries about comparing values, which we hypothesized would decrease outcome sensitivity, did not reach significance, but was the only query to qualitatively decrease this parameter. Therefore, focusing on single options, regardless of the specific query, may increase outcome sensitivity while focusing on comparing options may have an opposite effect.

Our sample included a wide range of depression severity, allowing us to examine whether self-reported depression symptoms influenced the effect of learning queries. We did not find an overall effect of depression or an interaction with the effect of any treatment, suggesting that this approach to changing learning is equally effective in depression-related psychopathology. In line with previous findings (Blanco, Otto, Maddox, Beevers, & Love, 2013; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008; but see Gradin et al., 2011; Rothkirch, Tonn, Kohler, & Sterzer, 2017), we did find significant relationships between depression severity and both performance on the learning task and choice accuracy. However, the effect sizes for these relationships were small, suggesting that even participants high in depression could complete all components of the task effectively. Providing further support for the feasibility of this task, our task was well-tolerated by participants as well as effective in changing learning. Participants rated the task as

engaging, interesting, and not difficult, and the vast majority of participants passed our quality checks, performed accurately on the learning task, and answered queries accurately. These results are encouraging for future applications of learning retraining-related tasks in treatment settings (Paulus, Huys, & Maia, 2016). Future work will need to build on the present findings to explore the feasibility and efficacy of longer sessions or those done on a repeat basis.

In addition to providing support for using learning queries to alter behavior, these findings also shed light on basic mechanisms of reinforcement learning. Participants' choices, driven by the expected value of these choices, changed in response to targeted queries, without any feedback-driven learning from the queries themselves. These shifts in value and learning suggest that retrieved value during decisions incorporates past experiences in a malleable way and adds to the growing literature on the interaction between reinforcement learning and memory representations (Gershman & Daw, 2017; Shohamy & Daw, 2015). However, these shifts did not always occur in predictable ways. For example, querying about prediction error, which most directly relates to the learning rate parameter, did not change learning rate. This facet of our results suggest that altering learning is a more complex process than previously thought and that therapeutic approaches to targeting learning (e.g. Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014) should thoroughly investigate how to bring about intended changes in learning.

In summary, we found that explicit learning-related queries are effective in changing learning parameters in specific ways, effects that are present across a range of depressive symptomatology and are highly acceptable to participants. These findings lay

the groundwork for future studies exploring the effects of repeated applications of learning retraining paradigms on real-world behavior and affect.

Methods

Participants

Participants were recruited via Amazon's Mechanical Turk platform. Per current Mechanical Turk requirements, participants were required to have an IP address based in the United States. To be eligible, participants affirmed that they had fluent English and were at least 18 years old. Participants provided informed consent and all procedures were approved by the Institutional Review Board at Virginia Tech. Participants were compensated with a base payment of $0.25 plus a performance bonus of the sum of three randomly selected outcomes from the learning task. Participants were paid the performance bonus if their performance passed certain eligibility screens (see below for details); the bonus averaged $1.55 (range $0.48 to $2.48).

Participants were recruited until at least 100 usable participants had completed each treatment (200 for the active control group). A priori power analyses were conducted by simulating effects of varying sizes (.2 to .8) and various numbers of participants on changes in parameters; parameter values for the control group were based on a previous study with a similar learning task and were set at (transformed) learning rate = -0.898 and outcome sensitivity = 1.162. Based on this simulation, 100 participants per treatment group was sufficient to detect effects sizes .4 (small/medium) and above.

Task

The task was composed of two parts: participants completed fifty trials of a probabilistic two-choice learning task commonly used to assess reinforcement learning (Pessiglione et al., 2006), interspersed every three trials (or every trial when assessing the effects of query frequency) with a query. The task and a brief description are shown in Figure 2.1. Upon enrollment, participants were randomly assigned to a treatment; each treatment's queries were unique and each treatment contained only one type of query (see Table 2.1 for a full list of treatments and their associated queries). This method of assignment meant that each participant was asked only one type of query repeatedly throughout the task.

On the learning portion of the task, participants were presented with two stimuli (a clover and a club) and instructed to choose one stimulus with their computer mouse or keyboard. After the selection, the chosen stimulus changed color and the outcome associated with that stimulus was shown. Higher outcomes ranged from $0.65 to $0.85 (chosen each trial from a uniform distribution in this range) and lower outcomes ranged from $0.15 to $0.35 (similarly chosen from a uniform distribution). One stimulus (the 'better option') was associated with a higher probability (75%) of the higher outcome and a lower probability (25%) of the lower outcome, while the other stimulus (the 'worse option') had reversed probabilities. Correct choices on the learning task were defined as choosing the better option, regardless of whether the resulting outcome was high or low.

After every third trial, participants were given a query after viewing the outcome of their choice. The query text was displayed at the top of the screen, the two stimuli in the middle, and a slider bar and a 'submit' button at the bottom. The response for the slider bar was randomly initialized at a different location on the slider bar for each query

presentation. The participant was instructed to move the bar to the desired response and to click the 'submit' button. Anchor values for the slider bar were displayed on each end and were based on the specific query (e.g. when queried about the value of the last chosen option the range was specified as between 0 and 100 cents, while when queried about the difference in value between the two options the range was -100 to +100 cents). For queries where the stimulus queried based on participants' behavior (e.g. the last chosen option), participants were asked to estimate the value of the symbol for that stimulus rather than being asked to estimate the last chosen option directly. This method of querying removed confounds related to remembering specific stimuli and to participants' knowledge of the specific query condition. For the active control group, participants were queried at an equal frequency but were simply asked to move the slider bar to a certain point as indicated by an arrow and were not queried about any task-related components.

Measures

Participants in the main analyses completed the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) before and after the task. Changes in these scales were defined as post-task minus pre-task, all divided by pre-task scores. Participants also completed the short (21 question) version of the Depression Anxiety and Stress Scale (DASS; Antony, Bieling, Cox, Enns, & Swinson, 1998; Lovibond & Lovibond, 1995), which includes subscales for depression, anxiety, and stress. Scores on each subscale were doubled to be consistent with the long (42 question) version of the DASS and clinical cutoffs were applied per the DASS manual. Participants also reported their level of engagement, interest, and difficulty with the task as well as basic demographic information.

Participants in the group recruited to test query frequency did not complete the DASS and pre-task PANAS.

Data cleaning and preprocessing

Participants were paid a participation bonus if they moved the slider during the task and chose the correct option at least 30% of the time. Sixty enrolled participants were excluded at this stage of processing for not meeting these performance criteria. An additional 43 participants completed the task but did not complete post-task questionnaires or demographic information and so were not paid the participation bonus. These participants were included in analyses except for those requiring demographic or questionnaire information. Of these 43 participants, 37 of these participants also did not provide ratings of engagement, interest, and difficulty. To ensure participants were attempting to learn the task sufficiently, additional data retention criteria included: completion of at least 30 (of a total of 50) learning trials, minimum accuracy of 35% on the learning task, and switching between options at least three times. Eighty-seven participants were excluded at this stage. These data cleaning steps resulted in 1,299 participants' data retained for analyses (90% of originally collected sample).

To quantify accuracy on queries, the correct answer to the queries was defined as what the participant had experienced up until that trial. Therefore, if the query asked to rate the chance that the better option stimulus would lead to the high value outcome, the correct answer was defined as the proportion of times this had occurred during the learning task to that point, rather than the predefined probability of 75%. If the correct answer could not be calculated for a specific query instance (e.g. that particular stimulus

had not been selected yet), that query instance was excluded for that participant (1 participant could not have query accuracy calculated for this reason and was excluded from analyses of query accuracy). The deviation from this correct answer was calculated per query (plotted in Figure 2.2). The absolute value of the distance from the correct answer was summed over all 16 instances of the query and used as the quantitative measure of query accuracy for analyses. Since a larger value for this measure indicates worse accuracy, correlations and t-tests are reported with signs reversed to aid in interpretation (e.g. a reported positive correlation between choice accuracy and query accuracy indicates that participants with higher choice accuracy had a lower summed absolute distance from the correct answer for queries, and therefore higher query accuracy).

Analyses of summary statistics

Analyses used Pearson correlations to assess relationships between continuous variables (choice accuracy, query accuracy, depression, change in affect, and subjective ratings). When assessing the effect of treatment, ANOVAs were run with treatment as a factor and the active control group as the reference group. If the ANOVA showed a significant relationship, a regression was run with each treatment group as a separate variable to assess the effects of individual treatments. Significant results were defined as alpha < .05.

Analyses of reinforcement learning parameters

Parameters were estimated using hierarchical Bayesian estimation as implemented in Stan (Carpenter et al., 2016). Learning rate and outcome sensitivity were specified as

normally distributed at the group level with a non-centered parameterization to aid in estimation. Learning rate was transformed with a logistic transformation after specification to constrain values between 0 and 1. The prior distribution for the mean of the group level parameter was set at N(0,2.5) for learning rate and N(0,10) for outcome sensitivity and the prior for the variance was set at Cauchy(0,2) and Cauchy(0,2.5), respectively. The individual variance, effect of each treatment, effect of covariates (i.e. depressive severity), and interaction of covariates and treatment effects were all specified with a prior of N(0,1). We note that priors in this type of Bayesian analyses are not intended to strongly effect the posterior estimates but rather to regularize estimates to be in an interpretable range for the variables being estimated (Gelman et al., 2014). As learning plateaued partway through the learning task, the first 25 trials were used for parameter estimates only. A changepoint analysis (Killick & Eckley, 2013) found one changepoint after trial 22, confirming differences in learning trajectories in the two halves of the task. Four MCMC chains were run for 4000 samples each for each analysis. The first 2000 samples of each chain were discarded as warm-up, resulting in 8000 samples for analysis. All chains were inspected for convergence and showed good mixing, with all values of the potential scale reduction factor below 1.1 (Gelman & Rubin, 1992).

The primary analysis included variables for each treatment (coded 1 for participants in that treatment and 0 otherwise, with the active control as the reference group) and the interaction of treatment with depression severity. Depression severity was z-scored prior to entering in the model. Significance was defined as the 95% credible interval of the posterior probability distribution of an effect falling completely outside 0 (e.g. entirely above or below 0).

References

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, *10*(2), 176–181. http://doi.org/10.1037/1040-3590.10.2.176

Arditte, K. A., Demet, C., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological Assessment*, *28*(6), 684–691. http://doi.org/10.1037/pas0000217

Atlas, L. Y., Doll, B. B., Li, J., Daw, N. D., & Phelps, E. A. (2016). Instructed knowledge shapes feedback- driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *eLife*, *5*(MAY2016), 1–26. http://doi.org/10.7554/eLife.15192

Beck, J. S. (2011). *Cognitive behavior therapy: Basics and beyond*. Guilford Press.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. http://doi.org/10.1038/nn1954

Blanco, N. J., Otto, A. R., Maddox, W. T., Beevers, C. G., & Love, B. C. (2013). The influence of depression symptoms on exploratory decision-making. *Cognition*, *129*(3), 563–568. http://doi.org/10.1016/j.cognition.2013.08.018

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, *8*(May 2015), 15958. http://doi.org/10.1038/ncomms15958

Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological Review*, *108*(1), 4–32.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2016). Stan: A probabilistic programming language. *Journal of Statistical Software*, *20*.

Chandler, J., & Shapiro, D. N. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*, 53–81. http://doi.org/10.1146/annurev-clinpsy-021815-093623

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. http://doi.org/10.1016/j.brat.2014.04.006

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning: Attention and performance XXIII* (pp. 3–38).

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, *68*(2), 118–24. http://doi.org/10.1016/j.biopsych.2010.01.027

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis* (Third). Boca Raton, FL: CRC Press.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–511. http://doi.org/10.1214/ss/1177011136

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. http://doi.org/10.1146/annurev-psych-122414-033625

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., … Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, *134*, 1751–64. http://doi.org/10.1093/brain/awr059

Herrnstein, R. J. (1974). Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior*, *21*(I), 159–164. http://doi.org/10.1901/jeab.1974.21-159

Huys, Q. J. M., Gölzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., & Dolan, R. J. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological Medicine*, *46*(5), 1027–35. http://doi.org/10.1017/S0033291715002597

Killick, R., & Eckley, I. (2013). changepoint: An R Package for changepoint analysis. *Lancaster University*, *58*(3), 1–15. http://doi.org/10.1359/JBMR.0301229

Li, J., Delgado, M. R., & Phelps, E. a. (2011). How instructed knowledge modulates the neural systems of reward learning. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(1), 55–60. http://doi.org/10.1073/pnas.1014938108

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales.* (2nd ed.). Sydney: Psychology Foundation.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic predictive hebbian learning. *The Journal of Neuroscience*, *76*(5), 1936–1947.

Niv, Y., & Langdon, A. (2016). Reinforcement learning with Marr. *Current Opinion in Behavioral Sciences*, 1–7. http://doi.org/10.1016/j.cobeha.2016.04.005

Paulus, M. P., Huys, Q. J. M., & Maia, T. V. (2016). A Roadmap for the Development of Applied Computational Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. http://doi.org/10.1016/j.bpsc.2016.05.001

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–5. http://doi.org/10.1038/nature05051

Pizzagalli, D. A., Iosifescu, D. V, Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced

hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87. http://doi.org/10.1016/j.jpsychires.2008.03.001

Pulcu, E., & Browning, M. (2017). Affective Bias as a Rational Response to the Statistics of Rewards and Punishmentsxt. *eLife*, 1–15.

Rothkirch, M., Tonn, J., Kohler, S., & Sterzer, P. (2017). Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Cerebral Cortex*, 1–11. http://doi.org/10.1093/cercor/bhw393

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9054347

Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*. Elsevier Ltd. http://doi.org/10.1016/j.cobeha.2015.08.010

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT Press.

Swan, A. J., Carper, M. M., & Kendall, P. C. (2016). In Pursuit of Generalization: An Updated Review. *Behavior Therapy*, *47*(5), 733–746. http://doi.org/10.1016/j.beth.2015.11.006

Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., … Claes, S. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biological Psychiatry*, *73*(7), 639–45. http://doi.org/10.1016/j.biopsych.2012.10.014

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and*

Social Psychology, *54*(6), 1063–1070. http://doi.org/10.1037/0022-3514.54.6.1063

Tables

Table 2.1: Treatment conditions. The query column shows the text presented to participants about the stimulus indicated in the stimulus queried column. Color indicates the color for that condition in figures.

Treatment	Query	Stimulus queried	Color	Valence	N
Control	Move the slider to the arrow.	--	Gray	--	199
Chance of high outcome	What is the chance of getting more than 50¢ if you choose...	Chosen on last trial	Light red	Positive	100
Chance of low outcome	What is the chance of getting less than 50¢ if you choose...	Chosen on last trial	Dark red	Negative	98
Most received	What is the most you have received for...	Chosen on last trial	Light orange	Positive	97
Least received	What is the least you have received for...	Chosen on last trial	Dark orange	Negative	102
Last chosen	On average, how much do you get when you choose...	Chosen on last trial	Light green	Positive	100
Last unchosen	On average, how much do you get when you choose...	Not chosen on last trial	Dark green	Negative	106
High value	On average, how much do you get when you choose...	Stimulus with higher probability of high outcome	Light blue	Positive	99
Low value	On average, how much do you get when you choose...	Stimulus with higher probability of low outcome	Dark blue	Negative	90
Value comparison	On average, how do the two options compare?	Both	Light purple	--	102
Value estimation	What is the average amount you expected to get when you picked...	Chosen on last trial	Medium purple	--	103
Prediction error (negative)	How much more or less did you expect to get when you picked...	Chosen on last trial	Dark purple	--	103

Table 2.2: Participant demographics.

Measure	Mean [SD]	Percent above clinical cutoff
Gender (# [%] female)	558 [44.4%]	--
Age	34.92 [13.85]	--
Education (# [%] college degree)	638 [50.8%]	--
Depression	7.90 [9.33]	31.9%
Anxiety	5.28 [7.21]	24.6%
Stress	9.07 [8.56]	22.1%
Positive Affect (before)	31.31 [8.82]	--
Positive Affect (after)	29.35 [10.35]	--
Negative Affect (before)	13.24 [5.18]	--
Negative Affect (after)	13.22 [5.10]	--

Depression, Anxiety, & Stress: subscales of Depression, Anxiety, and Stress Scale

Positive and Negative Affect: from pre-task (before) and post-task (after) administration of Positive and Negative Affect Schedule

note: 43 participants did not provide demographic data

Figures

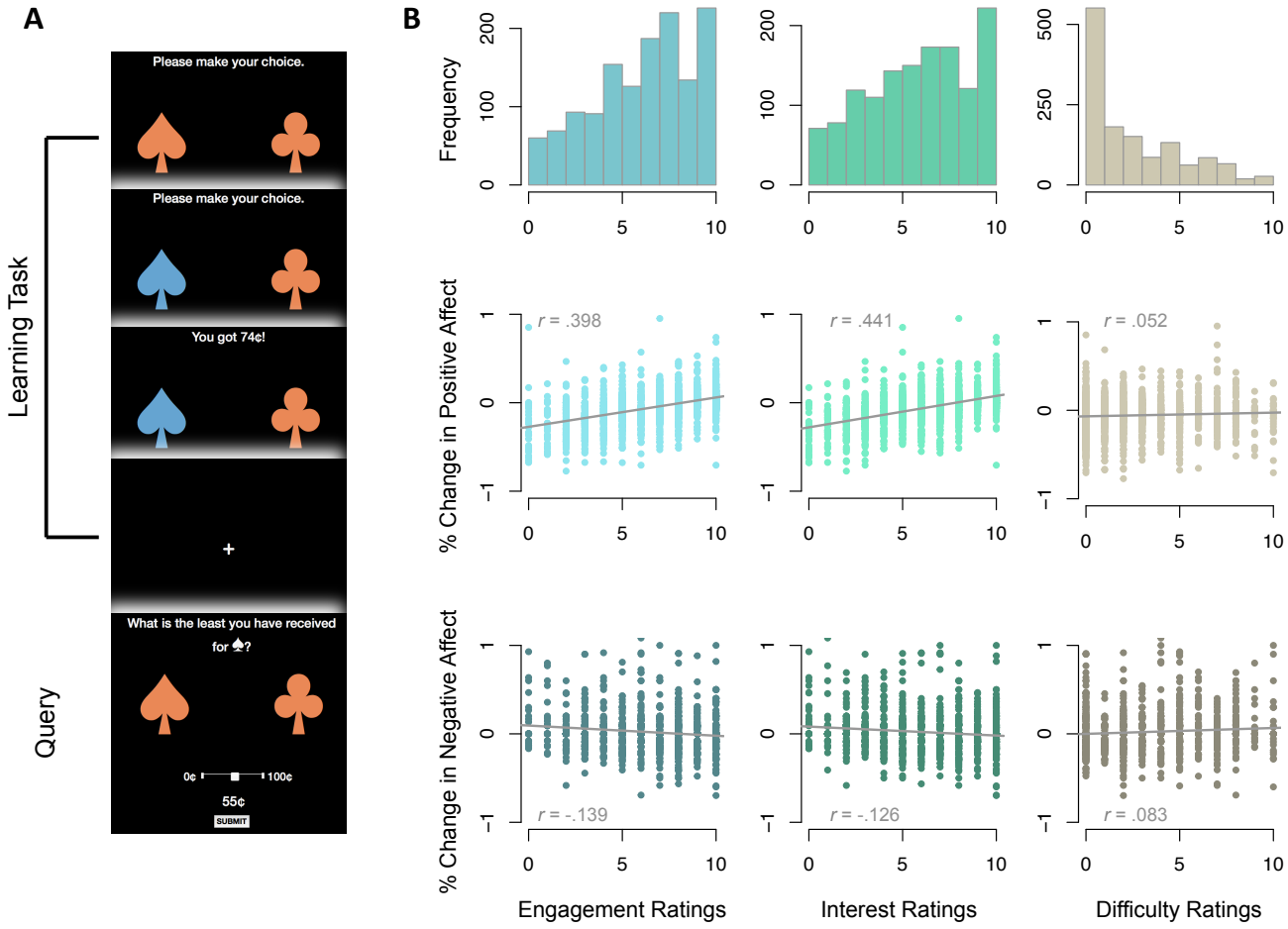

Figure 2.1: Schematic of learning task and query and related participant ratings. A) Participants completed 50 trials of a probabilistic two choice learning task where one outcome was associated with a higher probability (75%) of a larger reward (greater than 50 cents) and the other outcome was associated with a higher probability of a smaller reward (less than 50 cents). Every three trials, participants were queried on a specific aspect of the learning task and answered the query using a slider bar. Each participant was assigned to one treatment and each treatment used only one query throughout the task. Note that participants were not given feedback on their answers to queries. The query for

the control treatment asked participants to move the slider bar to a certain point without asking about any learning components. B) Participants rated their engagement (first column), interest (second column), and difficulty (third column) in the task. Overall, ratings were high on engagement and interest and low on difficulty (top row; histograms of ratings), indicating the task was acceptable to participants. Engagement and interest ratings were positively correlated with positive affect (middle row) and engagement and interest ratings were negatively correlated and difficulty ratings positively correlated with changes in negative affect (bottom row), confirming the validity of the ratings. For plots on bottom two rows, dots indicate individual values and gray lines indicate lines of best fit.

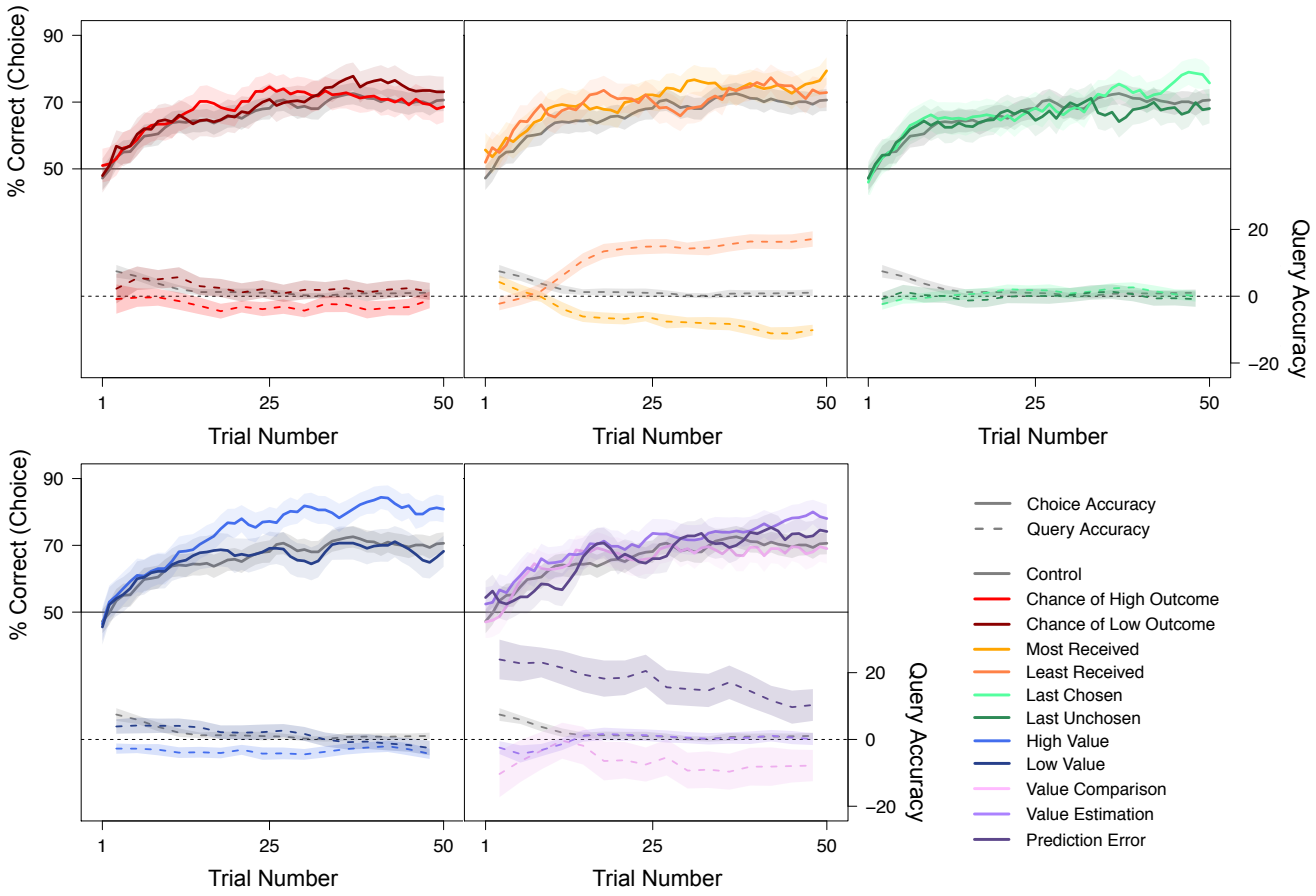

Figure 2.2: Learning curves and query accuracy over time by treatment. Solid lines (top) indicate percent correct choices on the learning task, while dotted lines (bottom) indicate distance from the correct value on queries (i.e., correct = 0); positive values on query accuracy indicate answers greater than the correct value while negative values indicate answers less than the correct value. Colored lines indicate treatments and the gray line indicates the control condition; values are group means and shading indicates one standard deviation. Across treatments, participants increased in accuracy from chance (50% correct) on the first trial, showing learning. A changepoint analysis confirmed that

learning asymptoted about halfway through the learning task (changepoint = trial 22), so trials in the first half of the learning task only were used to estimate learning parameters. Answers to queries were generally accurate but showed greater distance from 0 for queries requiring maintaining values in memory (most/least received treatments) or comparing values (value comparison & prediction error treatments).

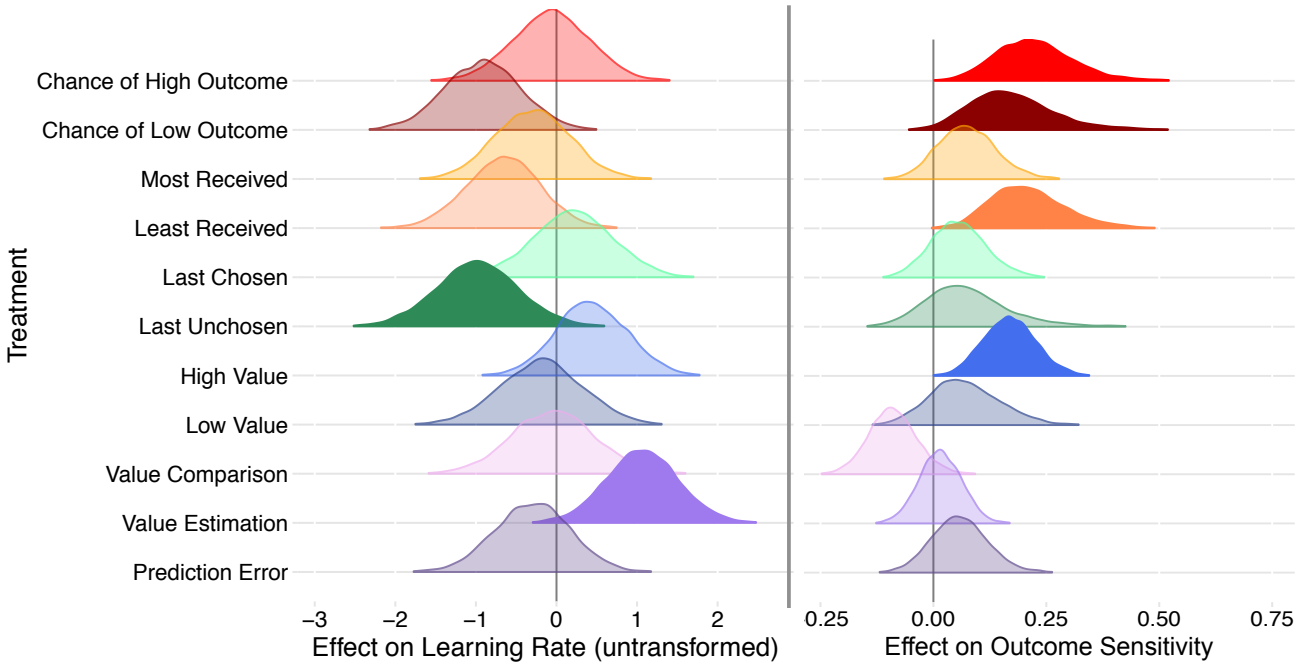

Figure 2.3: Parameter changes by treatment. Colored areas indicate posterior probability distributions of the effect of each treatment versus the control condition on each parameter. Solidly filled distributions have at least 95% of the distribution above or below 0, indicating a significant effect of that specific treatment on the parameter. Queries about the value of the unchosen option decreased learning rate, while querying estimates of expected value increased learning rate; queries about the option more likely to lead to the high value outcome, about the least ever received for an option, or about the probability of receiving either a high or low outcome all increased outcome sensitivity.

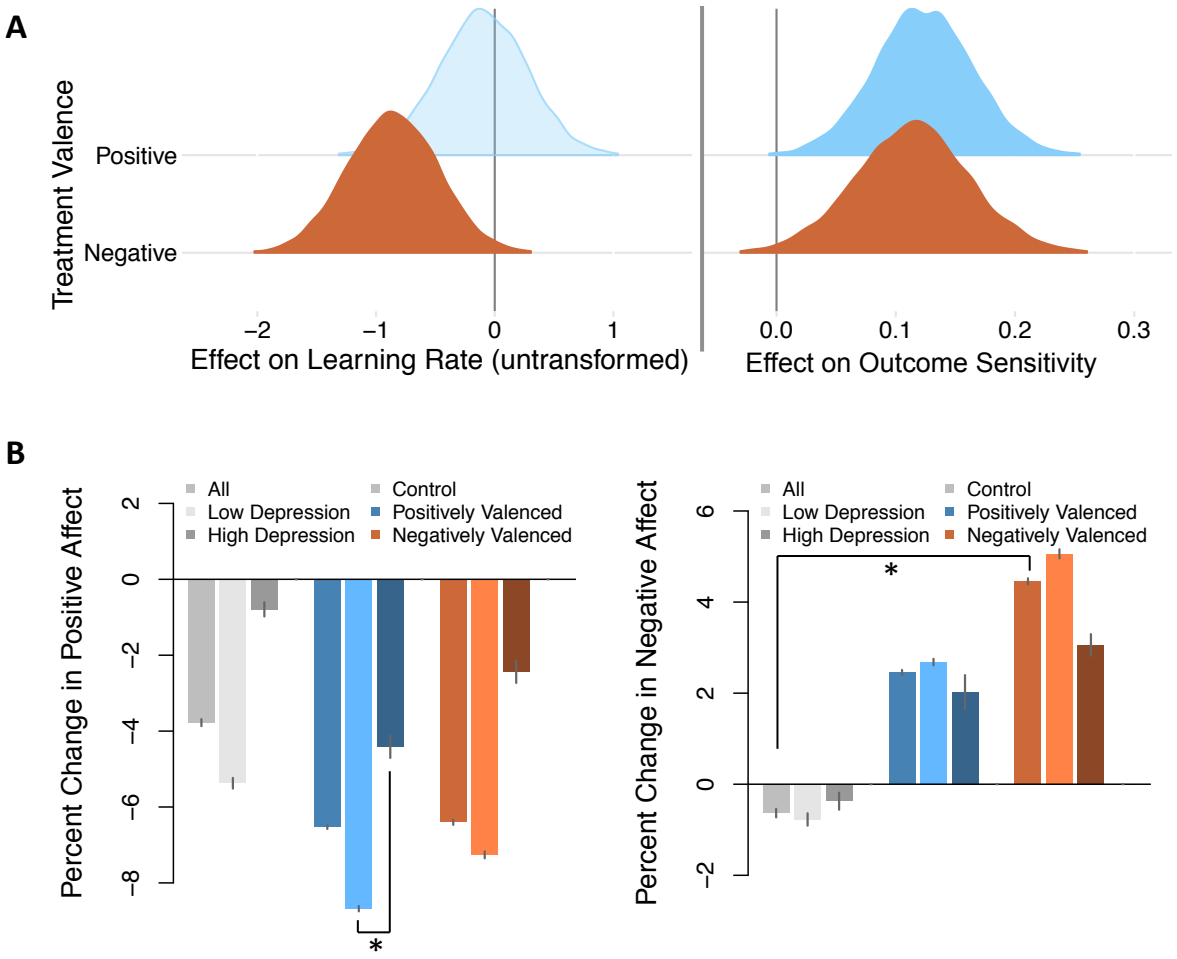

Figure 2.4. Parameter changes and changes in affect by treatment valence. A) Effects of treatment valence on learning parameters. Both positively and negatively valenced treatments increased outcome sensitivity, while negatively valenced treatments also decreased learning rate. B) Effects of treatment valence on changes in affect. Participants with depression levels above the clinical cutoff showed an improvement in positive affect after completing a positively valenced treatment, relative to participants with low levels of depression, while all participants showed a greater increase in negative affect in negatively valenced treatments relative to the control condition.

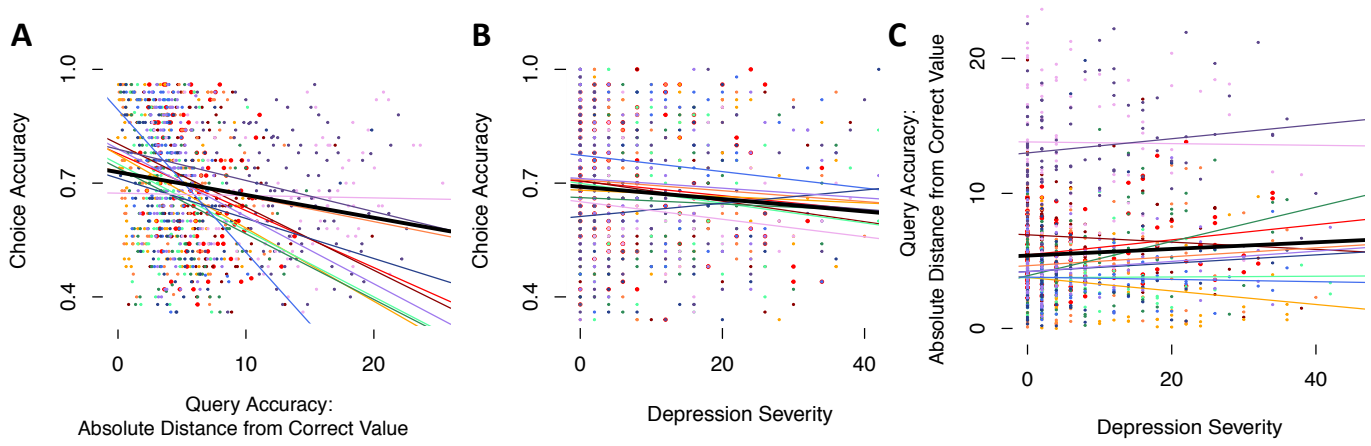

Figure 2.5: Relationship among query accuracy, choice accuracy on learning task, and depression. A) Relationship between query accuracy and choice accuracy. Dots are individual participants' mean values, colored lines are lines of best fit for individual treatments, and thick black line is line of best fit across all treatments. Higher values on query accuracy indicate a greater distance from the correct value (i.e., worse accuracy), while higher values on choice accuracy indicate better learning. Across all treatments, greater choice accuracy is related to better query accuracy (smaller distance from correct value); this relationship is present in most individual treatments as well. B) Relationship between depression severity and choice accuracy. Participants with higher levels of depression had worse performance on the learning task, although the size of the effect was small (R^2 = .006). C) Relationship between depression severity and query accuracy. Across treatments, depression severity was not related to query accuracy, but in individual treatments, higher depression was related to greater accuracy (smaller distance from correct) for the treatment querying the probability of the high value option and worse accuracy (larger distance from correct) when querying the value of the option not chosen most recently.

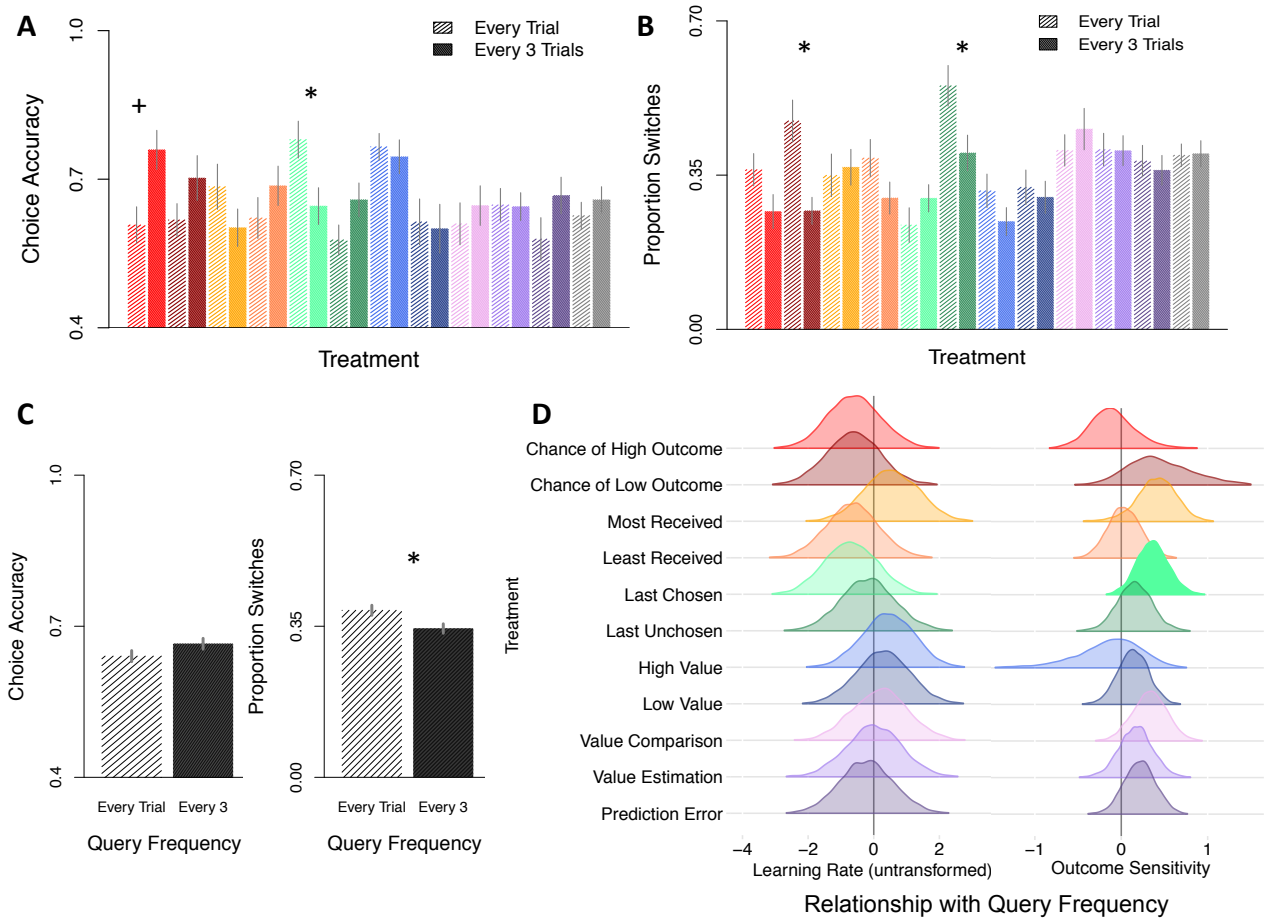

Figure 2.6: Effect of query frequency on learning. An additional ~20 participants per group completed the learning task with queries either every three trials (as in previous versions) or every trial. A) Effects of query frequency on choice accuracy on the learning task. Querying every trial resulted in worse accuracy when querying the probability of receiving a high outcome, a trend towards worse accuracy when querying the value of the option not chosen most recently, and better accuracy when querying the value of the most recently chosen option. B) Effects of query frequency on proportion of trials where participants switched choices. Querying every trial resulted in greater switches when querying the probability of receiving a low outcome and when querying about the value of the option not chosen most recently. C) Effect of query frequency on choice accuracy and

proportion switches collapsed across all treatments. Across all treatments, querying every trial significantly increased the proportion of switch trials. D) Effects of querying every trial versus every three trials, for each treatment compared to the control condition. Query frequency had little effect on learning rate but increased outcome sensitivity when querying about the last chosen option.

Discussion

Summary of studies

This set of studies aimed to add to the understanding of how computational models of reinforcement learning can inform knowledge of the mechanisms of depression and its treatment. In Study 1 (Paper 1), alterations in specific components of reward and loss learning according to a reinforcement learning model corresponded with symptoms of anhedonia and negative affect in depression on behavioral and neural levels. In Study 2, (Paper 1) these altered behavioral components changed along with improvements in symptoms after cognitive behavioral therapy, while pre-treatment neural indicators of reward learning predicted response to treatment. In Study 3 (Paper 2), targeted queries were successful in changing learning parameters in participants with a range of depressive symptoms, although some queries had effects that differed from what would be inferred from reinforcement learning theory.

In Study 1, behavioral reinforcement learning parameters and neural correlates of reinforcement learning during reward and loss learning differed by levels of anhedonia and negative affect symptoms. In reward learning, people with depression showed increased outcome sensitivity and reduced learning rate with increased anhedonia as well as disrupted relationships between neural measures of expected value and prediction error in the ventral striatum. These relationships between anhedonia and learning measures were only present in participants with a diagnosis of depression, suggesting that anhedonia in clinically depressed participants affects learning differently than participants with

anhedonia in the absence of mood-related psychopathology. In loss learning, greater levels of negative affect in all participants was related to a more negative valuation of losses as indexed by a more negative outcome shift parameter behaviorally and reduced subgenual anterior cingulate cortex and precuneus signaling to levels of outcome value neurally. Together, these findings show that depression affects learning from both rewards and losses, but that these effects differ by valence and symptom cluster.

In Study 2, participants previously diagnosed with depression completed the learning task again after completing a course of cognitive behavioral therapy to examine correlates and predictors of symptom change with treatment. The amount of improvement in anhedonia correlated with increased learning rate and decreased outcome sensitivity, while the amount of improvement in negative affect correlated with less negative outcome shift and reduced outcome sensitivity. When assessing these changes in light of the baseline differences found in Study 1, these findings indicate that symptom change with CBT led to remediation of altered learning patters found pre-treatment. People high in anhedonia also showed normalization of altered neural relationships between striatal signals of prediction error and expected value. Behavioral learning measures and striatal prediction error-expected value relationships did not show significant changes in control participants who did not undergo CBT, suggesting that these measures are stable over time when assessed independent of symptom change. Striatal signals of prediction error and expected value also predicted improvement in anhedonia, with greater prediction error signal and lesser expected value signal pre-treatment related to greater improvement in anhedonia with treatment. The combined results from Study 2 suggest that effective CBT is related to changes in disrupted learning parameters in both reward and loss learning and

that neural measures of reward learning may predict improvement in reward-related symptoms of anhedonia.

In Study 3, participants with a range of depressive symptoms completed the learning task with interspersed queries targeted at different aspects of learning. Queries asking about the expected value of the most recent choice increased learning rate while queries about the value of the unchosen option decreased learning rate. Negatively valenced queries, when assessed as a group, also decreased learning rate. In terms of outcome sensitivity, queries about the probability of outcomes, the value of high value options, and the least ever received when choosing an option. Both positively and negatively valenced queries when assessed as groups increased outcome sensitivity as well. Depression severity was not related to changes in learning parameters and participants' ratings and performance indicated the task was well tolerated.

General discussion

Together, these studies add important information to understanding, at a mechanistic level, alterations of reward and punishment learning in depression and how to treat these alterations. First of all, the specificity of findings from Studies 1 and 2 to symptom clusters, rather than an overall diagnosis of depression, provide insight into the underlying structure of depression. Depression has long been recognized as a heterogeneous category encompassing subtypes or dimensions of disease (e.g. Abramson, Metalsky, & Alloy, 1989; Fava et al., 1997). The current results suggest that, in the realm of learning dysfunctions and their treatment with CBT, continuous measures of negative affect and anhedonia, core symptoms of depression, map on to unique learning differences

in depression. These results are consonant with other recent findings showing specificity of neural and behavioral responses to reward and punishment with these symptom measures (Harlé et al., 2017; Luking et al., 2015; Young et al., 2016). This body of work, in addition to other studies suggesting problems with depression questionnaire sum scores as research constructs (Fried & Nesse, 2015), means that a focus on well-validated measures of symptom clusters, rather than on overall MDD diagnosis or depression severity, may be a more productive avenue for future research.

These studies also point to the utility of reinforcement learning models in understanding how learning differences characterize depression and treatment mechanisms. In these studies, reinforcement learning models precisely illustrated alterations and remediations in learning. In turn, this precision allowed for mappings between behavioral and neural measures and showed which parts of learning were disrupted (and subsequently remediated) and which were intact. For example, Study 1 found specific aspects of reward learning that were related to anhedonia, with the corollary finding that other aspects are unaffected by anhedonia. Therefore, rather than merely showing that anhedonia was related to altered reward learning, the modeling results showed that some aspects of learning were intact (e.g. representation of outcomes and expected values) and which were disrupted (updating expectations of reward based on prediction error). In turn, Studies 2 and 3 showed that CBT selectively remediated these disruptions, suggesting that CBT effectively targets the precise learning difficulties people high in anhedonia experience, and that similar changes in learning could be effected through certain kinds of queries (e.g. increasing learning rate by asking about the expected value of recent choices). This exact connection between alterations and treatment was

greatly facilitated by computational models that contain certain parameters or parts of the model for each possible process affected.

Lastly, these studies show the potential of using computational models to understand psychological treatments. Study 2 found that traditional CBT treatment which addresses maladaptive learning patterns appears to do so in a way that remediates pre-treatment learning differences, while Study 3 found that these learning patterns can be systematically altered in a retraining paradigm that targets aspects of reinforcement learning. Current psychological treatments for disorders like depression are moderately effective but their mechanisms of action are incompletely understood (Emmelkamp et al., 2014; Kazdin, 2009). Computational models, through their ability to provide specific, formalized findings and to connect levels of analysis, can add to the understanding of how current treatments work and assist in developing new, model-informed treatments. Further work using computational models to connect basic research on learning and affective processing to clinical applications will lead to refinements of existing treatments and development of novel treatments to ameliorate disorders like depression.

Limitations and future directions

The present set of studies assessed how depression and treatment are related to a certain computational model of learning, namely model-free reinforcement learning. Model-free reinforcement learning is a powerful approach to understand how humans learn from positive and negative feedback, but does not encompass all aspects of learning. This type of model may neglect other processes involved in dysfunction and treatment besides learning, such as memory processes (A. G. E. Collins et al., 2014; Shohamy & Daw,

2015) and attention (Niv et al., 2015). In addition, other, more complex forms of learning may play a role in depression, which the present tasks and model would have difficulty capturing. Other work suggests problems with biased prior expectations (Huys, Daw, & Dayan, 2015), biased pruning of decision trees (Huys et al., 2012), or overgeneralized learning of negative states or schemas (Disner et al., 2011; Gershman, Norman, & Niv, 2015) in depression; these lines of work could be incorporated into the altered model free learning found here to extend and clarify the present findings in light of these other forms of learning.

The current work found that symptom improvement in CBT was related to changes in learning parameters immediately post-treatment. This work suggests that CBT is causally related to learning changes, a result supported by results from Study 3 that learning parameters could be manipulated based on targeted queries similar to those employed in CBT techniques. However, the number of participants not receiving CBT was too small at follow up to permit a comparison of the group receiving CBT to a non-treatment condition. In addition, other effective treatments, including antidepressants such as selective serotonin reuptake inhibitors and other psychological approaches such as interpersonal therapy, which is efficacious but is hypothesized to work through different mechanisms (Markowitz & Weissman, 2004), were not investigated. Similarly, both Studies 2 and 3 only investigated changes immediately post-intervention and did not assess long term changes in learning or symptoms of depression. As a result, whether the learning changes seen immediately post-treatment persist and continue to be related to depressive symptoms is unclear. Therefore, longer-term studies involving waitlist or other treatment groups are needed to determine whether learning changes are specific to learning-based

treatments like CBT or learning retraining or if learning changes are common to improvements in symptoms regardless of treatment. Further work could also investigate the relationship between learning changes and development of depressive symptoms as well as study the evolution of symptoms and learning differences with more frequent measurements. Study 3 found that it was feasible to administer learning-related measurements and interventions outside of laboratory settings, and future studies could incorporate these online approaches during longer-term treatments to have finer-grained temporal sampling of changes during treatment.

Conclusion

In summary, the present work shows the utility of using reinforcement learning to understand and treat psychopathologies such as depression. Symptom clusters of depression, anhedonia and negative affect, were related to specific learning disruptions on behavioral and neural levels. Symptom reduction after treatment with cognitive behavioral therapy was related to remediation in learning parameters and these learning parameters could be directly modified via queries in a learning retraining paradigm. These findings add important information to our understanding of depression and lay the groundwork for more effective treatments of this and related disorders.

References

Abramson, L. Y., Metalsky, F. I., & Alloy, L. B. (1989). Hopelessness depression: A theory based subtype of depression. *Psychological Review*, *96*(2), 358–372.

Asnaani, A., McLean, C. P., & Foa, E. B. (2016). Updating Watson & Marks (1971): How has our understanding of the mechanisms of extinction learning evolved and where is our field going next? *Behavior Therapy*, (1971). http://doi.org/10.1016/j.beth.2016.02.003

Bellman, R. (1952). On the theory of dynamic programming. *Proceedings of the National Academy of Sciences*, *38*, 716–719.

Berkman, E. T., Hutcherson, C. A., Livingston, J. L., Kahn, L. E., & Inzlicht, M. (2017). Self-control as value-based choice. *Current Directions in Psychological Science*, 1–16.

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., … Holmes, E. A. (2015). Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: A randomized controlled trial. *Clinical Psychological Science*, *3*(1), 91–111. http://doi.org/10.1177/2167702614560746

Bush, R. R., & Mosteller, F. (1953). A stochastic model with applications to learning. *The Annals of Mathematical Statistics*, 559–585.

Chase, H. W., Nusslock, R., Almeida, J. R., Forbes, E. E., Labarbara, E. J., & Phillips, M. L. (2013). Dissociable patterns of abnormal frontal cortical activation during anticipation of an uncertain reward or loss in bipolar versus major depression. *Bipolar Disorders*, (2), 1–16. http://doi.org/10.1111/bdi.12132

Collier, A., & Siegle, G. J. (2015). Individual differences in response to prediction bias

training. *Clinical Psychological Science*, *3*(1), 79–90. http://doi.org/10.1177/2167702614560747

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

Collins, K. A., Westra, H. A., Dozois, D. J. A., & Burns, D. D. (2004). Gaps in accessing treatment for anxiety and depression: Challenges for the delivery of care. *Clinical Psychology Review*, *24*(5), 583–616. http://doi.org/10.1016/j.cpr.2004.06.001

Cooper, J. A., Gorlick, M. A., Denny, T., Worthy, D. A., Beevers, C. G., & Maddox, W. T. (2014). Training attention improves decision making in individuals with elevated self-reported depressive symptoms. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 729–41. http://doi.org/10.3758/s13415-013-0220-4

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*. http://doi.org/10.1192/bjp.bp.114.146761

Davidson, R. J., Irwin, W., Anderle, M. J., & Kalin, N. H. (2003). The neural substrates of affective processing in depressed patients treated with venlafaxine. *The American Journal of Psychiatry*, *160*(1), 64–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12505803

Delaveau, P., Jabourian, M., Lemogne, C., Allaïli, N., Choucha, W., Girault, N., … Fossati, P. (2015). Antidepressant short-term and long-term brain effects during self-referential processing in major depression. *Psychiatry Research - Neuroimaging*, *247*, 17–24.

http://doi.org/10.1016/j.pscychresns.2015.11.007

Diener, C., Kuehner, C., Brusniak, W., Ubl, B., Wessa, M., & Flor, H. (2012). A meta-analysis of neurofunctional imaging studies of emotion and cognition in major depression. *NeuroImage*, *61*(3), 677–685. http://doi.org/10.1016/j.neuroimage.2012.04.005

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *Journal of Neuroscience*, *31*(16), 6188–6198. http://doi.org/10.1523/JNEUROSCI.6486-10.2011

Downar, J., Geraci, J., Salomons, T. V, Dunlop, K., Wheeler, S., McAndrews, M. P., … Giacobbe, P. (2013). Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. *Biological Psychiatry*. http://doi.org/10.1016/j.biopsych.2013.10.026

Emmelkamp, P. M. G., David, D., Beckers, T., Muris, P., Cuijpers, P., Lutz, W., … Vervliet, B. (2014). Advancing psychotherapy and evidence-based psychological interventions. *International Journal of Methods in Psychiatric Research*, *23*(1), 58–91. http://doi.org/10.1002/mpr

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, *16*, 693–700. http://doi.org/10.1038/nrn4044

Fava, M., Uebelacker, L. A., Alpert, J. E., Nierenberg, A. A., Pava, J. A., & Rosenbaum, J. F. (1997). Major depressive subtypes and treatment response. *Biological Psychiatry*, *42*(7), 568–576. http://doi.org/10.1016/S0006-3223(96)00440-4

Foxx, R. M. (2008). Applied behavior analysis treatment of autism: the state of the art. *Child and Adolescent Psychiatric Clinics of North America*, *17*(4), 821–34, ix.

http://doi.org/10.1016/j.chc.2008.06.007

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16311–6. http://doi.org/10.1073/pnas.0706111104

Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), 1–11. http://doi.org/10.1186/s12916-015-0325-4

Gershman, S. J., Norman, K. A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, *5*, 43–50. http://doi.org/10.1016/j.cobeha.2015.07.007

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*(6), 940–58. http://doi.org/10.1037/a0024355

Henriques, J. B., & Davidson, R. J. (2000). Decreased responsiveness to reward in depression. *Cognition & Emotion*, *14*(5), 711–724. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/026999300501176 84

Hollon, S. D., Stewart, M. O., & Strunk, D. R. (2006). Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. *Annual Review of Psychology*, *57*, 285–315. http://doi.org/10.1146/annurev.psych.57.102904.190044

Huys, Q. J. M., Daw, N. D., & Dayan, P. (2015). Depression: A decision-theoretic analysis. *Annual Review of Neuroscience*, (January), 1–23. http://doi.org/10.1146/annurev-neuro-071714-033928

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3). http://doi.org/10.1371/journal.pcbi.1002410

Huys, Q. J. M., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood and Anxiety Disorders*, *20*, 1–29. Retrieved from http://www.biolmoodanxietydisord.com/content/3/1/12/abstract

Jacobson, N. S., Martell, C. R., & Dimidjian, S. (2006). Behavioral activation treatment for depression: Returning to contextual roots. *Clinical Psychology: Science and Practice*, *8*(3), 255–270. http://doi.org/10.1093/clipsy.8.3.255

Johnsen, T. J., & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin*, *141*(4), 747–768. http://doi.org/10.1037/bul0000015

Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research*, *19*(4–5), 418–428. http://doi.org/10.1080/10503300802448899

Keshavan, M. S., Vinogradov, S., Rumsey, J., Sherrill, J., & Wagner, A. (2014). Cognitive training in mental disorders: update and future directions. *The American Journal of Psychiatry*, *171*(May), 510–522. http://doi.org/10.1176/appi.ajp.2013.13081075

Maia, T. V, & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, *14*(2), 154–162. http://doi.org/10.1038/nn.2723

Markowitz, J. C., & Weissman, M. M. (2004). Interpersonal psychotherapy: Principles and applications. *World Psychiatry*, *3*(3), 136–139.

McFarland, B. R., & Klein, D. N. (2009). Emotional reactivity in depression: Diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance. *Depression and Anxiety*, *26*(2), 117–22. http://doi.org/10.1002/da.20513

McGurk, S. R., Twamley, E. W., Sitzer, D. I., McHugo, G. J., & Mueser, K. T. (2007). A meta-analysis of cognitive remediation in schizophrenia. *The American Journal of Psychiatry*, *164*(12), 1791–1802. http://doi.org/10.1176/appi.ajp.2007.07060906.A

Miller, J. M., Schneck, N., Siegle, G. J., Chen, Y., Ogden, R. T., Kikuchi, T., … Parsey, R. V. (2013). fMRI response to negative words and SSRI treatment outcome in major depressive disorder: A preliminary study. *Psychiatry Research: Neuroimaging*, *214*(3), 296–305. http://doi.org/10.1016/j.pscychresns.2013.08.001

Motter, J. N., Pimontel, M. A., Rindskopf, D., Devanand, D. P., Doraiswamy, P. M., & Sneed, J. R. (2016). Computerized cognitive training and functional recovery in major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *189*, 184–191. http://doi.org/10.1016/j.jad.2015.09.022

Müller, V. I., Cieslik, E. C., Serbanescu, I., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2016). Altered brain activity in unipolar depression revisited. *JAMA Psychiatry*, 1–9. http://doi.org/10.1001/jamapsychiatry.2016.2783

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157.

http://doi.org/10.1523/JNEUROSCI.2978-14.2015

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*(2), 97–8. http://doi.org/10.1038/nn802

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biological Psychiatry*, *57*(4), 319–327. http://doi.org/10.1016/j.biopsych.2004.11.026

Rigoux, L., Stephan, K. E., Friston, K. J., Daunizeau, J., Penny, W. D., Daunizeau, J., … Friston, K. J. (2013). Bayesian Model Selection for group studies. *NeuroImage*, *46*(4), 1004–1017. http://doi.org/10.1016/j.neuroimage.2009.03.025.Bayesian

Ritchey, M., Dolcos, F., Eddington, K. M., Strauman, T. J., & Cabeza, R. (2011). Neural correlates of emotional processing in depression: Changes with cognitive behavioral therapy and predictors of treatment response. *Journal of Psychiatric Research*, *45*(5), 577–87. http://doi.org/10.1016/j.jpsychires.2010.09.007

Rush, A. J. (2007). The varied clinical presentations of major depressive disorder. *Journal of Clinical Psychiatry*, *68*(suppl 8), 4–10. http://doi.org/10.1017/S000748530002229X

Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hrynkiewicz, L., … Dolan, R. J. (2017). Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA Psychiatry*, 10–12. http://doi.org/10.1001/jamapsychiatry.2017.1713

Sherdell, L., Waugh, C. E., & Gotlib, I. H. (2012). Anticipatory pleasure predicts motivation for reward in major depression. *Journal of Abnormal Psychology*, *121*(1), 51–60. http://doi.org/10.1037/a0024945

Siegle, G. J., Thompson, W. K., Carter, C. S., Steinhauer, S. R., & Thase, M. E. (2007). Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biological Psychiatry*, *61*(2), 198–209. http://doi.org/10.1016/j.biopsych.2006.05.048

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished response to pleasant stimuli by depressed women. *Journal of Abnormal Psychology*, *110*(3), 488–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11502092

Smoski, M. J., Felder, J. N., Bizzell, J., Green, S. R., Ernst, M., Lynch, T. R., & Dichter, G. S. (2009). fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *Journal of Affective Disorders*, *118*(1–3), 69–78. http://doi.org/10.1016/j.jad.2009.01.034

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. http://doi.org/10.1037/a0028727

Steele, J. D., Meyer, M., & Ebmeier, K. P. (2004). Neural predictive error signal correlates with depressive illness severity in a game paradigm. *NeuroImage*, *23*(1), 269–80. http://doi.org/10.1016/j.neuroimage.2004.04.023

Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Sederberg, P. B., & Steyvers, M. (2013). A Bayesian framework for simultaneously modeling neural and behavioral data. *NeuroImage*, *72*, 193–206. http://doi.org/10.1016/j.neuroimage.2013.01.048

Watkins, E. R., Baeyens, C. B., & Read, R. (2009). Concreteness training reduces dysphoria: Proof-of-principle for repeated cognitive bias modification in depression. *Journal of*

Abnormal Psychology, *118*(1), 55–64. http://doi.org/10.1037/a0013642

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Geneva.

Yechiam, E., Busemeyer, J. R., & Stout, J. C. (2004). Using cognitive models to map relations between neuropsychological disorders and human decision making deficits using cognitive models to map relations between neuropsychological disorders and human decision making seficits. *Psychological Science*, *16*(12), 973–978.

Tables

Table 2.1: Treatment conditions. The query column shows the text presented to participants about the stimulus indicated in the stimulus queried column. Color indicates the color for that condition in figures.

Treatment	Query	Stimulus queried	Color	Valence	N
Control	Move the slider to the arrow.	--	Gray	--	199
Chance of high outcome	What is the chance of getting more than 50¢ if you choose...	Chosen on last trial	Light red	Positive	100
Chance of low outcome	What is the chance of getting less than 50¢ if you choose...	Chosen on last trial	Dark red	Negative	98
Most received	What is the most you have received for...	Chosen on last trial	Light orange	Positive	97
Least received	What is the least you have received for...	Chosen on last trial	Dark orange	Negative	102
Last chosen	On average, how much do you get when you choose...	Chosen on last trial	Light green	Positive	100
Last unchosen	On average, how much do you get when you choose...	Not chosen on last trial	Dark green	Negative	106
High value	On average, how much do you get when you choose...	Stimulus with higher probability of high outcome	Light blue	Positive	99
Low value	On average, how much do you get when you choose...	Stimulus with higher probability of low outcome	Dark blue	Negative	90
Value comparison	On average, how do the two options compare?	Both	Light purple	--	102
Value estimation	What is the average amount you expected to get when you picked...	Chosen on last trial	Medium purple	--	103

Table 2.2: Participant demographics.

Measure	Mean [SD]	Percent above clinical cutoff
Gender (# [%] female)	558 [44.4%]	--
Age	34.92 [13.85]	--
Education (# [%] college degree)	638 [50.8%]	--
Depression	7.90 [9.33]	31.9%
Anxiety	5.28 [7.21]	24.6%
Stress	9.07 [8.56]	22.1%
Positive Affect (before)	31.31 [8.82]	--
Positive Affect (after)	29.35 [10.35]	--
Negative Affect (before)	13.24 [5.18]	--
Negative Affect (after)	13.22 [5.10]	--

Depression, Anxiety, & Stress: subscales of Depression, Anxiety, and Stress Scale

Positive and Negative Affect: from pre-task (before) and post-task (after) administrati

Positive and Negative Affect Schedule

note: 43 participants did not provide demographic data

Figures

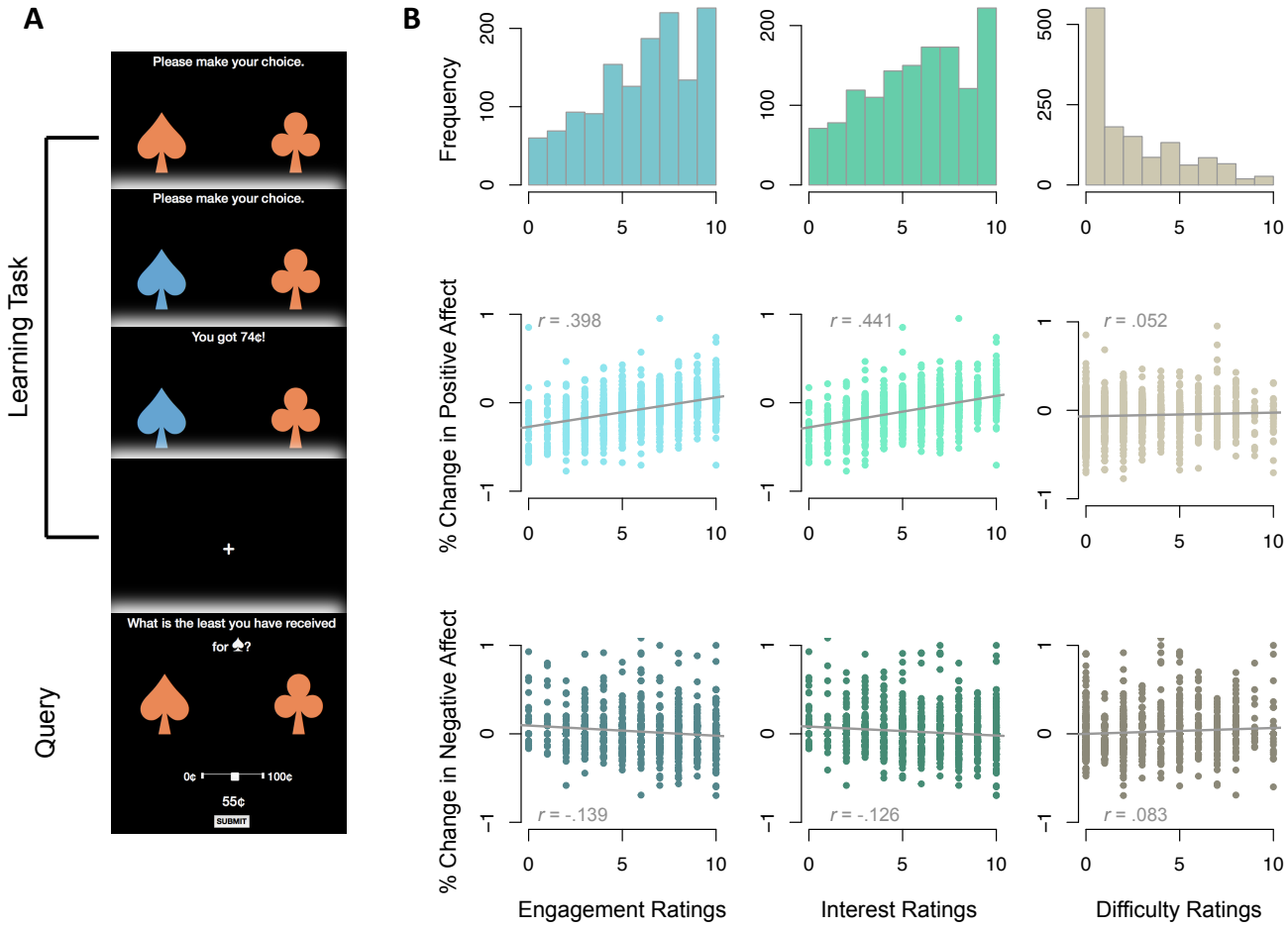

Figure 2.1: Schematic of learning task and query and related participant ratings. A) Participants completed 50 trials of a probabilistic two choice learning task where one outcome was associated with a higher probability (75%) of a larger reward (greater than 50 cents) and the other outcome was associated with a higher probability of a smaller reward (less than 50 cents). Every three trials, participants were queried on a specific aspect of the learning task and answered the query using a slider bar. Each participant was assigned to one treatment and each treatment used only one query throughout the task. Note that participants were not given feedback on their answers to queries. The query for

the control treatment asked participants to move the slider bar to a certain point without asking about any learning components. B) Participants rated their engagement (first column), interest (second column), and difficulty (third column) in the task. Overall, ratings were high on engagement and interest and low on difficulty (top row; histograms of ratings), indicating the task was acceptable to participants. Engagement and interest ratings were positively correlated with positive affect (middle row) and engagement and interest ratings were negatively correlated and difficulty ratings positively correlated with changes in negative affect (bottom row), confirming the validity of the ratings. For plots on bottom two rows, dots indicate individual values and gray lines indicate lines of best fit.

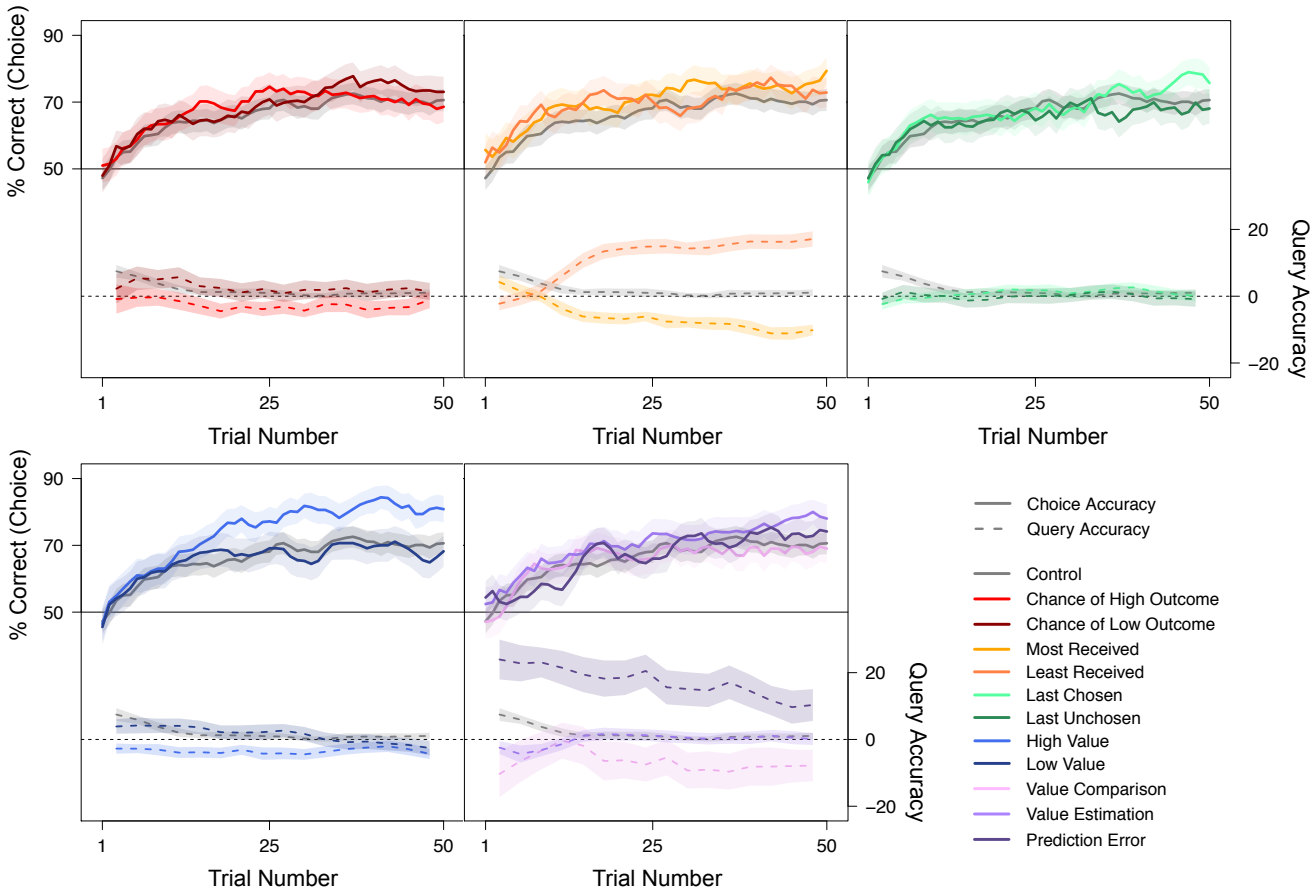

Figure 2.2: Learning curves and query accuracy over time by treatment. Solid lines (top) indicate percent correct choices on the learning task, while dotted lines (bottom) indicate distance from the correct value on queries (i.e., correct = 0); positive values on query accuracy indicate answers greater than the correct value while negative values indicate answers less than the correct value. Colored lines indicate treatments and the gray line indicates the control condition; values are group means and shading indicates one standard deviation. Across treatments, participants increased in accuracy from chance (50% correct) on the first trial, showing learning. A changepoint analysis confirmed that

learning asymptoted about halfway through the learning task (changepoint = trial 22), so trials in the first half of the learning task only were used to estimate learning parameters. Answers to queries were generally accurate but showed greater distance from 0 for queries requiring maintaining values in memory (most/least received treatments) or comparing values (value comparison & prediction error treatments).

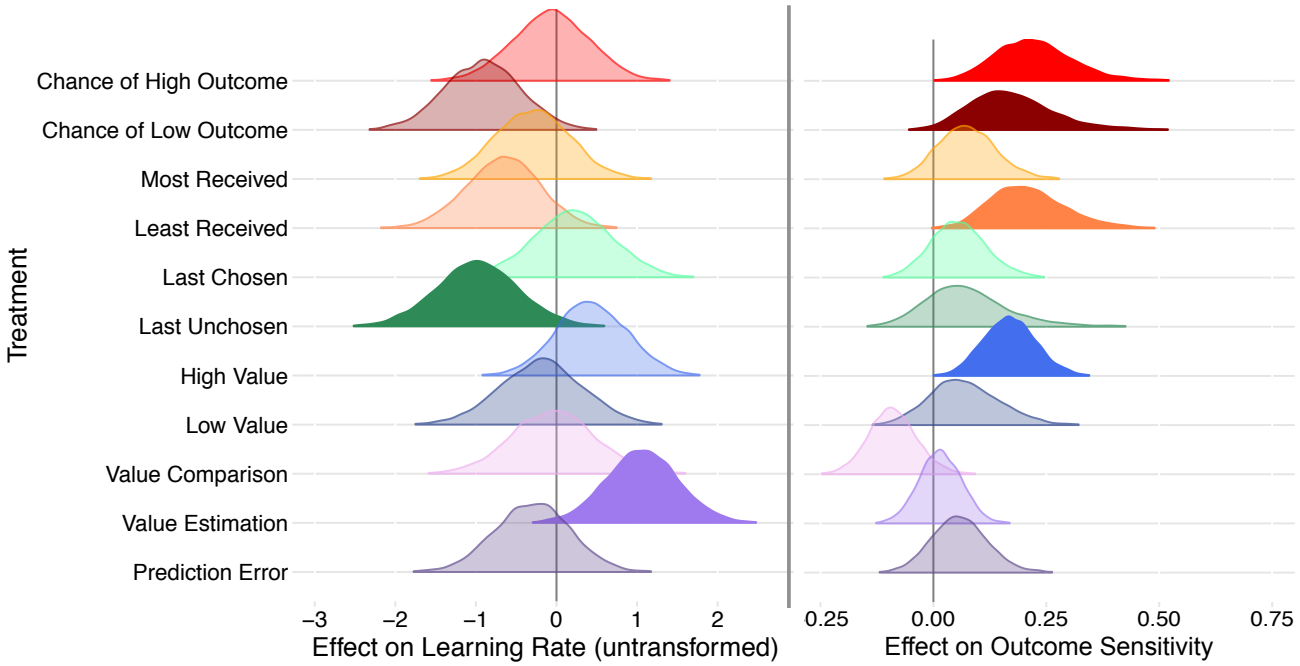

Figure 2.3: Parameter changes by treatment. Colored areas indicate posterior probability distributions of the effect of each treatment versus the control condition on each parameter. Solidly filled distributions have at least 95% of the distribution above or below 0, indicating a significant effect of that specific treatment on the parameter. Queries about the value of the unchosen option decreased learning rate, while querying estimates of expected value increased learning rate; queries about the option more likely to lead to the high value outcome, about the least ever received for an option, or about the probability of receiving either a high or low outcome all increased outcome sensitivity.

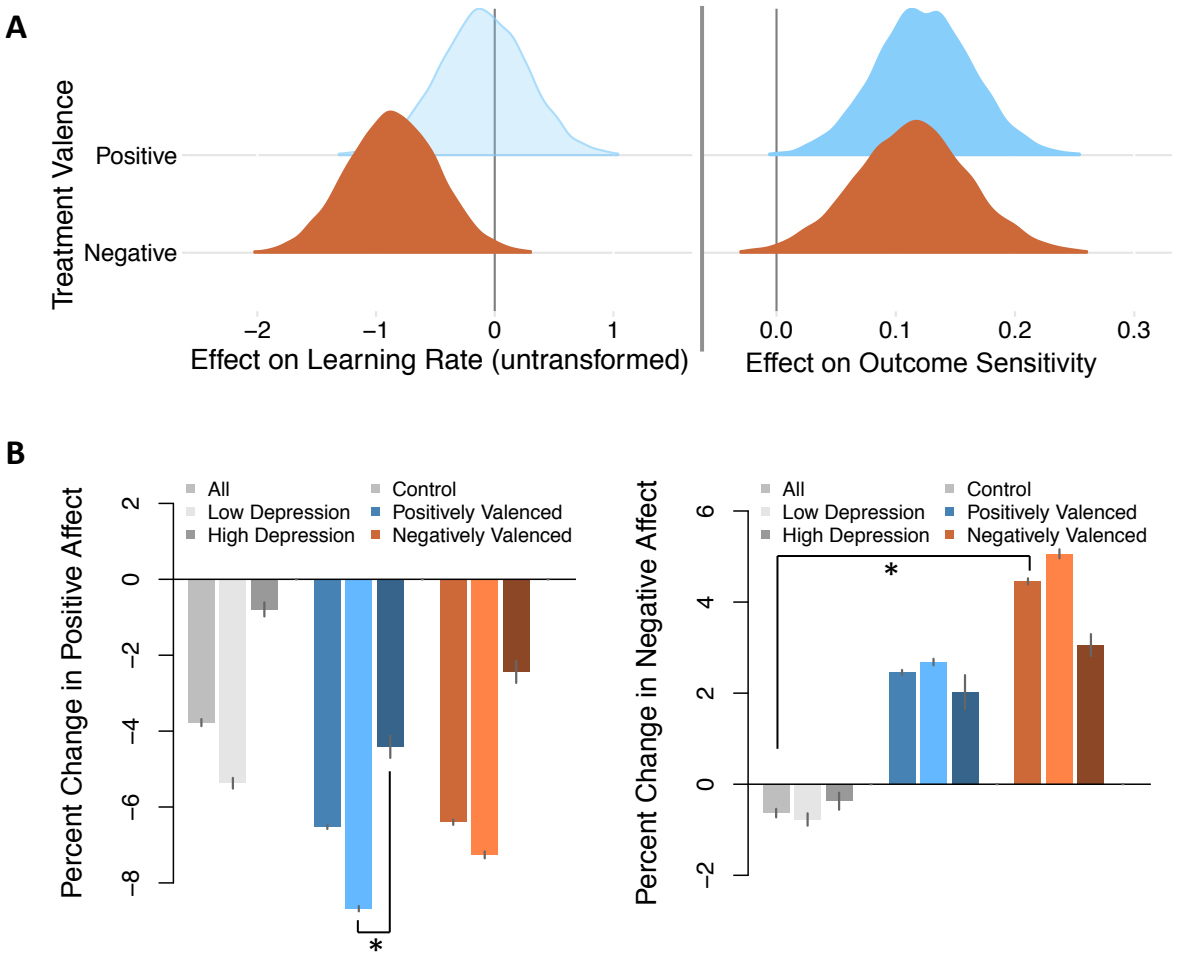

Figure 2.4. Parameter changes and changes in affect by treatment valence. A) Effects of treatment valence on learning parameters. Both positively and negatively valenced treatments increased outcome sensitivity, while negatively valenced treatments also decreased learning rate. B) Effects of treatment valence on changes in affect. Participants with depression levels above the clinical cutoff showed an improvement in positive affect after completing a positively valenced treatment, relative to participants with low levels of depression, while all participants showed a greater increase in negative affect in negatively valenced treatments relative to the control condition.

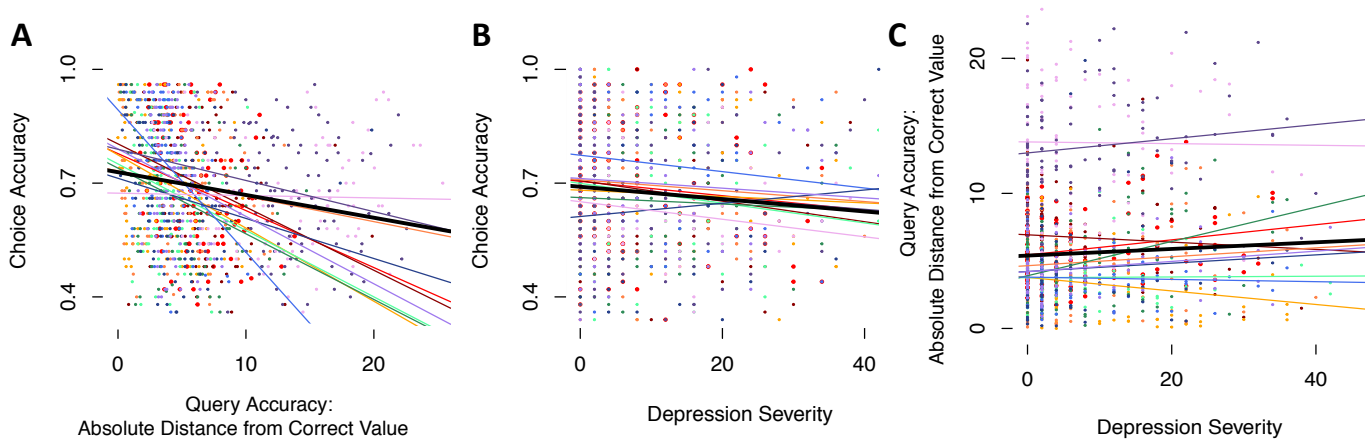

Figure 2.5: Relationship among query accuracy, choice accuracy on learning task, and depression. A) Relationship between query accuracy and choice accuracy. Dots are individual participants' mean values, colored lines are lines of best fit for individual treatments, and thick black line is line of best fit across all treatments. Higher values on query accuracy indicate a greater distance from the correct value (i.e., worse accuracy), while higher values on choice accuracy indicate better learning. Across all treatments, greater choice accuracy is related to better query accuracy (smaller distance from correct value); this relationship is present in most individual treatments as well. B) Relationship between depression severity and choice accuracy. Participants with higher levels of depression had worse performance on the learning task, although the size of the effect was small (R^2 = .006). C) Relationship between depression severity and query accuracy. Across treatments, depression severity was not related to query accuracy, but in individual treatments, higher depression was related to greater accuracy (smaller distance from correct) for the treatment querying the probability of the high value option and worse accuracy (larger distance from correct) when querying the value of the option not chosen most recently.

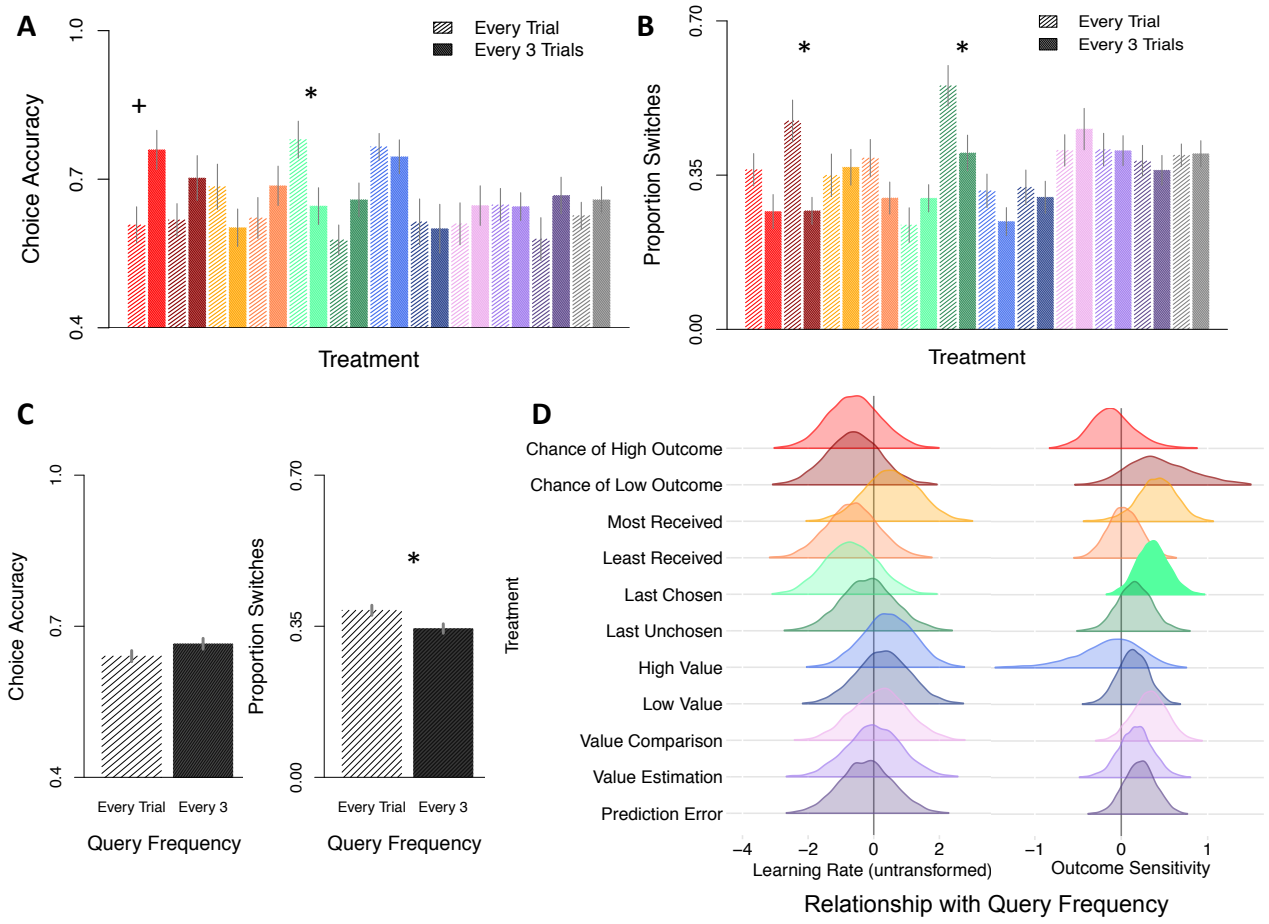

Figure 2.6: Effect of query frequency on learning. An additional ~20 participants per group completed the learning task with queries either every three trials (as in previous versions) or every trial. A) Effects of query frequency on choice accuracy on the learning task. Querying every trial resulted in worse accuracy when querying the probability of receiving a high outcome, a trend towards worse accuracy when querying the value of the option not chosen most recently, and better accuracy when querying the value of the most recently chosen option. B) Effects of query frequency on proportion of trials where participants switched choices. Querying every trial resulted in greater switches when querying the probability of receiving a low outcome and when querying about the value of the option not chosen most recently. C) Effect of query frequency on choice accuracy and

proportion switches collapsed across all treatments. Across all treatments, querying every trial significantly increased the proportion of switch trials. D) Effects of querying every trial versus every three trials, for each treatment compared to the control condition. Query frequency had little effect on learning rate but increased outcome sensitivity when querying about the last chosen option.

Discussion

Summary of studies

This set of studies aimed to add to the understanding of how computational models of reinforcement learning can inform knowledge of the mechanisms of depression and its treatment. In Study 1 (Paper 1), alterations in specific components of reward and loss learning according to a reinforcement learning model corresponded with symptoms of anhedonia and negative affect in depression on behavioral and neural levels. In Study 2, (Paper 1) these altered behavioral components changed along with improvements in symptoms after cognitive behavioral therapy, while pre-treatment neural indicators of reward learning predicted response to treatment. In Study 3 (Paper 2), targeted queries were successful in changing learning parameters in participants with a range of depressive symptoms, although some queries had effects that differed from what would be inferred from reinforcement learning theory.

In Study 1, behavioral reinforcement learning parameters and neural correlates of reinforcement learning during reward and loss learning differed by levels of anhedonia and negative affect symptoms. In reward learning, people with depression showed increased outcome sensitivity and reduced learning rate with increased anhedonia as well as disrupted relationships between neural measures of expected value and prediction error in the ventral striatum. These relationships between anhedonia and learning measures were only present in participants with a diagnosis of depression, suggesting that anhedonia in clinically depressed participants affects learning differently than participants with

anhedonia in the absence of mood-related psychopathology. In loss learning, greater levels of negative affect in all participants was related to a more negative valuation of losses as indexed by a more negative outcome shift parameter behaviorally and reduced subgenual anterior cingulate cortex and precuneus signaling to levels of outcome value neurally. Together, these findings show that depression affects learning from both rewards and losses, but that these effects differ by valence and symptom cluster.

In Study 2, participants previously diagnosed with depression completed the learning task again after completing a course of cognitive behavioral therapy to examine correlates and predictors of symptom change with treatment. The amount of improvement in anhedonia correlated with increased learning rate and decreased outcome sensitivity, while the amount of improvement in negative affect correlated with less negative outcome shift and reduced outcome sensitivity. When assessing these changes in light of the baseline differences found in Study 1, these findings indicate that symptom change with CBT led to remediation of altered learning patters found pre-treatment. People high in anhedonia also showed normalization of altered neural relationships between striatal signals of prediction error and expected value. Behavioral learning measures and striatal prediction error-expected value relationships did not show significant changes in control participants who did not undergo CBT, suggesting that these measures are stable over time when assessed independent of symptom change. Striatal signals of prediction error and expected value also predicted improvement in anhedonia, with greater prediction error signal and lesser expected value signal pre-treatment related to greater improvement in anhedonia with treatment. The combined results from Study 2 suggest that effective CBT is related to changes in disrupted learning parameters in both reward and loss learning and

that neural measures of reward learning may predict improvement in reward-related symptoms of anhedonia.

In Study 3, participants with a range of depressive symptoms completed the learning task with interspersed queries targeted at different aspects of learning. Queries asking about the expected value of the most recent choice increased learning rate while queries about the value of the unchosen option decreased learning rate. Negatively valenced queries, when assessed as a group, also decreased learning rate. In terms of outcome sensitivity, queries about the probability of outcomes, the value of high value options, and the least ever received when choosing an option. Both positively and negatively valenced queries when assessed as groups increased outcome sensitivity as well. Depression severity was not related to changes in learning parameters and participants' ratings and performance indicated the task was well tolerated.

General discussion

Together, these studies add important information to understanding, at a mechanistic level, alterations of reward and punishment learning in depression and how to treat these alterations. First of all, the specificity of findings from Studies 1 and 2 to symptom clusters, rather than an overall diagnosis of depression, provide insight into the underlying structure of depression. Depression has long been recognized as a heterogeneous category encompassing subtypes or dimensions of disease (e.g. Abramson, Metalsky, & Alloy, 1989; Fava et al., 1997). The current results suggest that, in the realm of learning dysfunctions and their treatment with CBT, continuous measures of negative affect and anhedonia, core symptoms of depression, map on to unique learning differences

in depression. These results are consonant with other recent findings showing specificity of neural and behavioral responses to reward and punishment with these symptom measures (Harlé et al., 2017; Luking et al., 2015; Young et al., 2016). This body of work, in addition to other studies suggesting problems with depression questionnaire sum scores as research constructs (Fried & Nesse, 2015), means that a focus on well-validated measures of symptom clusters, rather than on overall MDD diagnosis or depression severity, may be a more productive avenue for future research.

These studies also point to the utility of reinforcement learning models in understanding how learning differences characterize depression and treatment mechanisms. In these studies, reinforcement learning models precisely illustrated alterations and remediations in learning. In turn, this precision allowed for mappings between behavioral and neural measures and showed which parts of learning were disrupted (and subsequently remediated) and which were intact. For example, Study 1 found specific aspects of reward learning that were related to anhedonia, with the corollary finding that other aspects are unaffected by anhedonia. Therefore, rather than merely showing that anhedonia was related to altered reward learning, the modeling results showed that some aspects of learning were intact (e.g. representation of outcomes and expected values) and which were disrupted (updating expectations of reward based on prediction error). In turn, Studies 2 and 3 showed that CBT selectively remediated these disruptions, suggesting that CBT effectively targets the precise learning difficulties people high in anhedonia experience, and that similar changes in learning could be effected through certain kinds of queries (e.g. increasing learning rate by asking about the expected value of recent choices). This exact connection between alterations and treatment was

greatly facilitated by computational models that contain certain parameters or parts of the model for each possible process affected.

Lastly, these studies show the potential of using computational models to understand psychological treatments. Study 2 found that traditional CBT treatment which addresses maladaptive learning patterns appears to do so in a way that remediates pre-treatment learning differences, while Study 3 found that these learning patterns can be systematically altered in a retraining paradigm that targets aspects of reinforcement learning. Current psychological treatments for disorders like depression are moderately effective but their mechanisms of action are incompletely understood (Emmelkamp et al., 2014; Kazdin, 2009). Computational models, through their ability to provide specific, formalized findings and to connect levels of analysis, can add to the understanding of how current treatments work and assist in developing new, model-informed treatments. Further work using computational models to connect basic research on learning and affective processing to clinical applications will lead to refinements of existing treatments and development of novel treatments to ameliorate disorders like depression.

Limitations and future directions

The present set of studies assessed how depression and treatment are related to a certain computational model of learning, namely model-free reinforcement learning. Model-free reinforcement learning is a powerful approach to understand how humans learn from positive and negative feedback, but does not encompass all aspects of learning. This type of model may neglect other processes involved in dysfunction and treatment besides learning, such as memory processes (A. G. E. Collins et al., 2014; Shohamy & Daw,

2015) and attention (Niv et al., 2015). In addition, other, more complex forms of learning may play a role in depression, which the present tasks and model would have difficulty capturing. Other work suggests problems with biased prior expectations (Huys, Daw, & Dayan, 2015), biased pruning of decision trees (Huys et al., 2012), or overgeneralized learning of negative states or schemas (Disner et al., 2011; Gershman, Norman, & Niv, 2015) in depression; these lines of work could be incorporated into the altered model free learning found here to extend and clarify the present findings in light of these other forms of learning.

The current work found that symptom improvement in CBT was related to changes in learning parameters immediately post-treatment. This work suggests that CBT is causally related to learning changes, a result supported by results from Study 3 that learning parameters could be manipulated based on targeted queries similar to those employed in CBT techniques. However, the number of participants not receiving CBT was too small at follow up to permit a comparison of the group receiving CBT to a non-treatment condition. In addition, other effective treatments, including antidepressants such as selective serotonin reuptake inhibitors and other psychological approaches such as interpersonal therapy, which is efficacious but is hypothesized to work through different mechanisms (Markowitz & Weissman, 2004), were not investigated. Similarly, both Studies 2 and 3 only investigated changes immediately post-intervention and did not assess long term changes in learning or symptoms of depression. As a result, whether the learning changes seen immediately post-treatment persist and continue to be related to depressive symptoms is unclear. Therefore, longer-term studies involving waitlist or other treatment groups are needed to determine whether learning changes are specific to learning-based

treatments like CBT or learning retraining or if learning changes are common to improvements in symptoms regardless of treatment. Further work could also investigate the relationship between learning changes and development of depressive symptoms as well as study the evolution of symptoms and learning differences with more frequent measurements. Study 3 found that it was feasible to administer learning-related measurements and interventions outside of laboratory settings, and future studies could incorporate these online approaches during longer-term treatments to have finer-grained temporal sampling of changes during treatment.

Conclusion

In summary, the present work shows the utility of using reinforcement learning to understand and treat psychopathologies such as depression. Symptom clusters of depression, anhedonia and negative affect, were related to specific learning disruptions on behavioral and neural levels. Symptom reduction after treatment with cognitive behavioral therapy was related to remediation in learning parameters and these learning parameters could be directly modified via queries in a learning retraining paradigm. These findings add important information to our understanding of depression and lay the groundwork for more effective treatments of this and related disorders.

References

Abramson, L. Y., Metalsky, F. I., & Alloy, L. B. (1989). Hopelessness depression: A theory based subtype of depression. *Psychological Review*, *96*(2), 358–372.

Asnaani, A., McLean, C. P., & Foa, E. B. (2016). Updating Watson & Marks (1971): How has our understanding of the mechanisms of extinction learning evolved and where is our field going next? *Behavior Therapy*, (1971). http://doi.org/10.1016/j.beth.2016.02.003

Bellman, R. (1952). On the theory of dynamic programming. *Proceedings of the National Academy of Sciences*, *38*, 716–719.

Berkman, E. T., Hutcherson, C. A., Livingston, J. L., Kahn, L. E., & Inzlicht, M. (2017). Self-control as value-based choice. *Current Directions in Psychological Science*, 1–16.

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., … Holmes, E. A. (2015). Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: A randomized controlled trial. *Clinical Psychological Science*, *3*(1), 91–111. http://doi.org/10.1177/2167702614560746

Bush, R. R., & Mosteller, F. (1953). A stochastic model with applications to learning. *The Annals of Mathematical Statistics*, 559–585.

Chase, H. W., Nusslock, R., Almeida, J. R., Forbes, E. E., Labarbara, E. J., & Phillips, M. L. (2013). Dissociable patterns of abnormal frontal cortical activation during anticipation of an uncertain reward or loss in bipolar versus major depression. *Bipolar Disorders*, (2), 1–16. http://doi.org/10.1111/bdi.12132

Collier, A., & Siegle, G. J. (2015). Individual differences in response to prediction bias

training. *Clinical Psychological Science*, *3*(1), 79–90. http://doi.org/10.1177/2167702614560747

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

Collins, K. A., Westra, H. A., Dozois, D. J. A., & Burns, D. D. (2004). Gaps in accessing treatment for anxiety and depression: Challenges for the delivery of care. *Clinical Psychology Review*, *24*(5), 583–616. http://doi.org/10.1016/j.cpr.2004.06.001

Cooper, J. A., Gorlick, M. A., Denny, T., Worthy, D. A., Beevers, C. G., & Maddox, W. T. (2014). Training attention improves decision making in individuals with elevated self-reported depressive symptoms. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 729–41. http://doi.org/10.3758/s13415-013-0220-4

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*. http://doi.org/10.1192/bjp.bp.114.146761

Davidson, R. J., Irwin, W., Anderle, M. J., & Kalin, N. H. (2003). The neural substrates of affective processing in depressed patients treated with venlafaxine. *The American Journal of Psychiatry*, *160*(1), 64–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12505803

Delaveau, P., Jabourian, M., Lemogne, C., Allaïli, N., Choucha, W., Girault, N., … Fossati, P. (2015). Antidepressant short-term and long-term brain effects during self-referential processing in major depression. *Psychiatry Research - Neuroimaging*, *247*, 17–24.

http://doi.org/10.1016/j.pscychresns.2015.11.007

Diener, C., Kuehner, C., Brusniak, W., Ubl, B., Wessa, M., & Flor, H. (2012). A meta-analysis of neurofunctional imaging studies of emotion and cognition in major depression. *NeuroImage*, *61*(3), 677–685. http://doi.org/10.1016/j.neuroimage.2012.04.005

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *Journal of Neuroscience*, *31*(16), 6188–6198. http://doi.org/10.1523/JNEUROSCI.6486-10.2011

Downar, J., Geraci, J., Salomons, T. V, Dunlop, K., Wheeler, S., McAndrews, M. P., … Giacobbe, P. (2013). Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. *Biological Psychiatry*. http://doi.org/10.1016/j.biopsych.2013.10.026

Emmelkamp, P. M. G., David, D., Beckers, T., Muris, P., Cuijpers, P., Lutz, W., … Vervliet, B. (2014). Advancing psychotherapy and evidence-based psychological interventions. *International Journal of Methods in Psychiatric Research*, *23*(1), 58–91. http://doi.org/10.1002/mpr

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, *16*, 693–700. http://doi.org/10.1038/nrn4044

Fava, M., Uebelacker, L. A., Alpert, J. E., Nierenberg, A. A., Pava, J. A., & Rosenbaum, J. F. (1997). Major depressive subtypes and treatment response. *Biological Psychiatry*, *42*(7), 568–576. http://doi.org/10.1016/S0006-3223(96)00440-4

Foxx, R. M. (2008). Applied behavior analysis treatment of autism: the state of the art. *Child and Adolescent Psychiatric Clinics of North America*, *17*(4), 821–34, ix.

http://doi.org/10.1016/j.chc.2008.06.007

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16311–6. http://doi.org/10.1073/pnas.0706111104

Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), 1–11. http://doi.org/10.1186/s12916-015-0325-4

Gershman, S. J., Norman, K. A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, *5*, 43–50. http://doi.org/10.1016/j.cobeha.2015.07.007

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*(6), 940–58. http://doi.org/10.1037/a0024355

Henriques, J. B., & Davidson, R. J. (2000). Decreased responsiveness to reward in depression. *Cognition & Emotion*, *14*(5), 711–724. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/026999300501176 84

Hollon, S. D., Stewart, M. O., & Strunk, D. R. (2006). Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. *Annual Review of Psychology*, *57*, 285–315. http://doi.org/10.1146/annurev.psych.57.102904.190044

Huys, Q. J. M., Daw, N. D., & Dayan, P. (2015). Depression: A decision-theoretic analysis. *Annual Review of Neuroscience*, (January), 1–23. http://doi.org/10.1146/annurev-neuro-071714-033928

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3). http://doi.org/10.1371/journal.pcbi.1002410

Huys, Q. J. M., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood and Anxiety Disorders*, *20*, 1–29. Retrieved from http://www.biolmoodanxietydisord.com/content/3/1/12/abstract

Jacobson, N. S., Martell, C. R., & Dimidjian, S. (2006). Behavioral activation treatment for depression: Returning to contextual roots. *Clinical Psychology: Science and Practice*, *8*(3), 255–270. http://doi.org/10.1093/clipsy.8.3.255

Johnsen, T. J., & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin*, *141*(4), 747–768. http://doi.org/10.1037/bul0000015

Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research*, *19*(4–5), 418–428. http://doi.org/10.1080/10503300802448899

Keshavan, M. S., Vinogradov, S., Rumsey, J., Sherrill, J., & Wagner, A. (2014). Cognitive training in mental disorders: update and future directions. *The American Journal of Psychiatry*, *171*(May), 510–522. http://doi.org/10.1176/appi.ajp.2013.13081075

Maia, T. V, & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, *14*(2), 154–162. http://doi.org/10.1038/nn.2723

Markowitz, J. C., & Weissman, M. M. (2004). Interpersonal psychotherapy: Principles and applications. *World Psychiatry*, *3*(3), 136–139.

McFarland, B. R., & Klein, D. N. (2009). Emotional reactivity in depression: Diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance. *Depression and Anxiety*, *26*(2), 117–22. http://doi.org/10.1002/da.20513

McGurk, S. R., Twamley, E. W., Sitzer, D. I., McHugo, G. J., & Mueser, K. T. (2007). A meta-analysis of cognitive remediation in schizophrenia. *The American Journal of Psychiatry*, *164*(12), 1791–1802. http://doi.org/10.1176/appi.ajp.2007.07060906.A

Miller, J. M., Schneck, N., Siegle, G. J., Chen, Y., Ogden, R. T., Kikuchi, T., … Parsey, R. V. (2013). fMRI response to negative words and SSRI treatment outcome in major depressive disorder: A preliminary study. *Psychiatry Research: Neuroimaging*, *214*(3), 296–305. http://doi.org/10.1016/j.pscychresns.2013.08.001

Motter, J. N., Pimontel, M. A., Rindskopf, D., Devanand, D. P., Doraiswamy, P. M., & Sneed, J. R. (2016). Computerized cognitive training and functional recovery in major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *189*, 184–191. http://doi.org/10.1016/j.jad.2015.09.022

Müller, V. I., Cieslik, E. C., Serbanescu, I., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2016). Altered brain activity in unipolar depression revisited. *JAMA Psychiatry*, 1–9. http://doi.org/10.1001/jamapsychiatry.2016.2783

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157.

http://doi.org/10.1523/JNEUROSCI.2978-14.2015

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*(2), 97–8. http://doi.org/10.1038/nn802

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biological Psychiatry*, *57*(4), 319–327. http://doi.org/10.1016/j.biopsych.2004.11.026

Rigoux, L., Stephan, K. E., Friston, K. J., Daunizeau, J., Penny, W. D., Daunizeau, J., … Friston, K. J. (2013). Bayesian Model Selection for group studies. *NeuroImage*, *46*(4), 1004–1017. http://doi.org/10.1016/j.neuroimage.2009.03.025.Bayesian

Ritchey, M., Dolcos, F., Eddington, K. M., Strauman, T. J., & Cabeza, R. (2011). Neural correlates of emotional processing in depression: Changes with cognitive behavioral therapy and predictors of treatment response. *Journal of Psychiatric Research*, *45*(5), 577–87. http://doi.org/10.1016/j.jpsychires.2010.09.007

Rush, A. J. (2007). The varied clinical presentations of major depressive disorder. *Journal of Clinical Psychiatry*, *68*(suppl 8), 4–10. http://doi.org/10.1017/S000748530002229X

Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hrynkiewicz, L., … Dolan, R. J. (2017). Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA Psychiatry*, 10–12. http://doi.org/10.1001/jamapsychiatry.2017.1713

Sherdell, L., Waugh, C. E., & Gotlib, I. H. (2012). Anticipatory pleasure predicts motivation for reward in major depression. *Journal of Abnormal Psychology*, *121*(1), 51–60. http://doi.org/10.1037/a0024945

Siegle, G. J., Thompson, W. K., Carter, C. S., Steinhauer, S. R., & Thase, M. E. (2007). Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biological Psychiatry*, *61*(2), 198–209. http://doi.org/10.1016/j.biopsych.2006.05.048

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished response to pleasant stimuli by depressed women. *Journal of Abnormal Psychology*, *110*(3), 488–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11502092

Smoski, M. J., Felder, J. N., Bizzell, J., Green, S. R., Ernst, M., Lynch, T. R., & Dichter, G. S. (2009). fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *Journal of Affective Disorders*, *118*(1–3), 69–78. http://doi.org/10.1016/j.jad.2009.01.034

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. http://doi.org/10.1037/a0028727

Steele, J. D., Meyer, M., & Ebmeier, K. P. (2004). Neural predictive error signal correlates with depressive illness severity in a game paradigm. *NeuroImage*, *23*(1), 269–80. http://doi.org/10.1016/j.neuroimage.2004.04.023

Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Sederberg, P. B., & Steyvers, M. (2013). A Bayesian framework for simultaneously modeling neural and behavioral data. *NeuroImage*, *72*, 193–206. http://doi.org/10.1016/j.neuroimage.2013.01.048

Watkins, E. R., Baeyens, C. B., & Read, R. (2009). Concreteness training reduces dysphoria: Proof-of-principle for repeated cognitive bias modification in depression. *Journal of*

Abnormal Psychology, *118*(1), 55–64. http://doi.org/10.1037/a0013642

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Geneva.

Yechiam, E., Busemeyer, J. R., & Stout, J. C. (2004). Using cognitive models to map relations between neuropsychological disorders and human decision making deficits using cognitive models to map relations between neuropsychological disorders and human decision making seficits. *Psychological Science*, *16*(12), 973–978.