심리학석사 학위논문

Investigating the effect of anxiety on model-based reinforcement learning in reward and punishment conditions: a computational modeling study

보상 및 처벌조건에서 불안이 모델기반 강화학습에 미치는 영향 탐구: 계산모델링 연구

2021 년 8 월

서울대학교 대학원

심리학과 계산임상심리 전공

허 지 현

Investigating the effect of anxiety on model-based reinforcement learning in reward and punishment conditions: a computational modeling study

보상 및 처벌조건에서
불안이 모델기반 강화학습에 미치는 영향 탐구:
계산모델링 연구

지도교수 안 우 영

이 논문을 심리학석사 학위논문으로 제출함

2021 년 4 월

서울대학교 대학원

심리학과 계산임상심리 전공

허 지 현

허지현의 심리학석사 학위논문을 인준함

2021 년 6 월

위 원 장 ____________________

부위원장 ____________________

위 원 ____________________

Abstract

Decision-making and learning in anxiety is volatile. There have been inconsistent findings on the learning patterns in highly anxious people. For example, anxiety level often showed a negative correlation with adaptive learning measures during aversive tasks but not during neutral ones. Recently, psychologists investigated learning in anxiety based on two decision-making systems: model-free (MF) and model-based (MB) control. MF control reinforces habitual behavior by repeating previously rewarded actions, whereas MB control governs goal-directed behavior by internalizing the structure of a task. Previous findings examined that anxiety was not associated with MB control deficits under monetary gains. However, it remains unclear whether this null association is domain-general or subject to change depending on outcome valence. In this study, I probed whether state anxiety, trait anxiety, and worry are distinctly associated with MB control deficits under monetary losses. I recruited non-clinical adults and asked them to perform a multi-step decision-making task in both reward and punishment conditions. Using computational modeling, I estimated individual MB control measures and tested for between-condition differences in the association between MB control and three anxiety levels. Here, worry showed a significantly negative correlation with MB control in the punishment condition. Also, refocus on planning strategy moderated the relationship between worry and MB control. This study suggests that worry negatively impacts MB control in punishment when it interacts with certain emotion regulation strategies.

Keywords: Anxiety, Computational modeling, Reinforcement learning, Model-based decision-making

Student Number: 2019-21129

Contents

List of Figures

List of Tables

Introduction

Reinforcement learning (RL) has blossomed in the fields of computer science, neuroscience and psychology. In cognitive and clinical psychology, RL theories have been used to explain instrumental learning in humans and animals and connect observable learning behaviors with the underlying neural correlates (Collins & Cockburn, 2020; Sutton & Barto, 1998). Psychological research has investigated decision-making in humans and animals by utilizing the framework of such computational RL models (Niv, 2009). Most of the previous literature on learning and decision-making were based on two parallel systems of RL: model-free and model-based systems (Daw et al., 2005). Model-free (MF) RL requires no explicit model of the environment and leads to a choice based on the aggregated reward history, or cached value (Daw et al., 2011; Daw et al., 2005; Gläscher et al., 2010). It is closely tied with habitual behaviors, repetitive actions due to a strong stimulus-response association after overtraining. These habitual and inflexible behaviors tend to persist regardless of the change in the valence of outcomes. In contrast, model-based (MB) RL occurs when an agent has an internal structure of the environment and uses it to flexibly adjust its action in order to accumulate reward. MB system modulates goal-directed control that allows us to successfully regulate our behaviors and prospectively make decisions by taking the learned model of the environment into account (Daw et al., 2005; Gillan et al., 2016; Gläscher et al., 2010; Voon et al., 2017). As the MF and MB RL systems have been often mapped to the habitual and goal-directed behaviors, respectively, I will use these terms interchangeably here.

These dichotomized RL systems have been used to understand the behaviors and symptoms of each psychological disorder. Especially, MB control has been suggested as a 'dimensional construct' that underlies several psychological disorders associated with obsession and compulsion (Voon et al., 2017). A large online study examined that eating disorders, alcohol addiction, and OCD were associated with goal-directed control deficits that often lead to spontaneous and compulsive behaviors (Gillan et al., 2016). Also, the OCD patients showed higher habitual but lower goal-directed control scores compared to the healthy control during a multi-step decision-making task (Voon, Baek, et al., 2015). Recently, however, a few studies revealed that psychological constructs such as depression may also be related to the lack of ability in goal-directed decision-making under certain situations. For example, higher depression scores predicted more decreases in the use of MB control during a decision-making task, but only when the participants were experiencing social stress (Heller et al., 2018). This implies that the MF-MB control dimension can be a potential construct that underlies multiple psychological disorders characterized by not only high impulsivity but also mood instability if researchers study it with a proper experiment paradigm.

To this date, however, there lacks a research that examined these decision-making controllers in anxiety or other mood disorders in various contexts and emotional states facing different valences of outcome. The association between mood disturbance and goal-directed behaviors has not been fully probed yet. Consequently, only OCD, BED, and substance use disorders were reported as the main disorders showing MB control deficits. This might be because most of the studies that concluded anxiety groups did not show any MB control deficits compared to the healthy control used a reward-based multi-step decision-making task (e.g., Deserno et al., 2015; Gillan et al., 2020; Gillan et al., 2016; Otto et al., 2013). Also, there is a possibility that MB controller in anxiety might have seemed intact because it was estimated in an environment without any significant inner or outer aversive stimuli. Supporting this argument, several studies examined that contexts, emotional states, and outcome valence in which learning occurred significantly affected decision-making of anxious participants. For example, emotional disturbance due to acute stress and fear altered decision-making in anxious participants (Aylward et al., 2019; Browning et al., 2015; Sebold et al., 2019). It is clear that further research is needed to set up more valid experimental environments that resemble those where anxious people are likely to think and behave aberrantly.

Since only few studies manipulated the task structure or environment, it is difficult to exactly pinpoint which context, emotional state, and outcome valence is responsible for such alteration in decision-making. Indeed, most of the previous literature investigating both habitual and goal-directed behaviors exploited the limited range of task paradigms that incur positive reinforcement with probabilistic monetary reward in an emotionally neutral state. The lack of variation in a task design can be attributed to the absence of a systematic framework to consult when developing or selecting an appropriate task and environment. In the following subsections, I will summarize the definitions of MF and MB RL in psychology first and briefly review papers that have utilized MF and MB RL tasks to understand psychological disorders in human participants. Next, I will focus on anxiety and its aberrant decision-making and learning patterns that have been revealed by previous literature through neuroscientific and computational modeling approaches. After explaining how contexts, emotional states, and outcome valence interact with psychiatric symptoms including anxiety and result in distinct patterns of decision-making, I will propose my thesis idea and explain its study design to elucidate decision-making and learning in anxiety under four hypotheses.

Reinforcement Learning (RL) in psychology

Model-Free and Model-Based (MF and MB) control

RL theory originated from computer science and consists of five elements: agent, environment, a policy, reward function, and value function (Sutton & Barto, 1998). An agent is a learner and decision maker. It interacts with the environment to learn to maximize its total reward, which is one of the key features in a RL problem. A policy defines the probability of each action at a certain state. The reward function governs a reward given after each action, and it determines the negativity or positivity of the event. Lastly, the value function is similar to the reward function, but it reflects total reward in the long-term by including the estimation of the future rewards that will be given in the following states. These five elements can construct numerous RL problems by having an agent interact with the environment to learn, make a series of choices, and accumulate rewards. Even though the RL model widely used in psychology had been adopted from this RL theory, there are some differences in their definitions. In psychology, the RL model has been used to describe learning and decision-making or strategies for acting in humans and animals (Dolan & Dayan, 2013). One of the major mappings is that the MF-MB dichotomy formalized a dual system theory of behavior, habitual and goal-directed behavior, respectively (Balleine & Dickinson, 1998; Daw et al., 2005; Sutton & Barto, 1998). These two RL methods provide different predictions on future rewards, leading to distinct actions.

Habitual behavior is defined as repetitive and inflexible actions based on stimulus-response associations, relying less on the outcomes (Daw et al., 2005; Voon, Derbyshire, et al., 2015). This inflexible behavior is mostly governed by MF control that estimates the value based on cumulative history of the past rewards. In the RL theory, the temporal difference RL algorithm in which value is updated by taking reward prediction error (RPE)-the difference between prediction and actual reward-in the previous trial accounts for MF control. By utilizing this temporal difference algorithm in neuroscience, the neural mechanisms underlying habitual learning had been successfully revealed: both positive and negative RPEs are encoded in dopamine neurons and act as positive or negative reinforcers of actions (Schultz et al., 1997).

Goal-directed control is characterized by an inner representation of the environment or the transition probability. A MB agent develops an internal model or a cognitive map of the environment and flexibly changes its behavior to adjust to a volatile environment (Daw et al., 2005). Compared to habitual behavior, goal-directed behavior does not depend solely on reward history but simulates possible outcomes after a series of actions by incorporating the state transition information into a decision-making process. This prospective prediction guides more flexible decisions in the volatile environment but requires high computational costs (Loosen & Hauser, 2020).

RL and psychological disorders

Various RL tasks have been devised to study specific learning and decision-making features of each psychological disorder. Using a simple reward-learning paradigm, researchers found that patients with positive psychosis symptoms showed significantly weakened neural activation associated with RPE in substantia nigra/ventral tegmental when compared with healthy controls (Murray et al., 2008). Other latent measures of learning have also been estimated by computational modelling. One of the computational RL models, the Rescorla-Wagner RL model, formalizes how the value is updated by trial-by-trial RPE and has been widely used in computational psychiatry. It calculates trial-level prediction error values and the degree to which a subject reflects the prediction error to his or her choice, which is called learning rate. In a recent finding, unmedicated mood and anxiety patients learned faster about punishment outcomes and based their choices on more recent negative outcomes, compared to the healthy group (Aylward et al., 2019). This altered decision-making, however, was not shown in the reward domain. Such interactions between outcome valence and psychological constructs have created innumerable combinations of possible task paradigms, leading to inconsistent and incommensurable findings. This could be problematic because the confounded findings hamper our investigation to pinpoint which factor causes a difference. For example, researchers found higher trait anxiety scores were associated with less adaptability to a volatile environment in an aversive learning task, but there was no such association in the same task but with reward (Browning et al., 2015). In the study, the participants were punished with electric shock in the aversive version, whereas with money in the reward version. It is unclear whether the outcome valence (e.g. reward versus punishment) by itself or its interaction with acute emotional disturbance (e.g. fear induced by electric shock) caused the altered learning behaviors in participants with higher trait anxiety scores. To dissociate these factors and probe decision-making in psychological disorders more systematically and tailored to each disorder, it is necessary to understand the types of dimension that can lead to systematic changes in the decision-making behavior of people with psychological disorders. In the following subsections, I will specifically look into anxiety and its decision-making and learning patterns that vary under different emotional states and outcome valences.

Decision-making in anxiety

Origins of anxiety: neuroscientific evidence

Anxiety is one of the necessary emotional states or responses for humans and animals to survive (Delgado et al., 2006; LeDoux, 2012). It has been closely tied with fear as anxiety share a lot of commonalities with fear. However, many research findings have claimed that they are distinct from each other (e.g., Sylvers et al., 2011). To this date, these two emotional states have been conceptualized in many different ways. Now, most of the affective psychologists seem to reach a consensus on the definitions: fear is short-lived and elicited by a specific stimulus and short-lived, whereas anxiety, also known as a 'sustained fear,' can be incurred by not experiencing a direct stimulus and does not dissipate quickly (Davis, 1998; Hartley & Phelps, 2012; Sylvers et al., 2011). Fear has been easier for neuroscientists to identify its underlying neural mechanisms as it can be manually elicited with acute and specific sensory stimuli (Tovote et al., 2015). However, localizing specific brain areas that generate anxiety has been relatively harder due to its complexity and ambiguity as it can be evoked by not-present but anticipated and abstract threats. In fact, it was argued that what we usually call 'emotions' like anxiety, happiness, or sadness may mean nothing more than our subjective introspections on emotional states, 'feelings' (LeDoux, 2012). Due to this subjectivity involved in its definition, anxiety has been less understood than fear.

Regardless of this limitation, the neural mechanisms associated with both fear and anxiety have been increasingly investigated. Crucially, the clinical research on anxiety revealed that the process of experiencing anxiety is very similar to fear conditioning (Mineka & Zinbarg, 2006). Fear conditioning explains through which process fear can be elicited. One of the famous fear conditioning processes is Pavlovian conditioning, which associates a neutral stimulus with fear responses by connecting it with a fear-generating stimulus like electric shock or aversive sound. Using a Pavlovian fear conditioning paradigm, research to discover the neural correlates of fear and anxiety has become much easier and more scientific.

The initial focus of the research was to identify specific brain regions generating anxiety. For example, the amygdala has been widely known to be the main area responsible for fear acquisition and processing. The amygdala consists of subareas including lateral, basal, and central nuclei, and it was argued that each nucleus plays a specific role. The lateral nucleus directly receives sensory inputs from sensory cortices and thalamus, and it transmits the information to the central nucleus where the outputs are sent out to modulatory systems, periaqueductal gray or hypothalamus eliciting physiological responses. In the middle of this process, the basal nucleus connects the aforementioned two nuclei. Interestingly, it also projects to striatal regions that control behaviors or actions (LeDoux, 2007). As the past neuroimaging studies focused on the role of the amygdala, its association with anxiety has become clearer. Some studies examined that trait anxiety scores or social anxiety symptoms were correlated with hyperactivation in the amygdala (Indovina et al., 2011; Phan et al., 2006). Another study probed the impact of a short-term exposure therapy on the amygdalic activation and found that the 2-week exposure therapy on the spider phobics significantly reduced the hyperactivation in their amygdalas (Goossens et al., 2007). Apart from the amygdala, the hippocampus, insula and dorsal anterior cortex are also known to modulate fear acquisition (Fanselow, 2000; Milad et al., 2007).

However, most of the recent findings suggest that anxiety recruit more than one brain area, and such emotional states are the products of multiple brain regions that form neuronal circuits or networks. These networks can cover either distanced cortical areas or local neurons (Tovote et al., 2015). For one example of the long-range pathways, the amygdala-ventromedial prefrontal cortex (vmPFC) network has caught researchers' attention for its regulatory role on fear and anxiety. Specifically, the vmPFC engages in regulating the conditioned fear responses by directly projecting to the amygdala (Garcia et al., 1999). Using the resting-state fMRI data, neuroscientists showed higher anxiety scores were associated with weaker amygdala-vmPFC functional connectivity (M. J. Kim et al., 2011). This amygdala-vmPFC pathway also includes other regions in the temporal lobe such as the bed nucleus of the stria terminalis (BNST) and the ventral hippocampus (vHPC) (Calhoon & Tye, 2015). The BNST is known as a main output pathway of the amygdala and transmits the sensory information from the amygdala to vHPC, which eventually reaches the medial prefrontal cortex (mPFC). The reversal projection occurs as well. Especially, the reciprocal connection between the basolateral amygdala (BLA) and the vHPC modulates anxiety-related behaviors: the activation in the BLA-vPHC synapses increases anxiety-related responses (Felix-Ortiz et al., 2013).

To sum up, the initial research focus was to localize a specific brain area for fear learning and responses. Most of the early studies investigated the role of the amygdala on fear processing by itself. However, recent findings revealed that anxiety and fear conditioning is rather associated with the distributed neural networks across the brain. The major regions in these networks are the amygdala, BNST, vHPC, and mPFC. The four areas form bidirectional connections to transmit back and forth the fear-related information and mediate behavioral and physiological responses like freezing and feeling of conditioned fear. Recently, some neuroscientists raised their voices to assert that despite these concrete findings on the neural circuitry of fear and anxiety, clinical practice has not been fully benefited from them (LeDoux & Pine, 2016). To improve clinical outcomes, a novel approach to understand fear and anxiety in both neuroscience and clinical psychology might be required. Recently, computational modeling has been suggested as one of the solutions to close this gap between advanced neuroscientific evidence and clinical practices.

Computational modeling: a novel approach to probe decision-making and learning

Computational modeling is a modeling of human cognition and behavior through computational models that are mathematical equations with either known or unknown parameters and values (W.-Y. Ahn et al., 2017; Lewandowsky & Farrell, 2011). Modeling of human cognition and behavior serves various roles in psychology. For example, researchers seek to simply describe observed behavior by developing or finding the best model to fit the data. While successful description of the data provides valuable insights into our latent cognitive processes, researchers expand this approach to move beyond simple description and eventually predict and explain future behavior and cognition. Here, the terms cognition and behavior broadly represent complex and numerous decision-making, information processing, or perception in the brain, on which researchers make inferences based on their final outputs, behavior. The unobservable and latent nature of cognition has obscured our understanding on which mechanisms humans think, learn and make decisions. However, with the advent of computational models, it became significantly easier to clarify our general and individual-specific cognitive processes. This novel approach has served to fill in our gaps in the knowledge of psychiatric symptoms and disorders, which led a creation of a new area of field called computational psychiatry (W.-Y. Ahn et al., 2017; Friston et al., 2014; Huys et al., 2016; Montague et al., 2012; Paulus, 2019). In the realm of computational psychiatry, psychologists and neuroscientists have extensively worked to probe various psychological concepts by applying formal mathematical or computational models. Using computational models to describe, predict, and ultimately explain aberrant behaviors under psychiatric conditions has several advantages over other simple descriptive methods. First, it clarifies abstract psychological notions and theories that were suggested a few decades ago by providing us with their formulas. For example, one of the fundamental questions imposed on drug addicts around the 1950s was why they persist in utilizing substances that lead to regretful or unwanted psychological and physical states (Jellinek, 1949). This question had been investigated by cognitive psychologists using decision-making tasks like the Iowa Gambling task and computational models to reveal impaired learning in drug abusers (W.-Y. Ahn et al., 2016; Bechara et al., 1994; Busemeyer & Stout, 2002). Second, computational modeling enables us to quantify the latent cognitive processes and estimates individual- and group-level parameter values that can be applied to track the neural encodings of these processes (W.-Y. Ahn & Busemeyer, 2016). Despite some remaining challenges, computational psychiatry has been regarded as a medium to bring and apply neuroscientific findings to the clinical settings (Huys et al., 2016).

Decision-making and RL in anxiety

Not only psychological constructs like addiction that show apparent symptoms in a behavioral level, mood disturbance and anxiety have also been one of the main focuses in computational psychiatry. Anxiety has long been characterized with its excessive risk-avoidance decision-making, and a number of theories have suggested possible mechanisms governing such systematic risk-avoidance especially under aversive environments (e.g., Maner et al., 2007; Maner & Schmidt, 2006). Prospect theory is one of the well-known models that describe decision-making under uncertainty (Kahneman & Tversky, 1979). Using this model and computational modeling, researchers have found that anxiety patients tend to exhibit high risk-aversion (e.g., Charpentier et al., 2017). During my master's program, I have participated in one of the laboratory-based studies to investigate distortion in the perception of outcome probabilities using a modified gambling task (S. Kim et al., prep). In this study, three computational models with different probability functions were applied to the data collected from both reward and punishment domains to quantify individual perception of outcome probabilities. We examined that people with higher anxiety scores showed more distortion in probability perception, meaning higher anxiety was associated with overestimation on the likelihood of low-probable events in both reward and punishment domains.

Various learning tasks have been developed in order to study learning in anxiety. Focusing on the anxiety patients' attentional bias towards negative information, researchers have utilized simple RL tasks to reveal anxiety was associated with the increased usage of the recent outcome history or less adaptability to a changing environment during aversive learning (e.g., Aylward et al., 2019; Browning et al., 2015). Recently, there have been a few articles establishing a comprehensive MB control profile under various psychological constructs that obsession/compulsion and impulsivity are correlated with MB control deficits, whereas anxiety and depression are not (Gillan et al., 2020; Gillan et al., 2016). However, learning involves complex and multi-dimensional mechanisms in which various factors play a role and interact with each other to form similar but distinct states that can significantly alter people's choices. In the following subsections, I will list three major components by which learning can be impacted both within and between individuals and explain research gaps that I found in regards to anxiety and its goal-directed planning and learning.

What alters learning in people with psychiatric symptoms: contexts, emotional states, and outcome valence

Contexts: spaces where learning happens

Learning can occur in various contexts. Here, a context means a situation or space where the learning occurs, and it shapes the specific goal of decision-making. For example, people can be put in a context where they have to learn to maximize total monetary gain, which is the most common. There is another case when they are given rewards at the beginning and have to learn not to lose them. Also, people can be put in contexts associated with punishment as well. These contexts can be divided into 'avoidance' and 'escape.' In the avoidance condition, people learn to avoid punishment, while in the escape condition, they learn to escape from an aversive state.

Participants showed distinct learning behaviors based on which context they were in, and the within-individual or group difference in the resultant decision-making outcomes was associated with certain psychiatric symptoms. Suicidal psychiatric participants and non-suicidal psychiatric controls showed similar learning behaviors in the avoid condition, but suicidal participants exhibited significantly higher escape bias than non-suicidal controls in the escape condition (Miller et al., 2019). This escape bias represented suicidal patients' tendency to take active actions to escape from their aversive states, which well differentiated them from non-suicidal patients. Learning context also influences the transition from goal-directed to habitual behavior, which will be discussed in the outcome valence subsection in detail. It seems crucial to select a context that aligns well with a distinct behavior or a cognitive bias of the mental disorder that researchers are interested in.

Emotional states

Emotional state of the agent, the learner, is another significant factor that alters decision-making. Such emotional states include stress, frustration or craving after abstinence, anxiety, and fear. Stress induced in the middle of the two blocks of a multistep decision-making task, which estimates whether a learner exhibits the MF or MB control, decreased participants' reliance on MB system, and higher depression scores were associated with more decreases in MB control (Heller et al., 2018). The effect of stress has been also observed in a non-psychiatric group. The higher susceptibility to stress in healthy women, measured by the change in cortisol level, predicted smaller model-based weights in the low working memory capacity group (Otto et al., 2013). It is noteworthy that stressors were not the same in these two papers: social stress was induced in the former, whereas physical stress with cold water was induced in the latter. The different nature of stressors can also have an impact on learning, so the stress-learning relationship requires more investigation. Emotional states after abstinence can significantly change strategies for decision-making in substance use disorder patients. Alcohol-dependent patients who abstained from alcohol for about 24 hours showed more reliance on habitual control than goal-directed control in both behavioral and neural analyses (Sjoerds et al., 2013). Even though the associated emotional states were not reported, the longer duration of alcohol abstinence was also correlated with more weights on MB control (Voon, Derbyshire, et al., 2015). In addition, in the face of possible threats, anxious participants tend to learn more slowly. Higher trait anxiety scores exhibited significantly less flexible adjustment in their learning rates in the threat condition with probabilistic electric shock, but not in the reward condition (Browning et al., 2015). This result might be consistent with the finding of the neural mechanisms of anxiety that the activation in lateral amygdala induced unconditioned freezing behavior when facing shock-predictive cues, which could lead to behavioral inflexibility (Calhoon & Tye, 2015).

Outcome valence

Outcome valence means the positivity or negativity that an outcome bears. This shares some aspects that were already discussed in the contexts part, but here I will delve more into the distinction between gain and loss (or reward and punishment) and how it possibly affects decision-making of psychiatric patients. Before specifying each domain, it is important to understand there are a number of possible outcome types: monetary, social, auditory, and tactile outcomes. A coin gain or loss exemplifies the monetary outcome, which is the most commonly used one. Due to the loss aversion feature of human beings, monetary gain and loss affects our decisions differently. For example, the OCD patients exhibited more habitual behavior in the reward domain compared to the healthy control but showed more goal-directed behavior in the loss domain as much as the healthy control did (Voon, Baek, et al., 2015). This result has an important implication on our understanding of psychiatric disorders: it could be an overgeneralization if we conclude high-impulsivity disorders show MB control deficits in every situation. It would be crucial to understand how different motivations in the gain and loss domains bring out different behavioral results.

The interaction between gain-loss and stimulus type also affects decision-making. Overtraining for cocaine addicts in the monetary reward domain enhanced the transition from their goal-directed to habitual behavior, which was indicated by their increased insensitivity to devaluation. However, such insensitivity was not found in the tactile loss domain with electric shock (Ersche et al., 2016). One caveat on interpreting this finding is that it is hard to dissociate whether the insensitivity was due to general loss aversion or fear from predicted electric shock. To more accurately draw a conclusion, it would be recommended to use the gain and loss domain with the same stimulus type and then compare it with another stimulus type.

Study goals

Research gaps and study design

Previously established decision-making and learning patterns in people with psychiatric symptoms can be modified by their emotional responses towards different outcome valences. Anxiety, one of the most common affective states that both general populations and psychiatric patients can experience, has been characterized by its aberrant decision-making and learning behaviors examined through cognitive tasks and computational modeling. Previous literature on decision-making in anxiety patients or people with high anxiety levels have mainly focused on their excessive avoidant behavior under uncertain environments (e.g., Hartley & Phelps, 2012; Maner et al., 2007). This risk-avoidant decision-making in anxiety could be attributed to either heightened physiological responses associated with anxiety that lead to threat-avoiding choices (i.e., risk aversion) or biased appraisals that anticipate negative outcomes more often (i.e., loss aversion) (Charpentier et al., 2017). In terms of learning, individuals with higher anxiety levels showed faster update of action value after receiving an aversive outcome such as electric shock or a picture with negative facial expression. In a recent computational modelling study, people with unmedicated mood and anxiety disorders exhibited that they based their choices more on the recent history of negative outcomes rather than the longer outcome trajectories. Unlike our expectation, they did not show greater sensitivity to punishment, meaning they evaluated negative outcomes similarly to the healthy control (Aylward et al., 2019).

MB control and its association with learning in people with high state and trait anxiety have been recently investigated using a multi-step decision-making task, or the two-step task (e.g., Daw et al., 2011; Gillan et al., 2020; Gillan et al., 2016). Previous studies showed no relationship between anxiety level and MB control deficits. Not only dispositional anxiety but also induced anxiety did not impair MB learning when experimenters manipulated individual anxiety level. These findings supported that most of the components of anxiety were not associated with difficulty in the usage of MB system. However, one of the major limitations of these studies was that the task they used was only framed in the reward domain. Given the non-negligible impact of negative information on decision-making and learning in anxiety, it is likely that this null association between anxiety level and MB control might be restricted to a learning environment that does not accompany any threatening or negative outcomes.

It still remains unclear whether anxiety level is not associated with MB control deficits under monetary loss. Monetary loss incurs negative affect like anger or disgust to the degree to which individuals are sensitive to loss or punishment. Emotional responses are critical to cognitive appraisals of the situation that can lead to systematic alteration in decision-making and learning (Maner & Schmidt, 2006). It is also plausible that sequential monetary losses increase cognitive loads during aversive learning due to people's general tendency to avoid losses.

Thus, there is a possibility that MB control deficits can be observed in people with high anxiety, but only under monetary loss. To fill in this research gap, I have proposed a thesis study to recruit non-clinical adult participants with a range of anxiety levels and ask them to perform the two-step task under both monetary rewards and punishments. Here, I also comprehensively look into the components of anxiety by taking worry into consideration.

Worry is one of the the main symptoms that are considered for the diagnosis of generalized anxiety disorder (American Psychiatric Association, 2013). Historically, researchers have attempted to dissociate worry and anxiety, but most of the research findings concluded that worry is a cognitive component of anxiety (e.g., Zebb & Beck, 1998). Regarding the broad spectrum of anxiety symptoms in terms of somatic, cognitive, and behavioral aspects, it is important to dissect the construct into distinct components in order to prevent obscuring the effects of each symptom dimension (Schouten et al., 2020; Wise & Dolan, 2020). The cognitive model of pathological worry characterizes worry with three cognitive components: emotional processing, attentional control, and its verbal form (Hirsch & Mathews, 2012). Worry is associated with biased emotional processing of negative information, impaired attentional control or vulnerability towards inner or outer distractors, and its intrusive and verbal representation in mind. As MB control requires more cognitive or mental loads to use compared to MF control (Kool et al., 2016), worry might play an important role in lowering the usage of the MB system by hampering adequate distribution of cognitive resources for successful task performance. Thus, I focused on worry along with state and trait anxiety in this study to comprehensively test for associations between anxiety scales and MB control measures.

Objectives and hypotheses

The main goal of this study is to elucidate the unknown relationship between MB control and anxiety level under monetary loss. To do so, I will conduct statistical analyses based on the following four hypotheses. First, anxiety level will not be associated with MB control deficits under monetary gain. Also, people with a higher anxiety level will show a tendency to update their action values faster based on the recent history of negative outcomes. These two hypotheses are to test whether I can replicate the previous findings that reveal distinct learning and decision-making patterns in anxiety. Next, the condition-specific effect of monetary loss on MB control will be examined. I hypothesize that there will be a significantly negative association between MB control and anxiety levels only in the punishment domain. Lastly, through an exploratory analysis, I will investigate a possible factor that alters the association between MB control and anxiety. Specifically, I expect that cognitive emotion regulation strategies will moderate the relationship between anxiety level and MB control score.

This study will bring the realm of anxiety MB decision-making and learning research to an unexplored domain, punishment with sequential monetary losses. It will provide empirical evidence to support the importance of incorporating both reward and punishment conditions when probing decision-making and learning in anxiety. The findings will also suggest which component of anxiety is more associated with MB control, which will advance our understanding of each of the anxiety symptoms. From computational modeling to an exploratory investigation on possible cognitive and emotional contents underlying the association, this thesis project will ultimately help other researchers and clinicians to better grasp the latent decision-making and learning processes of anxiety patients or people with high anxiety levels.

Methods

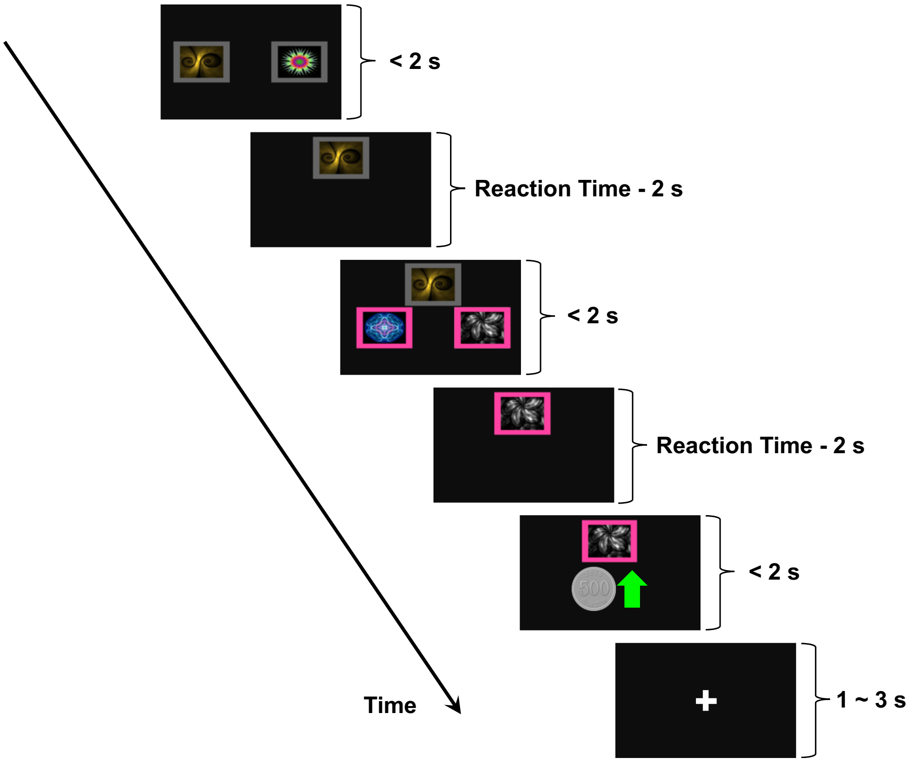
Participants

One-hundred eighty-three healthy adults were recruited through online and offline advertisements. They voluntarily agreed to fill out pre-screening survey questionnaires, which consisted of 18 questions to check participation eligibility and took about 5 minutes to complete. 15 participants failed to pass the screening because they either had histories of psychiatric disorders, were addicted to substances (e.g., nicotine), or were in a situation where involuntary participation might be possible. The pre-screening survey answers were manually checked by the experimenters, and the eligible participants were contacted individually through their preferred contact mediums (i.e., e-mail and phone message) for scheduling. Due to the unusual pandemic situation in 2020 and 2021, 39 of whom passed the screening did not respond to our request to visit the laboratory. 18 of them responded but either cancelled their visits or did not show-up. Thus, a total of 111 non-clinical adults aged from 18-35 years participated in the experiments. Each participant visited the laboratory once on the scheduled date and time. They were informed to take enough sleep during the night before the experiment day and refrain from drinking caffeinated beverage right before the experiment. All participants provided an informed consent by signing the agreement forms. They were compensated for their participation time at a rate of 10,000 ($6 in U.S dollar) per hour. Some participants received extra payment based on their performance on a cognitive task.

Cognitive task: two-step task

Figure 1

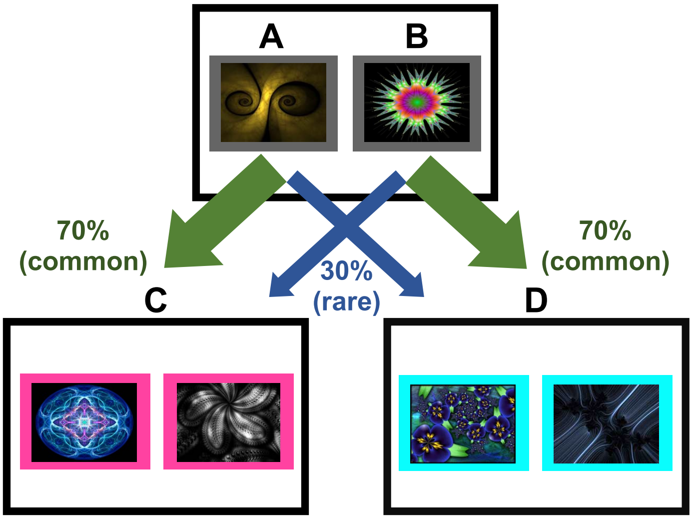
Schematic of two-step task trial in reward condition



Each participant conducted a multi-stage decision-making task. This task, also known as the two-step task, was developed to dissociate two systems of decision-making, model-free and model-based, and has been widely used in various psychology research (e.g., Daw et al., 2011; Gillan et al., 2016; Voon, Baek, et al., 2015). Each trial consists of two stages, and in each trial a participant is given with two different and

Figure 2

Task stage transition structure

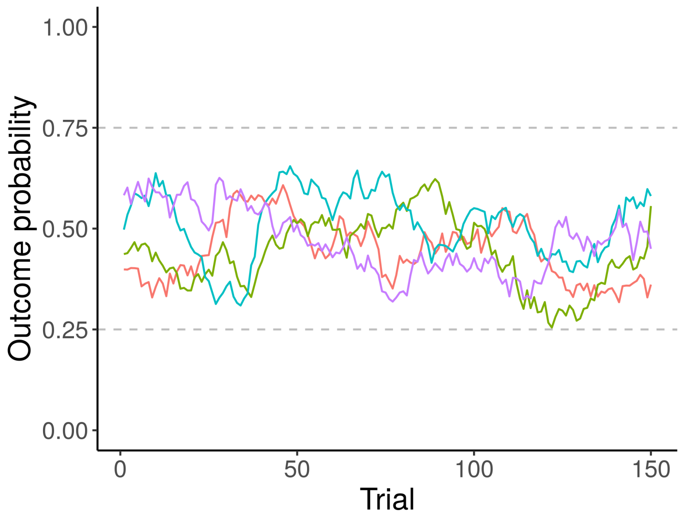


randomly selected fractal images. When the participant chooses one fractal image in the first stage within 2 seconds, the second stage, or the second pair of the fractal images, appears on the computer screen. After the participant selects one fractal image in the second stage, a final outcome appears on the screen (Figure 1).

Each of the two stimuli in the first stage is assigned with certain probabilities that determine the stimulus pair in the second stage (Figure 2). For example, if a participant chooses A in the first stage, there will be 70% chance that the stimuli with pink background (C in Figure 2) appear in the second stage and 30% change that the stimuli with skyblue background (D in Figure 2) appear. If B is selected, then vice versa. After the second stage pair is conditionally determined and

Figure 3

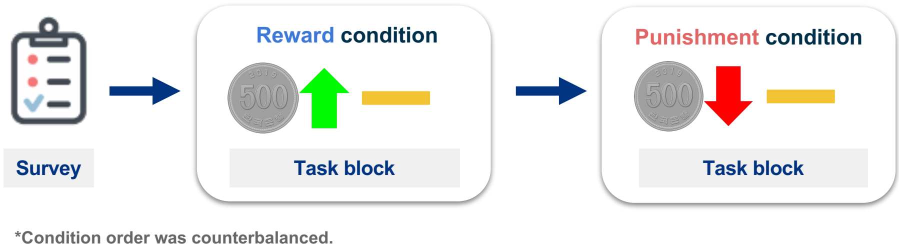
Task outcome probability



shown on the screen, the participant makes a choice to maximize the net monetary outcomes. Each of the second stage stimuli is associated with distinct probabilities of monetary outcomes, which were pre-programmed to slightly vary across the trials (Figure 3). Each of the four outcome probabilities was randomly generated with different initial points: 0.4, 0.45, 0.50, and 0.55. By using Gaussian Random Walks with the mean of 0 and standard deviation of 0.025, drifting outcome probabilities were determined with the lower and upper boundaries of 0.25 and 0.75, respectively. This variation was to make sure that the participants not only exploited the outcome histories when determining their next choices but also explored new second-stage options so that they did not stick with one first-stage or second-stage choice. To minimize any potential

Figure 4

Experimental procedure



learning biases caused by the probabilities assigned to each condition and stimulus, the same set of probabilities was used for the reward and punishment conditions and randomly assigned to each stimulus.

Most of the previous studies used the two-step task only with monetary rewards (e.g., Deserno et al., 2015; Gillan et al., 2020; Heller et al., 2018). In this study, however, all participants were asked to perform the two-step task in both reward and punishment conditions (Figure 4). In the reward condition, the participants gained either a coin reward of 500 ($0.45 in U.S. dollar) or no reward per trial. In the punishment condition, they either lost the same amount of coin or did not lose anything. The participants were informed that they would receive the net amount of money they earned during the task, and at the end of their experiment sessions, some participants whose net amounts of the task outcomes exceeded zero received extra payments. The order of the reward and punishment versions was counterbalanced across the participants and treated as a confounding variable in multiple regression analyses. All participants were required to finish 1 block of 15 practice trials per each condition to make sure they understood task instructions properly. They completed 2 blocks of 75 trials for each condition, with 1 minute break between block and 10 minutes break between condition. The entire task lasted for an hour and 10 minutes on average.

Inclusion and exclusion criteria

Screening survey

The screening survey included questions to check whether or not the subject had had neurological injuries, had hospitalized for psychiatric disorders within the last 5 years, had been medicated after being diagnosed with mental disorders within the last 1 year, had had addicted to alcohol or any other substances, was taking psychotropic drugs, had difficulty reading or listening task instructions, was not fluent in Korean, and had possibility to involuntarily participate in the experiment (e.g., a student who was taking a course from any experimenters). Participants must be able to answer no to all of the questions above to be eligible to participate in the laboratory experiment. Also, participants must be at least 18 years old and at most 35 years old. The questionnaires were distributed online through Qualtrics, and individuals who were interested voluntarily filled out the survey. Experimenters manually checked the collected answers and contacted those who were eligible to set up date and time for the laboratory experiment.

Behavioral data

Data with poor behavioral performance was not included in analyses. Poor performance was determined by the task accuracy. Specifically, I excluded the data from 17 participants who failed to demonstrate sensitivity to rewards (e.g., probability of stay after receiving rewards in common transition trials being less than 0.5) in the reward version of the two-step task. Also, 10 participants who chose the same first stage option over 95% of the entire trials in each condition were excluded as they were not deemed to adequately learn the outcome contingencies. These criteria were based on previous studies (Gillan et al., 2016; Otto et al., 2013) and reported in pre-registration. The total of 86 participants was included in the analyses.

Measures

Before conducting the two-step task in reward and punishment conditions, all participants were asked to fill out 9 survey questionnaires for about 20 minutes. The surveys included Liebowitz Social Anxiety Scale (LSAS) (Liebowitz, 1987), Penn State Worry Questionnaire (PSWQ) (Meyer et al., 1990), State-Trait Anxiety Inventory-Y (STAI-Y) (Spielberger, 1983), Beck Depression Inventory-2 (BDI-2) (Beck et al., 1996), Patient Health Questionnaire-9 (PHQ-9) (Spitzer et al., 1999), Cognitive Emotion Regulation Questionnaire (CERQ) (Garnefski & Kraaij, 2007), Yale-Brown Obsessive-Compulsive Scale-Symptom Scale (Y-BOCS-SC) (Goodman et al., 1989), and Barratt

Table 1

Summary of survey information

Construct	Survey	Number of questions
Anxiety	LSAS	24
	PSWQ	16
	STAI-Y	40
Depression	BDI-2	21
	PHQ-9	9
Emotion regulation	CERQ	36
Obsession & Compulsion	Y-BOCS-SC	58
Impulsivity	BIS-11	30

Impulsiveness Scale-11 (BIS-11) (Patton et al., 1995). The summary of the survey questionnaires is provided in Table 1. The distribution of survey scores for each questionnaire is included in Appendix A.

LSAS. LSAS measures social anxiety and avoidance with a total of 24 questions (S. Y. Park, 2003). This study used the Korean version of LSAS that was translated and validated (S. Y. Park, 2003). The internal consistency of the Korean version of LSAS (Cronbach's α) was .85-.91 (S. Y. Park, 2003).

PSWQ. PSWQ was developed to assess pathological worry level (Meyer et al., 1990). Pathological worry is defined as excessive and uncontrollable worry. The survey consists of 16 questions, and each item is scored with a 5-point Likert scale. Kim & Min translated and validated the Korean version of PSWQ (J. W. Kim & Min, 1998). The internal consistency of the Korean version of PSWQ (Cronbach's α) was .92 (J. W. Kim & Min, 1998).

STAI-Y. STAI-Y measures trait and state anxiety in non-clinical populations (Spielberger, 1983). It consists of 40 items in total. 20 questions assess state anxiety (STAI-S) and the other 20 ones assess trait anxiety (STAI-T) with a 4-point Likert scale. STAI-Y was translated into Korean and validated (Han et al., 1996). The internal consistency of the Korean version of STAI-Y (Cronbach's α) was .92 (Han et al., 1996).

BDI-2. BDI-2 is aimed to measure depressive symptoms with a total of 21 questions (Beck et al., 1996). Each question is assessed with a 4-point Likert scale of 0 to 3. This study used the Korean version of BDI-2, which was translated and validated (M.-S. Kim et al., 2007). The internal consistency of the Korean version of BDI-2 (Cronbach's α) was .80 (M.-S. Kim et al., 2007).

PHQ-9. PHQ-9 was developed for a diganosis of depressive disorders (Spitzer et al., 1999). The range of scores is 0 to 27, and the higher score represents the more severe symptoms. Each item is assessed by the frequency of each symptom with a 3-point Likert scale. This study used the Korean version of PHQ-9 that was translated and validated in 2020 (S.-J. Park et al., 2010). The internal consistency of the Korean version of PHQ-9 (Cronbach's α) was .81 (S.-J. Park et al., 2010).

CERQ. CERQ is aimed to measure individual differences in which cognitive emotion regulation strategies a person uses when experiencing negative events (Garnefski & Kraaij, 2007). It consists of 36 questions, and each item is scored at a 5-point Likert scale. CERQ was translated into Korean and validated with the internal consistency (Cronbach's α) of .84. However, the internal consistency of the 'acceptance' and 'putting into perspective' subscales were relatively low, .68 and .77, respectively (H.-n. Ahn et al., 2013).

Y-BOCS-SC. Y-BOCS-SC measures 8 categories of obsessions and 7 categories of compulsions (Goodman et al., 1989). A total of 58 questions are included in the survey, and each item is answered with either 'Yes' or 'No'. The internal consistency of Y-BOCS-SC (Cronbach's α) was .69-.91 (S., 1995). This study used the Korean version of Y-BOCS-SC that was translated and validated in 2004 (S. J. Kim et al., 2004).

BIS-11. BIS-11 measures impulsivity with a total of 30 items (Patton et al., 1995). Three constructs of impulsivity are measured by BIS-11: attentional, motor, and non-planning impulsivity. Each item is scored at a 4-point Likert scale. BIS-11 was translated and validated in Korean in 2012, and the internal consistency (Cronbach's α) was .58-.80 (Heo et al., 2012).

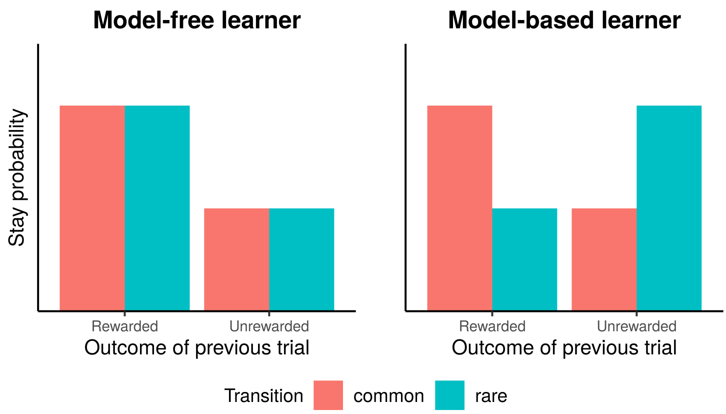
Data analyses

Behavioral analyses

Introduction to task logic. This task is based on an assumption that model-free and model-based learning strategies lead to different decision-making choice patterns. This choice pattern is mainly characterized by how the second-level choice impacts the first-level choice. Given the task structure where there are two possible transitions from the first- to the second-level (i.e., common and rare) and two possible second-level outcomes (i.e. rewarded versus not-reward and punished versus not-punished in the reward and punishment conditions, respectively), participants can incorporate either only past reward history or both reward and transition information into their choices on the next trial. If a participant exhibits more model-free learning, then the first-level choices will be made solely based on the previous reward history regardless of the transition type on the preceding trial. For example, assuming a model-free participant is rewarded after a rare transition on the previous trial, then he or she will be more likely to select the same first-level choice on the next trial. However, as the transition type is rare, it is probabilistically more plausible to view the second-level pair that includes the previously rewarded fractal stimulus. Thus, in the same scenario, a model-based learner is more likely to change its first-level choice to increase the likelihood of landing the same second-level that has the previously rewarded stimulus.

Figure 5

Stay probabilities of a purely model-free learner (left) and a model-based learner (right)



Stay probability. To visually assess the interaction between reward and transition and its impact on the first-level choice, stay probability on the current trial based on the reward and transition type on the previous trial was calculated separately for each of all possible reward and transition combinations. Also, the same four stay probabilities were calculated for both reward and punishment domains and compared. Figure 5 illustrates example patterns of stay probabilities of a model-free and a model-based learner. If a participant only uses completely model-free control, there will be no interaction effect between reward and transition on the stay probability, whereas if one is a complete model-based learner, a clear interaction effect will be inspected. Previous literature has suggested that humans exhibit both model-free and model-based strategies, which support a hybrid theory of decision-making (Daw et al., 2011; Gläscher et al., 2010). The hybrid theory specifies the human decision-making process that humans make choices based on the weighted combination of the action values that are driven by both controllers. In this study, stay probabilities were first calculated for both conditions, assessed whether they aligned with the hybrid theory, and visually inspected to check if there were any significant differences between conditions.

Mixed-effect logistic regression. To quantify model-based learning, mixed-effects logistic regression analyses were conducted using the *lme4* package in R. The logistic regression models were constructed to predict the first-level choice with the predictors including binary variables of whether reward was received and whether transition type was common or rare on the previous trial. The dependent variable, the first-level choice, was coded as 1 if a participant made the same first-level choice as the one in the previous trial, while 0 if switched. For the predictors, the rewarded (or not-punished) and common trials were coded as 1, while the non-rewarded (or punished) and rare trials as -1. Using the models, I examined the main effect of reward and interaction effect of reward and transition on the first-level choice. The main effect of reward indicated the stay propensity on the first-level was significantly governed by previous reward history or model-free learning. The interaction effect between reward and transition showed that the choice behavior was significantly controlled by model-based learning. To more accurately capture individual differences, the logistic regression models included within-subject factors as random effects, including the intercept, main effects of reward and transition, their interaction, condition, and order. Condition was coded as 1 for reward condition and as -1 for punishment condition, and order was coded as 1 if a participant was semi-randomly assigned to perform reward condition first, 0 if otherwise. The basic logistic regression model was tested while controlling for age (z-scored), order, and sex as fixed effects. The following is the model syntax specified in R:

Stay ~ Reward ∗ Transition ∗ Condition ∗ (Age + Order + Sex) + (Reward ∗ Transition ∗ Condition + Order + 1 | Subject)

I conducted additional tests including the state and trait anxiety and worry survey scores (all z-scored) in separate models:

Stay ~ Reward ∗ Transition ∗ Condition ∗ (Survey measure + Age + Order + Sex) + (Reward ∗ Transition ∗ Condition + Order + 1 | Subject)

Computational modeling

Reinforcement learning models. Three reinforcement learning models were fitted to the data from the reward and punishment conditions separately. The models are nested to each other, and the standard model, which consists of 7 parameters, was adopted from Daw et al., 2011. The model assumes that people make decisions using the weighted combination of the model-free and model-based systems. Given that there are two stages in the task, stimuli presented in each stage are updated distinctly as selection in each stage is followed by different types of outcome. A first-stage choice leads to a stimulus pair in the second stage, whereas a second-stage choice to a monetary outcome. Since monetary outcomes are presented after the second-stage choices, each of the four second-stage stimuli is updated based on the actual outcomes a participant observes across trials. This action value update in the second stage follows the model-free algorithm, SARSA(λ) temporal difference learning (Rummery & Niranjan, 1994). For each trial t, the action value V of each stimulus i ($i \in [1, 2, 3, 4]$ in the second stage) is updated following the equation below:

$$Vi_{s2}(t+1) = Vi_{s2}(t) + \alpha_2(r(t) - Vi_{s2}(t))$$

where $Vi_{s2}(t)$ is the action value of the chosen second-stage stimulus i, $r(t)$ represents actual reward (or punishment) presented on trial t. α_2 is second-stage learning rate that determines how much of the difference between the actual reward and action value is taken into account when updating the stimulus in the next trial, t+1. This reward prediction error is then added to the action value at the trial t to calculate the action value in the next trial.

The first-stage stimulus pair, however, is updated according to both model-free and model-based algorithms based on the hybrid model (Gläscher et al., 2010). The weighted sum of the action values distinctly calculated by the model-based and model-free algorithms is determined by a model-based weight parameter ω. The estimated ω close to 1 means that a participant acts like a pure model-based learner who updates his or her first-stage action value based mostly on the model-based algorithm. The model-free update of the first-stage action value follows the same SARSA algorithm like in the second stage. However, here the action value is updated by reward prediction error, the difference between the action value of the chosen second-stage stimulus and that of the first-stage stimulus. Also, the model assumes that there is a stage-skipping update that a certain portion of the difference between the outcome received in the second stage and the action value of the first-stage choice is added to the first-stage action value. This stage-skipping update is determined by both the first-stage learning rate parameter α_1 and eligibility trace parameter λ. The formal equation for the SARSA update in the first stage is the following ($i \in [1, 2]$ in the first stage):

$$Vi_{s1}^{MF}(t{+}1) = Vi_{s1}^{MF}(t) + \alpha_1(Vchosen_{s2}(t) \text{ - } Vi_{s1}^{MF}(t)) + \lambda\alpha_1(r(t) \text{ - } Vi_{s1}^{MF}(t))$$

One of the crucial characteristics of the model-based algorithm that it takes transition probabilities into account. Thus, the action value of each of the first-stage stimuli reflects not only the action values of each of the second-stage stimuli but also the transition probabilities. Thus, each of the maximum action values in each second-stage pair is weighted by the corresponding transition probability and then summed to calculate the model-based action value of the first-stage stimulus:

$$V1_{s1}^{MB} = 0.7 * \max(V1_{s2}^{MF}, V2_{s2}^{MF}) + 0.3 * \max(V3_{s2}^{MF}, V4_{s2}^{MF})$$

$$V2_{s1}^{MB} = 0.3 * \max(V1_{s2}^{MF}, V2_{s2}^{MF}) + 0.7 * \max(V3_{s2}^{MF}, V4_{s2}^{MF})$$

The weighted sum of $V1_{s1}^{MB}$ and $V1_{s1}^{MF}$ becomes the final action value of each stimulus in the first stage:

$$Vi_{s1}^{Hybrid} = \omega * Vi_{s1}^{MB} + (1\text{-}\omega) * Vi_{s1}^{MF}$$

where ω represents model-based weight parameter.

Based on the softmax function, the probability of selecting one stimulus (i.e. stimulus 1) in each stage is then calculated:

$$P_{s1}(1) = 1 \ / \ (1 + \exp(\text{-}\beta_1(V1_{s1}^{Hybrid} \text{ - } V2_{s1}^{Hybrid}) \text{ - } \pi(C(t)\text{-}C(t\text{-}1))))$$

$$P_{s2}(1) = 1 \ / \ (1 + \exp(\text{-}\beta_2(V1_{s2}^{MF} \text{ - } V2_{s2}^{MF})))$$

where perseverance parameter π reflects the tendency of repeating the first-stage choice in the previous trial (i.e., $C(t\text{-}1)$), and inverse temperature parameters β_1 and β_2 determine the randomness of choices in the first and second stage, respectively. This model has 7 parameters in total, which together form the standard model (7-par). I also modified the model by taking the eligibility parameter λ out, leading to 6 parameters (6-par). Lastly, I added the model with one learning rate and one inverse temperature parameter, resulting in 4 parameters (4-par).

Hierarchical Bayesian analysis. Model parameters were estimated with hierarchical Bayesian analysis (HBA). The parameter estimation using HBA is based on Bayesian inference, which is to re-distribute probabilities across candidate parameter values after observing the data (Kruschke, 2014). In Bayesian inference, parameter values are assigned with initial probabilities that reflect prior knowledge or beliefs. Using Bayes's rule, the priors are updated into posterior distributions that show the most probable parameter values for each individual given the data. Bayesian analysis has been suggested as an alternative of maximum likelihood estimation (MLE), a traditional parameter estimation method. One of their clear distinctions is that MLE gives point estimates that maximize the likelihood of data, whereas HBA provides full probability distributions of the parameter estimates. In comparison to MLE, HBA contains a more comprehensive information about candidate parameter values and is better at reflecting the randomness of data, estimation process, and parameter estimates.

Here, I also applied hierarchical modeling to each reinforcement learning model. Hierarchical estimation adds a hyper-constraint(s) on individual-level parameters to take both individual difference and similarity into account simultaneously. In other words, the individual parameters are estimated independently like in a non-hierarchical analysis, but they are constrained by the hyper- or group-level parameters (Lewandowsky & Farrell, 2011). This method is known to provide more reliable and stable estimates especially when the number of data or the amount of information (e.g., number of trials) for estimation is insufficient (W.-Y. Ahn et al., 2017).

HBA was conducted using the *hBayesDM* package in R (W.-Y. Ahn et al., 2017). The package uses an open-source platform for statistical modeling called Stan with a Hamiltonian Monte Carlo sampler (Carpenter et al., 2017). In this study, three reinforcement learning models that were already implemented in the *hBayesDM* package were fitted to the data collected from each condition. All parameter estimation procedures were preceded with 4 independent Markov chain Monte Carlo (MCMC) chains with 10,000 iterations after 2,500 warm-up samples. The Gelman-Rubin R-hat Statistics calculated by the package was used to check the convergence of the MCMC chains (Gelman & Rubin, 1992). The R-hat value close to 1 indicated successful convergence.

Model comparison. Model performance was compared based on Leave-One-Out Information Criterion (LOOIC). LOOIC values for each model were calculated using the *loo* package in R. The lower LOOIC value indicates the better model performance since it is a product of the expected log predictive distribution (elpd) times a negative integer (i.e., 2). The elpd is a utility function to measure the predictive accuracy of the model by taking the posterior distribution and likelihood into account. LOOIC has advantages over other simpler model comparison measures like Akaike Information Criterion or Bayesian Information Criterion (Vehtari et al., 2017). The fitting performance of three reinforcement learning models used in this study were compared with the LOOIC values provided by the *printfit* function in the *hBayesDM* package.

Main analyses

Correlational analysis. To assess simple correlations between anxiety scores and parameter values, the Pearson correlation coefficients (R) were calculated. The significance of each correlation coefficient was determined by evaluating its p-value.

Multiple linear regression analysis. In addition to simple linear correlation analyses, multiple linear regression models were used to test if the parameter values were associated with anxiety scores when age, gender, and other psychiatric symptom scores were controlled. The main dependent variables were model-based weights and second-stage learning rates estimated separately for each domain. State and trait anxiety and worry scores were individually included in the models. The model performance was examined based on its F-statistics, adjusted r-squared, and p-value for each estimated beta coefficient.

Moderation effect analysis. For exploratory analyses, regression models with interaction terms were tested to examine a potential moderation effect of cognitive emotion regulation strategies on the anxiety and model-based learning relationship. Independent variables were state anxiety, trait anxiety and worry scores, and the moderator was each of the emotion regulation strategy sub-score. The dependent variables were model-based learning weight parameter values estimated from the reward and punishment conditions. Independent analyses for each domain were performed using the same set of predictors.

Results

Demographic information

Eighty six participants who passed the pre-screening survey and performed a multi-stage decision-making task (see Methods for exclusion criteria) were included in the analyses. They filled out 9 self-report surveys and performed the cognitive task under the reward and punishment conditions in a psychology laboratory located in Seoul National University (see Table 2 for the full demographic information).

Survey measures

Nine self-report surveys were used in the study (Table 2). To comprehensively examine the impact of anxiety on decision-making and learning, anxiety was divided into three components: state anxiety, trait anxiety, and worry. In the analyses, I tried to rule out any potential impacts from other mood factors, mainly depression. Depression score was included in the regression analyses as one of the confounding variables. Depression was measured with two different surveys (i.e., BDI-2 and PHQ-9). Given the high co-morbidity between depression and anxiety, self-reported depression scores were highly correlated with anxiety scores in previous studies. Thus, in this study, the correlation coefficients between anxiety level and each of the two depression score were compared. The depression score that had lower correlation coefficients with anxiety level was included in the analyses to prevent multicollinearity. In Figure 6, the Pearson correlation coefficients for all pairs of the surveys are reported. The correlation coefficients between

Table 2

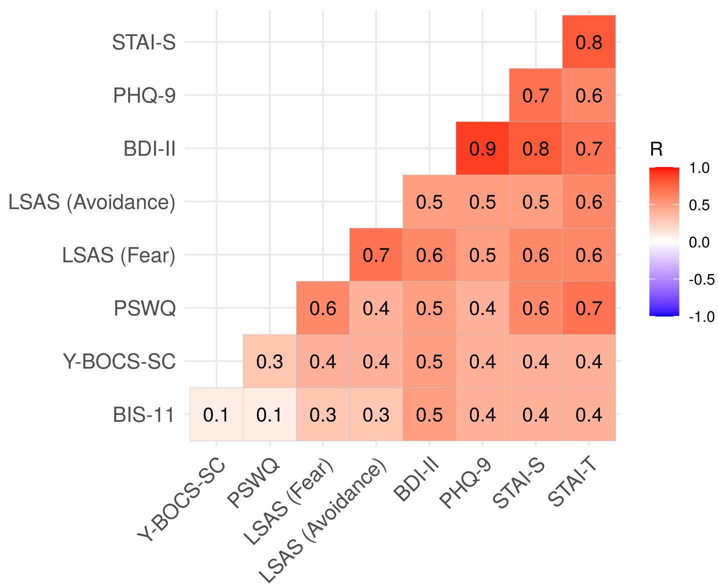
Demographics and survey measures

	Total (N = 86)	
Measures†	Mean	SD
Age	23.31	3.77
Education	15.19	2.04
BDI	10.69	8.55
BIS	62.58	9.24
LSAS	23.41	13.39
PHQ	5.06	4.57
PSWQ	46.22	11.72
STAI-S	41.60	11.08
STAI-T	41.81	10.73
Y-BOCS-SC	6.92	6.72
CERQ	-	-
Acceptance	1.32	0.26
Catastrophizing	0.93	0.40
Other-blame	1.06	0.34
Positive reappraisal	1.70	0.35
Positive refocusing	1.34	0.45
Putting into perspective	1.52	0.36
Refocus on planning	1.82	0.31
Rumination	1.39	0.40
Self-blame	1.39	0.35

† Education = years of education; BDI = Beck Depression Inventory-2; BIS = Barratt Impulsiveness Scale-11; LSAS = Liebowitz Social Anxiety Scale; PHQ = Patient Health Questionnaire-9; PSWQ = Penn State Worry Questionnaire; STAI-S = State Trait Anxiety Inventory-Y State; STAI-T = State Trait Anxiety Inventory-Y Trait; Y-BOCS-SC = Yale-Brown Obsessive-Compulsive Scale-Symptom Checklist; CERQ = Cognitive Emotion Regulation.

Figure 6

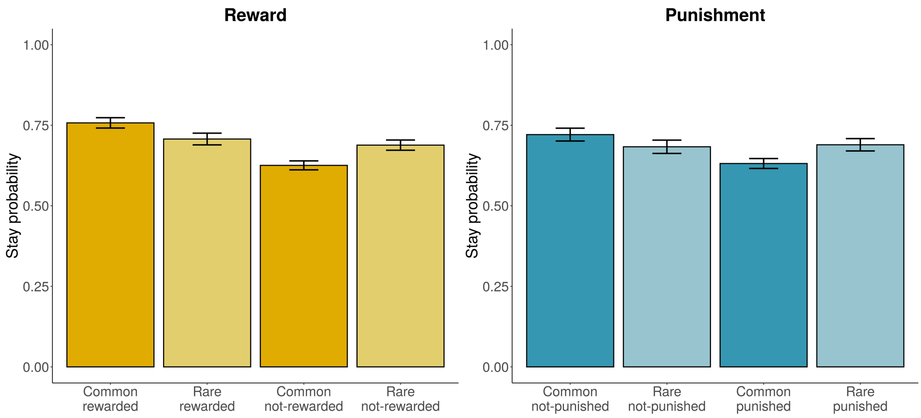
Pearson correlation coefficients (R) among the self-reported survey scores



PHQ-9 and all of the anxiety measures (r = .6, .7, and .4, respectively) were smaller than those between BDI-2 and anxiety measures (r = .7, .8, and .5, respectively). All of the correction coefficients were statistically significant ($p < .001$). Individual PHQ-9 score was used as a covariate to adjust for individual depression level. See Appendix A for the histogram of each of the survey scores.

Figure 7

Stay probabilities in the reward condition (left) and punishment condition (right)



Behavioral analyses

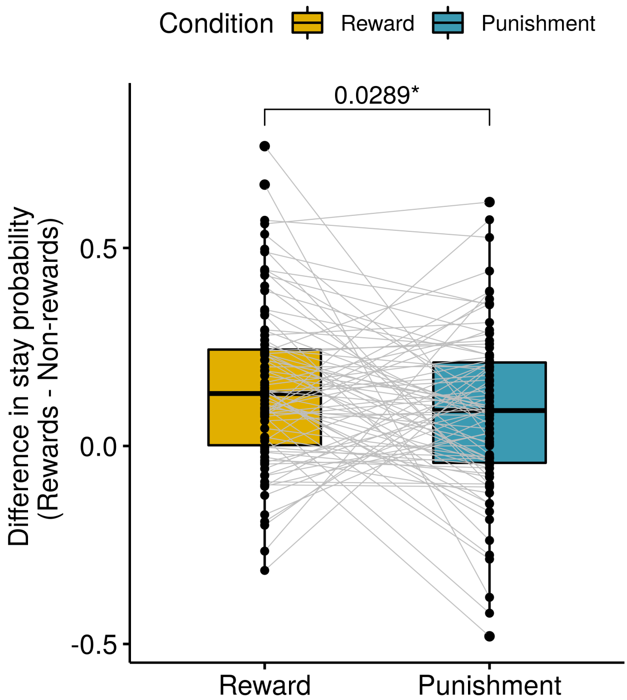
Stay probability

Stay probability was calculated for each of the four combinations of outcome and transition type in the previous trial, separately for each domain. In both reward and punishment conditions, participants performed the task using both model-based and model-free control as shown in Figure 7. These patterns of stay probabilities are consistent with the hybrid theory of decision-making, in which participants base their learning on the weighted combination of the model-based and model-free systems.

Despite this similarity in the overall stay probability patterns,

Figure 8

Comparison on the difference in stay probability after receiving rewards and after receiving non-rewards between the reward and punishment conditions



there was a significant difference in staying after receiving positive outcomes (i.e., a coin gain and or no loss in the reward and punishment condition, respectively) minus staying after receiving negative outcomes (i.e., no gain or a coin loss in the reward and punishment condition, respectively) between the two domains. In a two-tailed t-test for paired samples, the difference in the stay probability (rewards minus

Table 3

Results from mixed-effect logistic regression testing the association between stay probability and condition

Predictors	β (SE)	p
(Intercept)	0.676 (0.019)	**<.001**
Reward	0.033 (0.006)	**<.001**
Transition	-0.005 (0.005)	.302
Condition	-0.006 (0.008)	.447
Reward * Transition	0.024 (0.006)	**<.001**
Reward * Condition	0.010 (0.005)	.063
Transition * Condition	-0.002 (0.005)	.640
Reward * Transition * Condition	0.004 (0.005)	.469

non-rewards) was significantly higher in reward than punishment ($T(85) = 2.22$, $p < .05$; Figure 8). The reward-dependent discrepancy in stay choices became larger when participants were experiencing coin gains than coin losses. This difference could be attributed to either higher sensitivity to rewards than punishments or more exhibition of model-based control in the punishment than reward condition. To test which factor contributed to this difference, mixed-effect logistic regression analyses were conducted.

Mixed-effect logistic regression

Mixed-effect logistic regression analysis was conducted to test if there were a significant difference in model-free and model-based learning between the two domains (Table 3). After controlling for age, order, and sex, a significant main effect of reward on first-level choices

Table 4

Results from mixed-effect logistic regression testing the association among stay probability, state anxiety and condition

Predictors	β (SE)	p
(Intercept)	0.679 (0.019)	**<.001**
Reward	0.032 (0.006)	**<.001**
Transition	-0.005 (0.005)	.339
Condition	0.006 (0.008)	.443
STAI-S	-0.012 (0.013)	.394
Reward * Transition	0.025 (0.006)	**<.001**
Reward * Condition	-0.010 (0.005)	.061
Transition * Condition	-0.003 (0.005)	.537
Reward * STAI-S	0.006 (0.004)	.151
Transition * STAI-S	-0.003 (0.004)	.477
Condition * STAI-S	-0.000 (0.006)	.966
Reward * Transition * Condition	0.004 (0.005)	.398
Reward * Transition * STAI-S	-0.006 (0.005)	.222
Reward * Condition * STAI-S	-0.002 (0.004)	.661
Transition * Condition * STAI-S	0.004 (0.003)	.237
Reward * Transition * Condition * STAI-S	-0.004 (0.004)	.275

was found, meaning participants repeated their first-level choices more often when the previous trials were rewarded or not punished (β = 0.03, SE = 0.006, p < .001). The main effect of reward indicated that the participants exhibited model-free control during the task. There was also a significant reward * transition interaction effect on stay probability (β = 0.03, SE = 0.006, p < .001). The participants repeated their first-level choices more often when the previous trials were rewarded with common transition, indicating their usage of model-based control. The result also

Table 5

Results from mixed-effect logistic regression testing the association among stay probability, trait anxiety and condition

Predictors	β (SE)	p
(Intercept)	0.680 (0.019)	**<.001**
Reward	0.032 (0.006)	**<.001**
Transition	-0.005 (0.005)	.441
Condition	0.006 (0.008)	.608
STAI-T	-0.022 (0.013)	.107
Reward * Transition	0.025 (0.006)	**<.001**
Reward * Condition	0.010 (0.005)	.058
Transition * Condition	-0.003 (0.005)	.504
Reward * STAI-T	0.004 (0.004)	.376
Transition * STAI-T	0.001 (0.004)	.707
Condition * STAI-T	0.001 (0.006)	.835
Reward * Transition * Condition	0.003 (0.005)	.479
Reward * Transition * STAI-T	-0.007 (0.005)	.154
Reward * Condition * STAI-T	-0.002 (0.004)	.666
Transition * Condition * STAI-T	0.004 (0.003)	.219
Reward * Transition * Condition * STAI-T	0.000 (0.004)	.892

revealed a marginally significant reward * condition interaction effect (β = 0.01, SE = 0.008, $p < .10$), but there was no significant three-way interaction effect (reward * transition * condition). This result supported that the participants tended to show higher reward sensitivity in the reward than punishment condition.

Additional mixed-effect logistic regression analyses were performed to investigate whether anxiety scores would contribute to any significant differences in either model-free or model-based learning

Table 6

Results from mixed-effect logistic regression testing the association among stay probability, worry and condition

Predictors	β (SE)	p
(Intercept)	0.678 (0.019)	**<.001**
Reward	0.032 (0.006)	**<.001**
Transition	-0.005 (0.005)	.331
Condition	0.006 (0.008)	.435
PSWQ	-0.008 (0.014)	0.552
Reward * Transition	0.024 (0.006)	**<.001**
Reward * Condition	0.009 (0.005)	.079
Transition * Condition	-0.002 (0.005)	.649
Reward * PSWQ	0.005 (0.004)	.225
Transition * PSWQ	-0.002 (0.004)	.593
Condition * PSWQ	-0.002 (0.006)	.706
Reward * Transition * Condition	0.004 (0.005)	.442
Reward * Transition * PSWQ	-0.004 (0.005)	.407
Reward * Condition * PSWQ	0.003 (0.004)	.432
Transition * Condition * PSWQ	-0.000 (0.003)	.949
Reward * Transition * Condition * PSWQ	-0.002 (0.004)	.668

between the reward and punishment domains. State anxiety, trait anxiety, and worry were separately included in the mixed-effect logistic regression models with the same predictors (i.e., reward, transition, and their interaction term) and covariates (i.e., age, order, and sex). In all of the three analyses, there was a significant main effect of reward and interaction effect of reward * transition on first-level choices (Table 4 for state anxiety; Table 5 for trait anxiety; Table 6 for worry). A marginally significant reward * condition interaction effect was also found in all of

Table 7

Results from mixed-effect logistic regression testing the association between stay probability and obsession and compulsion in the reward condition

Predictors	β (SE)	p
(Intercept)	0.682 (0.018)	**<.001**
Reward	0.042 (0.008)	**<.001**
Transition	-0.008 (0.006)	.240
Y-BOCS-SC	-0.022 (0.013)	.097
Reward * Transition	0.027 (0.008)	**.001**
Reward * Y-BOCS-SC	0.004 (0.006)	.522
Transition * Y-BOCS-SC	0.001 (0.005)	.879
Reward * Transition * Y-BOCS-SC	-0.013 (0.006)	**.029**

the three analyses. However, there were no interaction effects between each of the survey scores and other predictors on stay probability.

In previous literature, it has been consistently reported that compulsivity is associated with model-based deficits in the reward condition (Gillan et al., 2016; Voon, Baek, et al., 2015). To test whether this finding could be replicated, I ran a supplementary mixed-effect logistic regression analysis with obsession and compulsion (OC) score and stay probability in the reward condition (Table 7). There was a significant main effect of reward (β = 0.042, SE = 0.008, $p < .001$) and interaction effect of reward * transition (β = 0.027, SE = 0.008, p = .001). Also, a significant three way interaction effect of reward * transition * OC score was found (β = -0.013, SE = 0.006, $p < .05$). One standard deviation increase in OC score was associated with a significant

Table 8

Model comparison results: LOOIC

Domain	Model[a]	LOOIC[b]
Reward	7-par	27,792.35
	6-par	27,788.55
	4-par	28,133.06
Punishment	7-par	28,048.23
	6-par	28,048.74
	4-par	28,235.52

[a] 7-par = 7-parameter; 6-par = 6-parameter; 4-par = 4-parameter.

[b] LOOIC = Leave-One-Out Information Criterion.

decrease in the interaction effect of reward * transition on stay probability, meaning this study successfully replicated the previous finding.

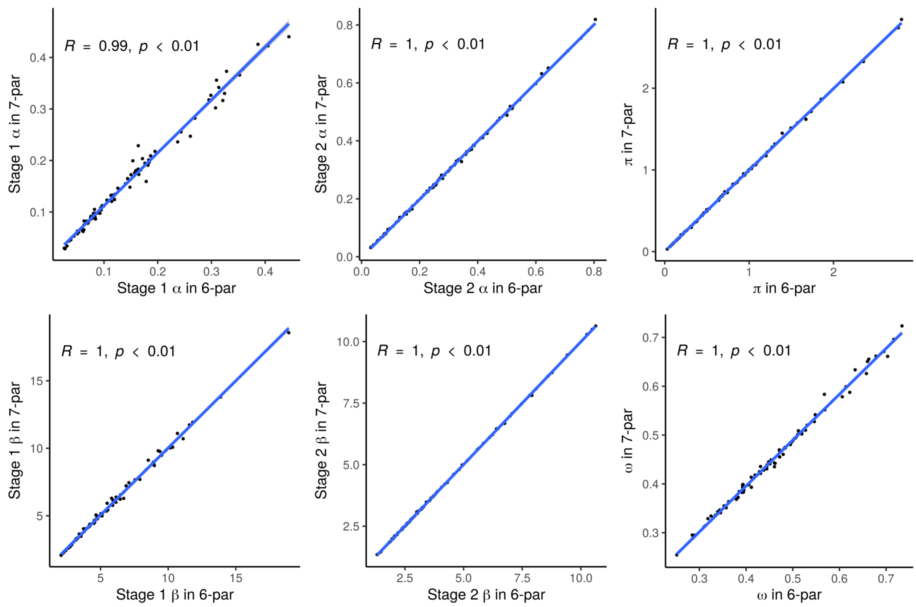
Computational modeling results

Model comparison

Three reinforcement learning models were fitted to the data from the reward and punishment conditions separately. The best model for each condition was determined based on the LOOIC value (Table 8). In the reward domain, 6-par was the winning model with the lowest LOOIC value. In the punishment domain, 6-par and 7-par showed nearly identical performance. Since 6-par and 7-par were nested, I performed post-hoc correlational analyses to check whether the individual-level model parameter values from the two models were significantly different. All of the mean estimates of the 6 overlapping

Figure 9

Correlation plots between individual-level parameter estimates from the 6-parameter (6-par) and 7-parameter (7-par) models



parameters showed significant correlations with the coefficient values close or equal to 1 (Figure 9). Given this similarity in the parameter values from the two models, the posterior mean estimates from 6-par were used in the further analyses for the punishment domain to match the winning models between the two conditions.

Note that all of the three models were identical to the ones developed by Daw and his colleagues in 2011 (Daw et al., 2011) and validated by previous literature (e.g., Gillan et al., 2016; Otto et al.,

Table 9

Group-level mean (standard deviation) parameter estimates on the winning model (6-par) for both reward and punishment conditions

Model parameters†	Reward	Punishment	Between-condition	HDI
6-par:				
α_1	0.13 (0.03)	0.10 (0.03)	-0.05	0.12
β_1	5.24 (0.53)	4.25 (0.52)	-0.46	2.46
α_2	0.17 (0.02)	0.25 (0.03)	**-0.15**	**-0.01**
β_2	4.41 (0.25)	3.47 (0.26)	**0.25**	**1.67**
π	0.50 (0.05)	0.45 (0.07)	-0.13	0.21
ω	0.50 (0.05)	0.46 (0.05)	-0.10	0.18

† α_1 = first-stage learning rate; β_1 = first-stage inverse temperature; α_2 = second-stage learning rate; β_2 = second-stage inverse temperature; π= perseverance; ω = model-based weight.

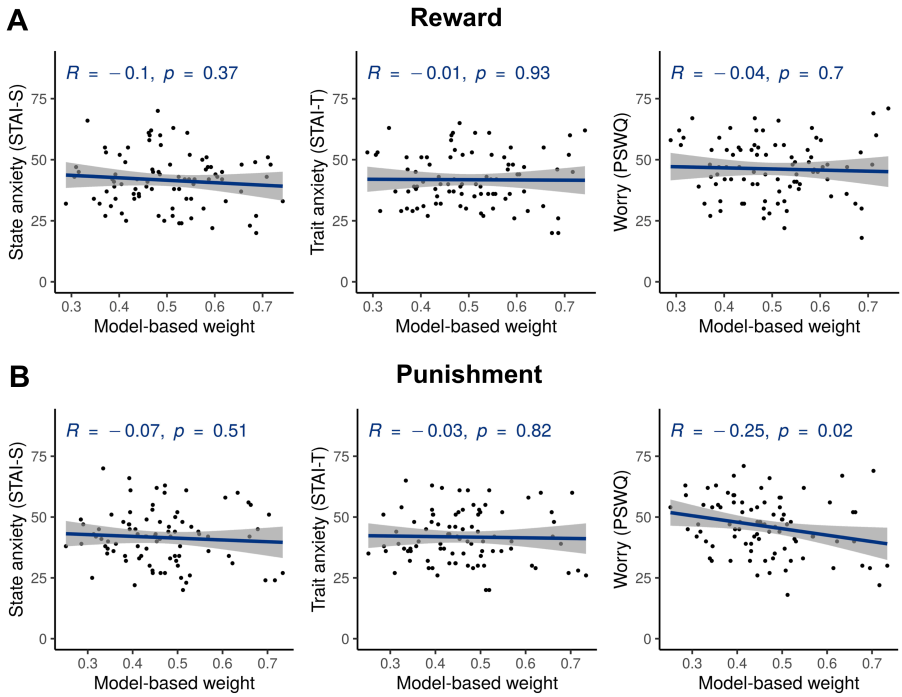
2013; Wunderlich et al., 2012). Thus, model validation using parameter recovery and posterior predictive check was skipped in this study.

Parameter estimates

The summary of group-level parameter estimates is provided in Table 9. The values under the reward and punishment columns represent the mean posterior estimates from the winning model (i.e., 6-par) fitted to each domain. For each of the parameters, between-condition 95% highest density interval (HDI) was also calculated. If HDI does not overlap zero, I conclude there is a significant difference between the two parameter estimates. All group-level parameters except second-stage inverse temperature (β_2) did not show significant differences between the

Figure 10

Correlation plots between anxiety level and individual-level model-based weight in the (A) reward condition and (B) punishment condition



two domains. The non-significant between-condition difference in model-based weight (ω) was consistent with the mixed-effect logistic regression analysis result that the model-based weight term (reward * transition) was not significantly moderated by condition.

Correlation analyses

Anxiety and model-based weight

Pearson correlation coefficients (r) were calculated to measure the association between anxiety level and model-based weight parameter ω (Figure 10). I mainly examined whether there would be a condition-dependent difference in the correlation between model-based weight and anxiety score. In the correlation analyses, the individual-level mean posterior estimates were used to quantify each participant's model-based weight. Anxiety was measured using self-reported state anxiety, trait anxiety and worry scores. State anxiety was not associated with model-based weight in both reward and punishment conditions (r = -.10, p = .37 in reward; r = -.07, p = .51 in punishment). Similarly, there was no significant association between trait anxiety and model-based weight in both conditions (r = -.01, p = .93 in reward; r = -.03, p = .82 in punishment). Worry was also not associated with model-based weight in the reward condition (r = -.04, p = .70). However, there was a significantly negative correlation between worry level and model-based weight in the punishment condition (r = -.25, $p < .05$).

I also calculated the False Discovery Rate (FDR) adjusted p-values for each pair of the correlations as three correlation analyses were performed simultaneously in each domain. Based on the FDR-adjusted p-values, state anxiety, trait anxiety, and worry scores were still not significantly associated with model-based weight in the reward condition as expected (FDR-corrected p = .93, .93, and .83, respectively). In the punishment condition, there were no significant correlations between state and trait anxiety and model-based weight (FDR-corrected $p = .76$ and $.82$, respectively), whereas worry had a marginally significant association with model-based weight in the punishment condition (FDR-corrected $p < .10$).

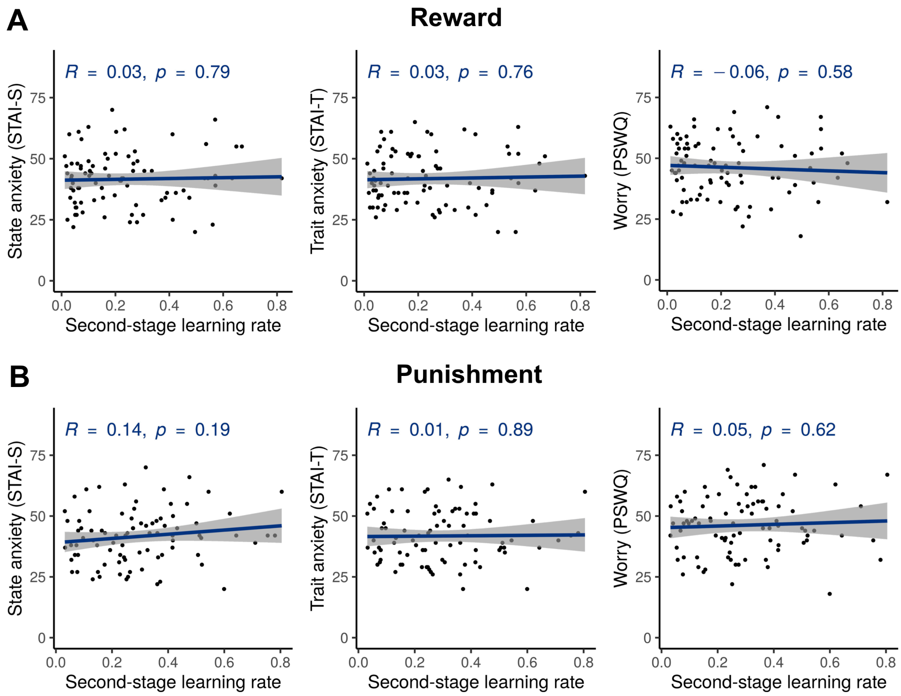
As a supplementary analysis, the Pearson correlation coefficient between each of all survey scores and model-based weight was calculated (Appendix B). None of the remaining survey measures (i.e. depression, impulsivity, obsession and compulsion, and social anxiety) showed a significant association with model-based weight in both conditions.

Anxiety and second-stage learning rate

The correlation between second-stage learning rate α_2 and anxiety level was calculated. Second-stage learning rate quantifies how fast an individual updates the second-stage action value after observing an outcome. I investigated whether the association between the learning rate parameter and anxiety level would be modulated by condition. In both conditions, all of the three anxiety scores did not show any significant associations with second-stage learning rate (Figure 11). Also, none of the remaining survey scores was associated with second-stage learning rate (Appendix B).

Figure 11

Correlation plots between anxiety level and individual-level second-stage learning rate in the (A) reward condition and (B) punishment condition



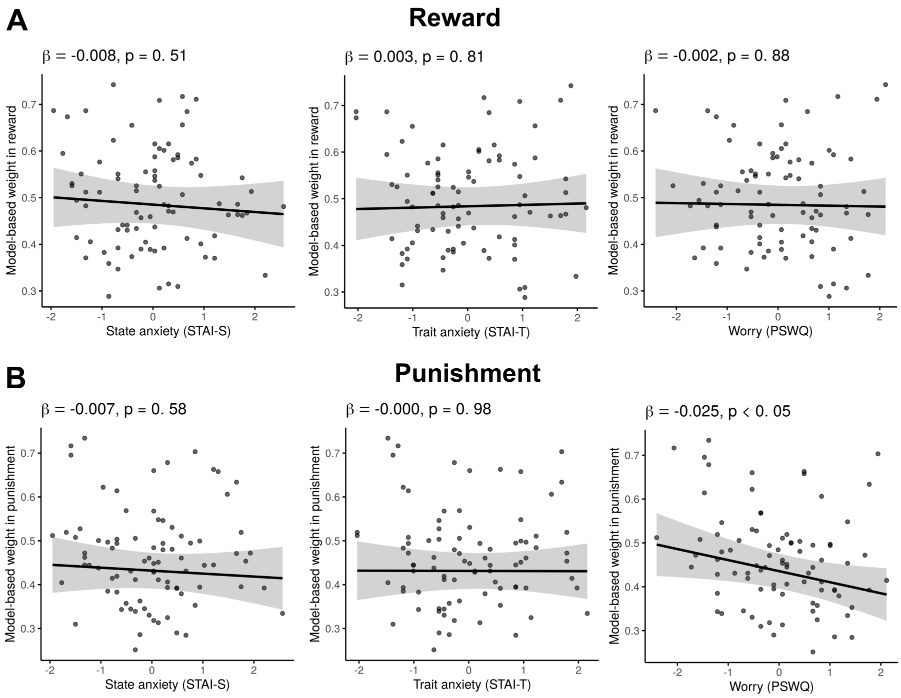
Multiple linear regression analyses

Independent variable I: model-based weight

Multiple linear regression analyses were performed to examine if anxiety was a significant predictor of model-based weight after adjusting for demographic information (i.e., age and sex), condition order, and other survey measures. Mainly, I tested whether worry remained as a

Figure 12

Main effect plots between anxiety level and individual-level model-based weight in the (A) reward condition and (B) punishment condition



significant predictor of model-based weight in punishment. The models were tested separately for each condition.

The first multiple linear regression model (Model 1) included the z-scored age and binary variables of sex and order as covariates along with the main predictor of interest, three anxiety scores as a dependent variable in separate models (Figure 12). The results were nearly identical to the correlation analyses. State anxiety was not a significant predictor

Table 10

Association between model-based weight in the reward condition and worry after adjusting for confounding variables

	Reward condition[a]							
	Model 1		Model 2		Model 3		Model 4	
Questionnaire[b]	β(SE)	p	β(SE)	p	β(SE)	p	β(SE)	p
Worry	-0.002(0.012)	.879	-0.003(0.013)	.815	-0.003(0.013)	.803	-0.004(0.013)	.780
Obsession			0.003(0.013)	.782			0.003(0.013)	.847
Depression					0.004(0.013)	.783	0.003(0.014)	.848
Adjusted R^2	-0.026		-0.038		-0.038		-0.050	
F-statistics[c]	0.466		0.384		0.384		0.322	

[a] All models are controlled for age (z-scored), order, and sex.

[b] Worry = Penn State Worry Questionnaire; Obsession = Yale-Brown Obsessive-Compulsive Score-Symptom Scale; Depression = Patient Health Questionnaire-9. All survey scores are z-scored.

[c] Notation for significance level: . ($p < .1$), * ($p < .05$), ** ($p < .01$), *** ($p < .001$)

of model-based weight in either domain (β = - 0.001, SE = 0.001, p = .51 in reward; β = - 0.001, SE = 0.001, p = .58 in punishment). Trait anxiety also did not significantly predict model-based weight in both reward and punishment conditions (β = 0.000, SE = 0.001, p = .81 in reward; β = -0.000, SE = 0.001, p = .98 in punishment). Similarly to state and trait anxiety measures, worry was not a significant predictor of model-based weight in the reward condition (β = -0.000, SE = 0.001, p = .88, Model 1 in Table 10). However, it significantly predicted model-based weight in the punishment condition (β = -0.002, SE = 0.001, $p < .05$, Model 1 in Table 11). Even after controlling for age, order and sex, 1 standard deviation increase in worry score was associated with a significant decrease in model-based weight by 0.002 (Figure 13).

Table 11

Association between model-based weight in the punishment condition and worry after adjusting for confounding variables

	Punishment condition[a]							
	Model 1		Model 2		Model 3		Model 4	
Questionnaire[b]	β(SE)	p	β(SE)	p	β(SE)	p	β(SE)	p
Worry	-0.025(0.012)	**.033**	-0.029(0.012)	**.020**	-0.025(0.013)	.050	-0.029(0.013)	**.035**
Obsession			0.012(0.012)	.322			0.014(0.013)	.300
Depression					0.000(0.013)	.971	-0.004(0.013)	.748
Adjusted R^2	0.060		0.060		0.048		0.049	
F-statistics[c]	2.359.		2.086.		1.864		1.736	

[a] All models are controlled for age (z-scored), order, and sex.

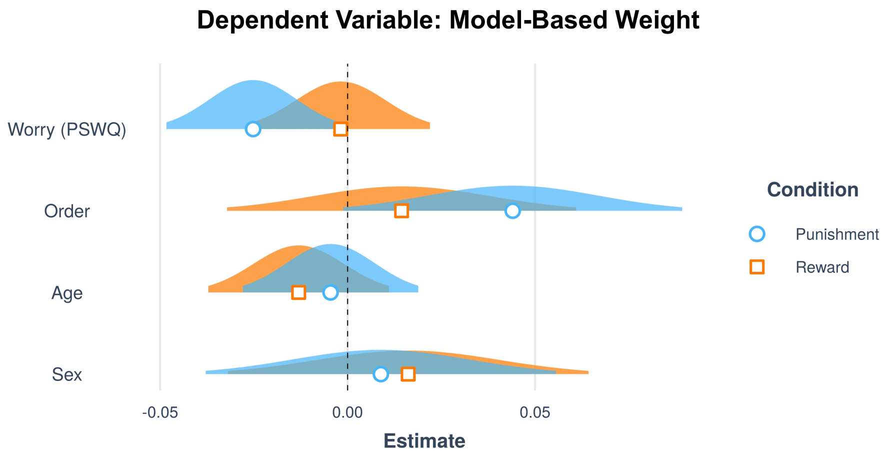
[b] Worry = Penn State Worry Questionnaire; Obsession = Yale-Brown Obsessive-Compulsive Score-Symptom Scale; Depression = Patient Health Questionnaire-9. All survey scores are z-scored.

[c] Notation for significance level: . ($p < .1$), * ($p < .05$), ** ($p < .01$), *** ($p < .001$)

I also investigated whether worry remained as a significant predictor when adjusting for other survey measures. I specifically looked into the effects of depression and OC and included them in the models as covariates along with age, order and sex. This was because these two constructs were the main psychological variables that previous literature have focused on when studying individual differences in model-based control (e.g., Gillan et al., 2016). I performed three additional multiple linear regression analyses (Model 2-4 in Table 10 for reward; Table 11 for punishment). Model 2 examined the association between worry and model-based weight while adjusting for OC level. Model 3 included worry and depression level, and Model 4 included all three survey scores. In all three models, worry did not predict model-based weight in reward

Figure 13

Beta distributions for the multiple regression analysis examining the association between worry and individual-level model-based weight in the punishment condition

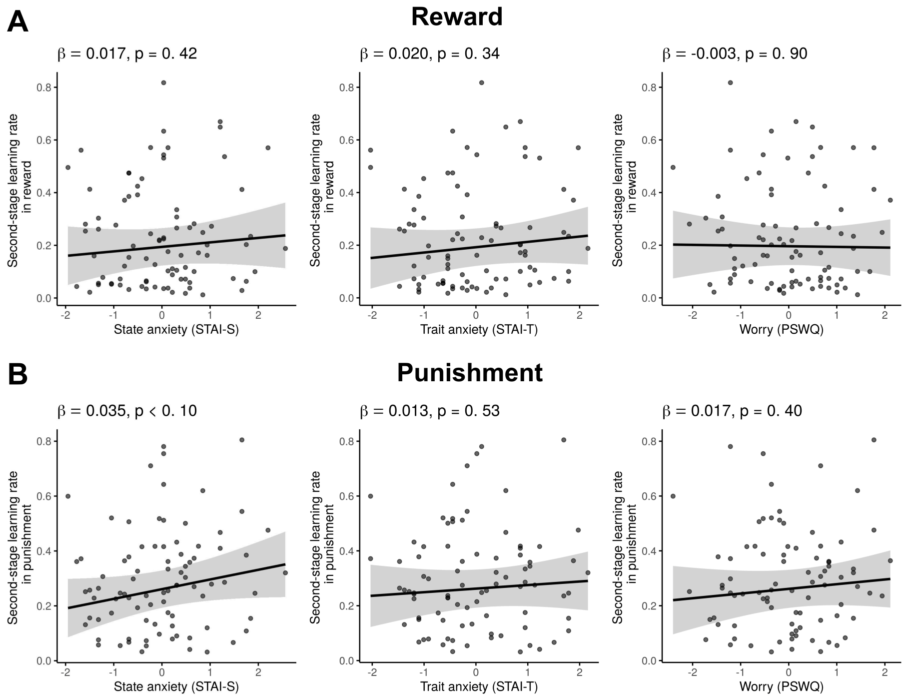


(β = -0.003, SE = 0.013, p = .82 in Model 2; β = -0.003, SE = 0.013, p = .80 in Model 3; β = -0.004, SE = 0.013, p = .78 in Model 4).

Unlike in the reward domain, worry was a significant predictor of model-based weight in punishment in all of the three models (β = -0.029, SE = 0.012, $p < .05$ in Model 2; β = -0.025, SE = 0.013, p = .05 in Model 3; β = -0.029, SE = 0.013, $p < .05$ in Model 4). When OC and depression levels were taken into account, worry still had a significantly negative association with model-based weight but only in the punishment condition.

Figure 14

Main effect plots between anxiety level and individual-level second-stage learning rate in the (A) reward condition and (B) punishment condition



Independent variable II: second-stage learning rate

Similar to the analyses to predict model-based weight ω, Model 1 included each anxiety level and covariates (i.e. age, order and sex) to predict second-stage learning rate α_2 estimated from each domain separately (Figure 14). In the reward condition, all three anxiety scores were not significant predictors of second-stage learning rate ($\beta = 0.017$,

Table 12

Association between second-stage learning rate in the reward condition and state anxiety after adjusting for confounding variables

	Reward condition[a]							
	Model 1		Model 2		Model 3		Model 4	
Questionnaire[b]	β(SE)	p	β(SE)	p	β(SE)	p	β(SE)	p
State anxiety	0.017(0.021)	.417	0.040(0.022)	.083	0.047(0.030)	.112	0.058(0.030)	.051
Obsession			-0.052(0.022)	**.019**			-0.048(0.022)	**.036**
Depression					-0.043(0.029)	.151	-0.029(0.029)	.321
Adjusted R^2	0.045		0.100		0.058		0.100	
F-statistics[c]	1.999		2.842*		2.041.		2.534*	

[a] All models are controlled for age (z-scored), order, and sex.

[b] State anxiety = State-Trait Anxiety Inventory-Y State; Obsession = Yale-Brown Obsessive-Compulsive Score-Symptom Scale; Depression = Patient Health Questionnaire-9. All survey scores are z-scored.

[c] Notation for significance level: . ($p < .1$), * ($p < .05$), ** ($p < .01$), *** ($p < .001$)

SE = 0.021, p = .42 for state anxiety; β = 0.000, SE = 0.002, p = .34 for trait anxiety; β = -0.000, SE= 0.002, p = .90 for worry). Similarly in the punishment condition, trait anxiety and worry were not associated with second-stage learning rate (β = 0.001, SE = 0.002, p = .53 for trait anxiety; β = 0.001, SE = 0.002, p = .40 for worry). Only state anxiety showed a marginally significant association with second-stage learning rate in punishment (β = 0.035, SE = 0.020, $p < .10$). There was a tendency that 1 standard deviation increase in state anxiety was associated with a significant increase in second-stage learning rate in punishment by 0.035 after age, order and sex were controlled.

As post-hoc analyses, I tested nearly identical models to predict second-stage learning rate except having state anxiety as a main

Table 13

Association between second-stage learning rate in the punishment condition and state anxiety after adjusting for confounding variables

	Punishment condition[a]							
	Model 1		Model 2		Model 3		Model 4	
Questionnaire[b]	β(SE)	p	β(SE)	p	β(SE)	p	β(SE)	p
State anxiety	0.035(0.020)	.082	0.043(0.022)	.056	0.074(0.028)	**.009**	0.076(0.028)	**.009**
Obsession			-0.017(0.022)	.415			-0.009(0.022)	.672
Depression					-0.055(0.028)	.051	-0.053(0.029)	.070
Adjusted R^2	0.029		0.025		0.063		0.049	
F-statistics[c]	1.641		1.441		2.143.		1.798	

[a] All models are controlled for age (z-scored), order, and sex.

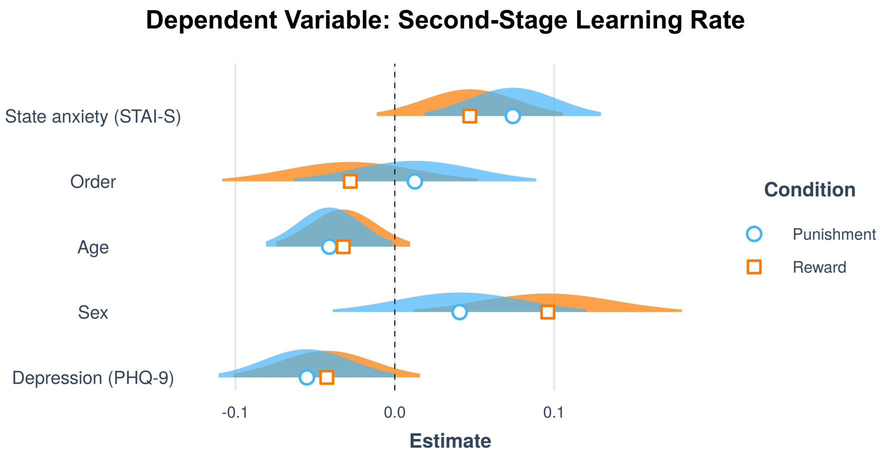
[b] State anxiety = State-Trait Anxiety Inventory-Y State; Obsession = Yale-Brown Obsessive-Compulsive Score-Symptom Scale; Depression = Patient Health Questionnaire-9. All survey scores are z-scored.

[c] Notation for significance level: . ($p < .1$), * ($p < .05$), ** ($p < .01$), *** ($p < .001$)

dependent variable (Model 2-4 in Table 12 for reward; Model 2-4 in Table 13 for punishment). I mainly investigated the association between state anxiety and second-stage learning rate, given the marginally significant association between state anxiety and second-stage learning rate in punishment. In Model 2 and Model 4 where state anxiety was tested with OC score only and both OC and depression, respectively, state anxiety was marginally associated with second-stage learning rate in reward ($\beta = 0.040$, SE = 0.022, $p < .10$ in Model 2; $\beta = 0.058$, SE = 0.030, $p < .10$ in Model 4). However, when controlling for depression only in Model 3, state anxiety no longer had a marginally significant association with second-stage learning rate in reward ($\beta = 0.047$, SE = 0.030, $p = .11$).

Figure 15

Beta distributions for the multiple regression analysis examining the association between state anxiety and individual-level second-stage learning rate in the punishment condition



Here, OC score was also a significantly positive predictor of second-stage learning rate in reward (β = -0.052, SE = 0.022, $p < .05$ in Model 2; β = -0.048, SE = 0.022, $p < .05$ in Model 4). As a post-hoc analysis, Model 1 with OC score as a sole dependent variable was tested. OC score was marginally associated with second-stage learning rate in reward (β = -0.037, SE = 0.020, $p < .10$).

In the punishment condition, state anxiety significantly predicted second-stage learning rate in all models (Table 13). Given the highest adjusted R-squared value of Model 3, Figure 15 illustrates the distributions of Model 3 beta coefficients separately for the reward and punishment conditions. State anxiety was possibly associated with second-stage learning rate in punishment but not in reward.

Exploratory analysis: moderation effect

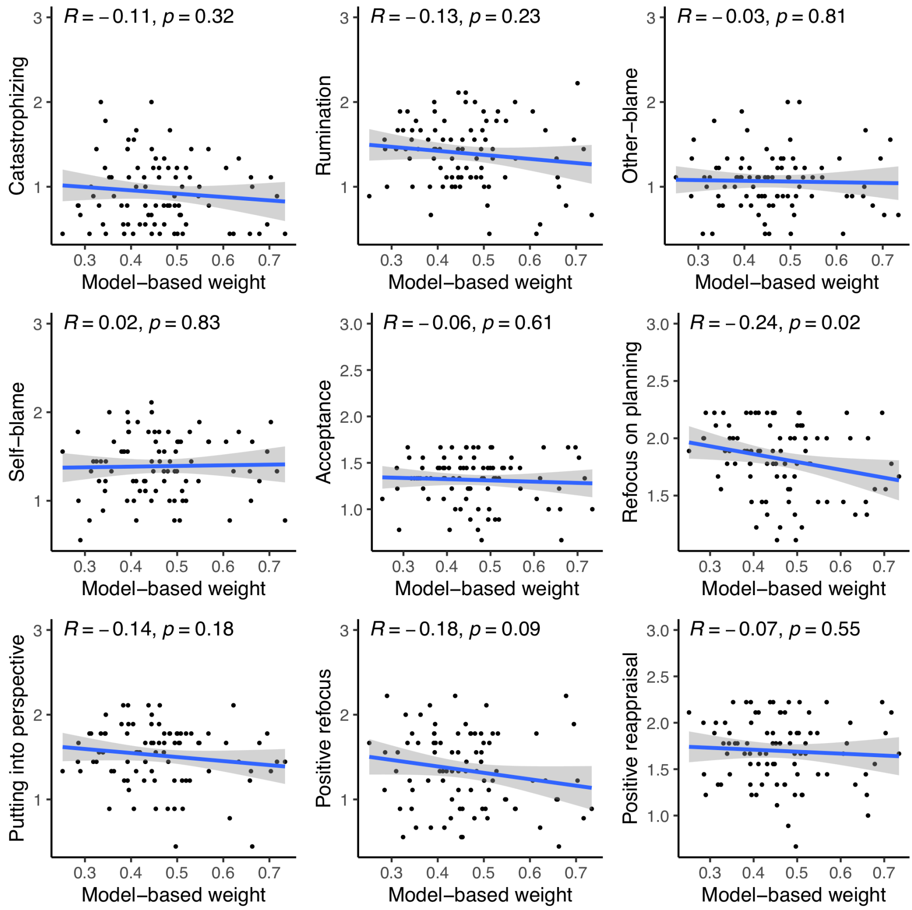
As exploratory analyses, I tested for a moderation effect to probe which cognitive content might have contributed to the negative association between worry and model-based weight in punishment. I used 9 sub-scores from the Cognitive Emotion Regulation Strategy (CERQ) survey, which assesses both adaptive and nonadaptive emotional regulation strategies (i.e., catastrophizing, rumination, other-blame, self-blame, acceptance, refocus on planning, putting into perspective, positive refocus, and positive reappraisal).

First, I examined whether there was a significant correlation between each sub-score and model-based weight in punishment (Figure 16). Only refocus on planning sub-score was negatively correlated with model-based weight in punishment (r = -.24, $p < .05$). All other sub-scores did not have significant correlations with model-based weight, except that positive refocus was marginally associated with model-based weight in punishment (r = -.18, $p < .10$).

Next, I tested if there was a moderation effect of refocus on planning on the worry and model-based weight relationship (Figure 17). An interaction term (worry * refocus on planning) was included in the model along with the main effect terms of worry and refocus on planning while controlling for age, order and sex. There were main effects of worry (β = -0.029, SE = 0.011, $p < .05$) and refocus on planning (β =

Figure 16

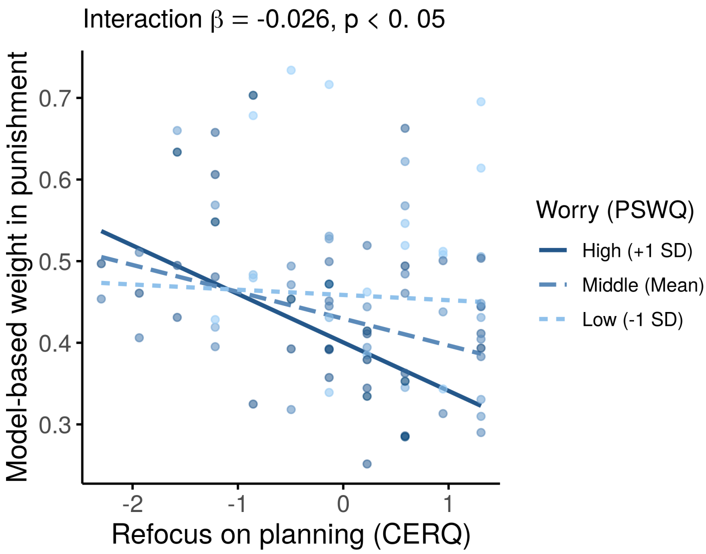
Correlation plots between for each Cognitive Emotion Regulation Strategies sub-score and model-based weight in the punishment condition



-0.033, SE = 0.011, $p < .01$) on model-based weight in punishment. Also, the relationship between worry and model-based weight in punishment was significantly moderated by refocus on planning score (β

Figure 17

Interaction effect plot from the moderation effect analysis testing the impact of the refocus on planning strategy on the association between worry and model-based weight in the punishment condition



= -0.026, SE = 0.013, $p < .05$). Only in the high worry group, participants with higher refocus on planning scores exhibited significantly less model-based control. This moderation effect was only significant in the punishment condition.

Discussion

In this study, I mainly tested for four hypotheses: 1) anxiety levels are not associated with model-based weight in the reward condition, 2) people with higher anxiety scores show higher learning rate in the punishment condition, 3) anxiety levels show significantly negative correlations with model-based weight in the punishment condition, and 4) certain cognitive emotion regulation strategies moderate the association between anxiety levels and model-based weight in punishment. Additionally, I investigated whether there were any systematic differences in decision-making and choice behavior between the reward and punishment condition by calculating and comparing behavioral and computational modeling measures. I also conducted analyses with other psychiatric symptom measures focusing on obsession and compulsion (OC) scores in order to replicate the previous findings.

Using mixed-effect logistic regression analyses, I found there was no significant interaction effect of anxiety scale on either reward or reward * transition interaction in the reward condition. This result was consistent with the correlation and multiple linear regression analyses using the parameter estimates from computational modeling. None of the anxiety measures had a significant association with model-based weight in the reward condition. This null finding remained even after age, sex, condition order, and other psychiatric symptom measures (i.e., OC and depression) were controlled. Overall, the results replicated the previous findings that the anxiety construct was not significantly associated with model-based weight in the reward domain (Gillan et al., 2020; Gillan et al., 2016).

My second replication analysis was focused on second-stage learning rate, which measures how fast an individual updates its action value based on the recent outcome history. A previous literature examined that unmedicated mood and anxiety patients showed significantly higher punishment learning rate than their control group, meaning they updated their action values abbertanly faster when experiencing aversive outcomes. However, their group-level mean reward learning rate was not significantly different from the control group's (Aylward et al., 2019). Based on this finding, I hypothesized that people with higher anxiety levels would exhibit similar learning patterns by updating their action values faster in the second-stage where they observed monetary outcomes. I expected that the increase in second-stage learning rate would be significant only in the punishment condition in which participants were experiencing sequential monetary losses.

Consistent with my hypothesis, one of the anxiety levels, state anxiety, showed a significantly positive association with second-stage learning rate only in the punishment condition. The significance got significantly improved when either depression or both depression and OC scales were controlled in the multiple linear regression analyses. One of the remaining questions, however, is that trait anxiety was not significantly associated with second-stage learning rate. Given the high correlation coefficient value between state and trait anxiety (r = .80, p < .001), this discrepancy in the results was rather unexpected. A possible explanation would be that there might have been an interaction effect between state anxiety and emotional disturbance caused by experiencing monetary losses. People who were experiencing high state anxiety might have become more sensitive to negative outcomes during the punishment condition, leading to faster changes in their choice behavior upon the recent negative outcomes. Further studies would be needed to conclude whether state anxiety rather than trait anxiety plays a more significant role in altering learning rate.

One of the main findings in this study is that people with a higher worry score exhibited less model-based control in punishment. Individual worry score was significantly correlated with model-based weight in punishment, and the relationship remained still even after controlling for age, sex, order, OC and depression. This consistency supported the robustness of the result. According to the cognitive model of pathological worry, two of the cognitive characteristics in worry are attentional bias towards threatening information and its quasi-verbal form (Hirsch & Mathews, 2012). Worry might have led highly worrying people to focus more on the negative outcome than the transition history in the previous history. Also, the verbal nature of worry might have caused a conflict with representing a cognitive map for model-based planning. It has been revealed that the dorsal hippocampus plays an essential role in model-based planning in rodents (Miller et al., 2017). As successful model-based planning involves a clear spatial or visual representation of the task structure, worry that most likely exists in a verbal form may be conflicted with utilizing model-based control especially in the punishment condition where people might have worried significantly more than in the reward condition. As the real-time worry was not assessed during the experiment sessions, further study would be required to better understand how monetary loss outcome impacts current worry level and alters the usage of model-based control in learning.

Still, the possibility that worry level had got significantly increased during the punishment condition was supported by the exploratory analyses I conducted. I performed correlational analyses to understand whether there were any significant associations between cognitive emotion regulation strategy sub-scores and model-based weights from both conditions. Refocus on planning, which is to actively switch attention to come up with some plans to get out of the aversive situations, was negatively correlated with model-based weight in punishment but not in reward. In addition, refocus on planning served as a moderator between the worry score and model-based weight. It was significantly associated with model-based weight only in the high worry group. Thus, a cognitive emotion regulation strategy like refocus on planning explained individual differences in the impact of worry on model-based learning. Not all highly worrying people exhibited lower model-based weight in punishment, but it was when they strategically tended to alter their attention to refocus on planning to escape from the aversive condition.

Refocus on planning is known as one of the adaptive cognitive emotion regulation strategies (Garnefski & Kraaij, 2007). This finding that an increase in the refocus on planning scale is significantly associated with a decrease in model-based weight in punishment only in the high worry group might be inconsistent with a previous finding. For example, refocus on planning was one of the positive predictors of academic achievement as refocus on planning score had a significantly positive correlation with GPA among university students (Al-badareen, 2016). However, it has been reported that refocus on planning had a positive correlation with other nonadaptive cognitive emotion regulation strategy subscales such as catastrophizing ($r = .23$) and self-blame ($r = .30$) (H.-n. Ahn et al., 2013). Previous literature also suggested that categorizing a certain cognitive emotion regulation strategy into either adaptive or nonadaptive is less recommended as adaptiveness is dependent on the situations (Garnefski & Kraaij, 2007). Thus, it is more likely that the adaptive cognitive emotion regulation strategies serve both positive and negative roles. This study supports this hypothesis that more reliance on the refocus on planing strategy could sometimes be associated with negative performance on various dimensions, including model-based learning.

In the additional analyses, I found that there was a significant difference in stay probabilities (rewards versus non-rewards) between the two conditions. After examining the results from the mixed-effect logistic regression analyses and computational modeling analyses, a tentative hypothesis that people marginally became more reward sensitive in the reward condition than in the punishment was supported. Lastly, the previous finding on impaired model-based control in the OC group under monetary reward was also replicated in this study, supporting the validity of the task and overall experimental procedures used in this study.

Limitations and future studies

There are two major limitations in this study. First, the inconsistent findings among the analyses with state anxiety, trait anxiety and worry were not thoroughly investigated. In this study, only worry scores had a significantly negative correlation with model-based weight in punishment. As worry showed significantly positive correlations with both state (r = .58, p < .001) and trait anxiety (r = .71, p < .001), it was expected that state and trait anxiety scores might have also had significant associations with model-based weight in punishment, which was not found in this study. One possible explanation for this inconsistency is that model-based weight in punishment is significantly associated only with cognitive components of anxiety. The questionnaire I used to measure state and trait anxiety, STAI-Y, measures both cognitive and physical symptoms of anxiety (Han et al., 1996; Spielberger, 1983). In a previous literature, researchers performed exploratory analyses using two subscales of anxiety, cognitive and physical symptoms scores, in order to prevent their analyses from being obscured by the sum scores (Wise & Dolan, 2020). As STAI-Y does not dissociate cognitive and physical components of anxiety but rather calculates an aggregate score of anxiety, a future study with a questionnaire that consists with different subscales of anxiety (e.g., the State and Trait Inventory of Cognitive and Somatic Anxiety (Grös et al., 2007)) is recommended.

Also, this study only included non-clinical participants, which might have contributed to insignificant relationships between state and trait anxiety scores and model-based weight parameters. Previous studies on investigating the relationship between anxiety and model-based weight mainly focused on non-clinical populations (Gillan et al., 2020; Gillan et al., 2016). To this date, it still remains unclear whether anxious patients would show model-based learning deficits compared to a non-clinical group. Due to the unusual circumstance (i.e., COVID-19 pandemic), this study was able to recruit the non-clinical samples only, who showed less hesitance to visit the lab and participate the experiment in-person. This situation might have caused a sampling bias that excluded people with higher or more severe anxiety symptoms. Thus, a future study with anxiety disorder patients is needed in order to probe possible group-level differences in model-based learning by between-group analyses.

References

Ahn, H.-n., Joo, H., & Lee, N. (2013). Validation of the cognitive emotion regulation questionnaire in a korean population. *Korea Journal of Counseling*, *14*, 1773–1794. https://doi.org/10.15703/kjc.14.3.201306.1773

Ahn, W.-Y., & Busemeyer, J. R. (2016). Challenges and promises for translating computational tools into clinical practice. *Current Opinion in Behavioral Sciences*, *11*, 1–7. https://doi.org/10.1016/j.cobeha.2016.02.001

Ahn, W.-Y., Dai, J., Vassileva, J., Busemeyer, J. R., & Stout, J. C. (2016). Chapter 3 Computational modeling for addiction medicine From cognitive models to clinical applications. *Progress in Brain Research*, *224*, 53–65. https://doi.org/10.1016/bs.pbr.2015.07.032

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*, *1*, 24–57. https://doi.org/10.1162/cpsy_a_00002

Al-badareen, G. (2016). Cognitive emotion regulation strategies as predictors of academic achievement among university students. *Journal of Educational and Psychological Studies*, *10*(4), 680–686.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: Diagnostic and statistical manual of mental disorders* (Fifth). https://doi.org/10.1176/appi.books.9780890425596.

Aylward, J., Valton, V., Ahn, W.-Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour*, *3*(10), 1116–1123. https://doi.org/10.1038/s41562-019-0628-0

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4-5), 407–419. https://doi.org/10.1016/s0028-3908(98)00033-1

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for the Beck Depression Inventory -II. San Antonio, TX: Psychological Corporation.

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596. https://doi.org/10.1038/nn.3961

Busemeyer, J. R., & Stout, J. C. (2002). A Contribution of Cognitive Decision Models to Clinical Assessment: Decomposing Performance on the Bechara Gambling Task. *Psychological Assessment*, *14*(3), 253–262. https://doi.org/10.1037/1040-3590.14.3.253

Calhoon, G. G., & Tye, K. M. (2015). Resolving the neural circuits of anxiety. *Nature Neuroscience*, *18*(10), 1394–1404. https://doi.org/10.1038/nn.4101

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Charpentier, C. J., Aylward, J., Roiser, J. P., & Robinson, O. J. (2017). Enhanced Risk Aversion, But Not Loss Aversion, in Unmedicated Pathological Anxiety. *Biological Psychiatry*, *81*(12), 1014–1022. https://doi.org/10.1016/j.biopsych.2016.12.010

Collins, A. G. E., & Cockburn, J. (2020). Beyond Dichotomies in Reinforcement Learning. *Nature Reviews Neuroscience*, *21*(10), 576–586. https://doi.org/10.1038/s41583-020-0355-6

Davis, M. (1998). Are different parts of the extended amygdala involved in fear versus anxiety? *Biological Psychiatry*, *44*(12), 1239–1247. https://doi.org/10.1016/s0006-3223(98)00288-1

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Delgado, M., Olsson, A., & Phelps, E. (2006). Extending animal models of fear conditioning to humans. *Biological Psychology*, *73*(1), 39–48. https://doi.org/10.1016/j.biopsycho.2006.01.006

Deserno, L., Wilbertz, T., Reiter, A., Horstmann, A., Neumann, J., Villringer, A., Heinze, H.-J., & Schlagenhauf, F. (2015). Lateral prefrontal model-based signatures are reduced in healthy individuals with high trait impulsivity. *Translational Psychiatry*, *5*(10), e659–e659. https://doi.org/10.1038/tp.2015.139

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Ersche, K. D., Gillan, C. M., Jones, P. S., Williams, G. B., Ward, L. H. E., Luijten, M., Wit, S. d., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2016). Carrots and sticks fail to change behavior in cocaine addiction. *Science*, *352*(6292), 1468–1471. https://doi.org/10.1126/science.aaf3700

Fanselow, M. S. (2000). Contextual fear, gestalt memories, and the hippocampus. *Behavioural Brain Research*, *110*(1-2), 73–81. https://doi.org/10.1016/s0166-4328(99)00186-2

Felix-Ortiz, A. C., Beyeler, A., Seo, C., Leppla, C. A., Wildes, C. P., & Tye, K. M. (2013). BLA to vHPC Inputs Modulate Anxiety-Related Behaviors. *Neuron*, *79*(4), 658–664. https://doi.org/10.1016/j.neuron.2013.06.016

Friston, K. J., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *The*

Lancet Psychiatry, *1*(2), 148–158. https://doi.org/10.1016/s2215-0366(14)70275-5

Garcia, R., Vouimba, R.-M., Baudry, M., & Thompson, R. F. (1999). The amygdala modulates prefrontal cortex activity relative to conditioned fear. *Nature*, *402*(6759), 294–296. https://doi.org/10.1038/46286

Garnefski, N., & Kraaij, V. (2007). The Cognitive Emotion Regulation Questionnaire: Psychometric features and prospective relationships with depression and anxiety in adults. *European Journal of Psychological Assessment*, *23*(3), 141–149.

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Sceince*, *7*(4).

Gillan, C. M., Vaghi, M. M., Hezemans, F. H., Grothe, S. v. G., Dafflon, J., Brühl, A. B., Savulich, G., & Robbins, T. W. (2020). Experimentally induced and real-world anxiety have no demonstrable effect on goal-directed behaviour. *Psychological Medicine*, 1–12. https://doi.org/10.1017/s0033291720000203

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, *5*, e11305. https://doi.org/10.7554/elife.11305

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement

Learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Goodman, W. L., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. II. Validity. *Archives of general psychiatry*, *46*(11), 1012–1016.

Goossens, L., Sunaert, S., Peeters, R., Griez, E. J., & Schruers, K. R. (2007). Amygdala Hyperfunction in Phobic Fear Normalizes After Exposure. *Biological Psychiatry*, *62*(10), 1119–1125. https://doi.org/10.1016/j.biopsych.2007.04.024

Grös, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the state-trait inventory for cognitive and somatic anxiety (sticsa): Comparison to the state-trait anxiety inventory (stai). *Psychological assessment*, *19*(4), 369–381.

Han, D. W., Lee, C. H., & Chon, K. K. (1996). Korean adaptation of spielberger's stai (k-stai). *Korean Journal of Health Psychology*, *1*(1), 1–14.

Hartley, C. A., & Phelps, E. A. (2012). Anxiety and Decision-Making. *Biological Psychiatry*, *72*(2), 113–118. https://doi.org/10.1016/j.biopsych.2011.12.027

Heller, A. S., Ezie, C. C., Otto, A. R., & Timpano, K. R. (2018). Model-based learning and individual differences in depression: The moderating role of stress. *Behaviour Research and Therapy*, *111*, 19–26. https://doi.org/10.1016/j.brat.2018.09.007

Heo, S., Oh, J. Y., & Kim, J. H. (2012). The korean version of the barratt impulsiveness scale, 11th version: Its reliability and validity. *Korean Journal of Psychology: General*, *31*(3), 769–782.

Hirsch, C. R., & Mathews, A. (2012). A cognitive model of pathological worry. *Behaviour Research and Therapy*, *50*(10), 636–646. https://doi.org/10.1016/j.brat.2012.06.007

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, *19*(3), 404–413. https://doi.org/10.1038/nn.4238

Indovina, I., Robbins, T. W., Núñez-Elizalde, A. O., Dunn, B. D., & Bishop, S. J. (2011). Fear-Conditioning Mechanisms Associated with Trait Vulnerability to Anxiety in Humans. *Neuron*, *69*(3), 563–571. https://doi.org/10.1016/j.neuron.2010.12.034

Jellinek, E. M. (1949). Phases in the Drinking History of Alcoholics. *The American Journal of Psychology*, *62*(4), 606. https://doi.org/10.2307/1418573

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Kim, J. W., & Min, B. B. (1998). Intolerance of uncertainty and problem orientation in worry. *The Korean Psychological Association*, *1998*(1), 83–92.

Kim, M. J., Gee, D. G., Loucks, R. A., Davis, F. C., & Whalen, P. J. (2011). Anxiety Dissociates Dorsal and Ventral Medial Prefrontal

Cortex Functional Connectivity with the Amygdala at Rest. *Cerebral Cortex*, *21*(7), 1667–1673. https://doi.org/10.1093/cercor/bhq237

Kim, M.-S., Lee, I. S., & Lee, C. S. (2007). The validation study 1 of korean bdi-2: In female university students sample. *Korean Journal of Clinical Psychology*, *26*(4), 997–1014.

Kim, S. J., Hong, H. J., Lee, H. S., & Kim, C.-H. (2004). Dimensional analyses of the yale-brown obsessive-compulsive scale and yale-brown obsessive-compulsive scale checklist. *Korean Journal of Psychopharmacology*, *15*(3), 339–345.

Kim, S., Hur, J., & Ahn, W.-Y. (prep). Greater distortion of perceived probability among people with high anxiety in both reward and punishment domains: A computational modeling study.

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When Does Model-Based Control Pay Off? *PLOS Computational Biology*, *12*(8), e1005090. https://doi.org/10.1371/journal.pcbi.1005090

Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and stan* (Second). Academi Press.

LeDoux, J. E. (2007). The amygdala. *Current Biology*, *17*(20), R868–R874. https://doi.org/10.1016/j.cub.2007.08.005

LeDoux, J. E. (2012). Rethinking the Emotional Brain. *Neuron*, *73*(4), 653–676. https://doi.org/10.1016/j.neuron.2012.02.004

LeDoux, J. E., & Pine, D. S. (2016). Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework.

American Journal of Psychiatry, *173*(11), 1083–1093. https://doi.org/10.1176/appi.ajp.2016.16030353

Lewandowsky, S., & Farrell, S. (2011). *Computational Modeling in Cognition: Principles and Practice*. Sage Publications.

Liebowitz, M. R. (1987). Social phobia. *Modern Problems in Pharmacopsychiatry*, 141–173.

Loosen, A. M., & Hauser, T. U. (2020). Towards a computational psychiatry of juvenile obsessive-compulsive disorder. *Neuroscience & Biobehavioral Reviews*, *118*, 631–642. https://doi.org/10.1016/j.neubiorev.2020.07.021

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences*, *42*(4), 665–675. https://doi.org/10.1016/j.paid.2006.08.016

Maner, J. K., & Schmidt, N. B. (2006). The Role of Risk Avoidance in Anxiety. *Behavior Therapy*, *37*(2), 181–189. https://doi.org/10.1016/j.beth.2005.11.003

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of Penn state worry questionnaire. *Behavior research and therapy*, *28*(6), 487–495.

Milad, M. R., Quirk, G. J., Pitman, R. K., Orr, S. P., Fischl, B., & Rauch, S. L. (2007). A Role for the Human Dorsal Anterior Cingulate Cortex in Fear Expression. *Biological Psychiatry*, *62*(10), 1191–1194. https://doi.org/10.1016/j.biopsych.2007.04.032

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. *Nature Neuroscience*, *20*(9), 1269–1276. https://doi.org/10.1038/nn.4613

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits Without Values. *Psychological Review*, *126*(2), 292–311. https://doi.org/10.1037/rev0000120

Mineka, S., & Zinbarg, R. (2006). A Contemporary Learning Theory Perspective on the Etiology of Anxiety Disorders. *American Psychologist*, *61*(1), 10–26. https://doi.org/10.1037/0003-066x.61.1.10

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

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., Jones, P. B., Bullmore, E. T., Robbins, T. W., & Fletcher, P. C. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, *13*(3), 267–276. https://doi.org/10.1038/sj.mp.4002058

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

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*, *110*(52), 20941–20946. https://doi.org/10.1073/pnas.1312011110

Park, S.-J., Choi, H.-R., Choi, J.-H., Kim, K., & Hong, J. P. (2010). Reliability and validity of the korean version of the patient health questionnaire-9 (phq-9). *Korean Academy of Anxiety and Mood*, *6*(2), 119–124.

Park, S. Y. (2003). The difference of social anxiety subtypes according to the characteristics of social anxiety-provoking situations (master' s thesis). *Seoul National University, Seoul, Korea.*

Patton, J., Stanford, M., & Barratt, E. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*, 768–74. https://doi.org/10.1002/1097-4679(199511)51:63.0.CO;2-1

Paulus, M. P. (2019). Driven by Pain not Gain: Computational Approaches to Aversion-Related Decision-Making In Psychiatry. *Biological Psychiatry*, *87*(4), 359–367. https://doi.org/10.1016/j.biopsych.2019.08.025

Phan, K. L., Fitzgerald, D. A., Nathan, P. J., & Tancer, M. E. (2006). Association between Amygdala Hyperactivity to Harsh Faces and Severity of Social Anxiety in Generalized Social Phobia. *Biological Psychiatry*, *59*(5), 424–429. https://doi.org/10.1016/j.biopsych.2005.08.012

Rummery, G. A., & Niranjan, M. (1994). *On-line q-learning using connectionist systems* (tech. rep.).

S., T. (1995). Assessment of obsessions and compulsions: reliability, validity, and sensitivity to treatment effects. *Clinical Psychology Review*, *15*, 261–296.

Schouten, R. W., Nadort, E., Harmse, V., Honig, A., Ballegooijen, W. v., Broekman, B. F., & Siegert, C. E. (2020). Symptom dimensions of anxiety and their association with mortality, hospitalization and quality of life in dialysis patients. *Journal of Psychosomatic Research*, *133*, 109995. https://doi.org/10.1016/j.jpsychores.2020.109995

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Sebold, M., Garbusow, M., Jetzschmann, P., Schad, D. J., Nebe, S., Schlagenhauf, F., Heinz, A., Rapp, M., & Romanczuk-Seiferth, N. (2019). Reward and avoidance learning in the context of aversive environments and possible implications for depressive symptoms. *Psychopharmacology*, *236*(8), 2437–2449. https://doi.org/10.1007/s00213-019-05299-9

Sjoerds, Z., Wit, S. d., Brink, W. v. d., Robbins, T. W., Beekman, A. T. F., Penninx, B. W. J. H., & Veltman, D. J. (2013). Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Translational*

Psychiatry, *3*(12), e337–e337.
https://doi.org/10.1038/tp.2013.107

Spielberger, C. (1983). Manual for the State-Trait Anxiety Inventory (STAI) (Form Y).

Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*, *282*(18), 1737–1744.

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

Sylvers, P., Lilienfeld, S. O., & LaPrairie, J. L. (2011). Differences between trait fear and trait anxiety: Implications for psychopathology. *Clinical Psychology Review*, *31*(1), 122–137. https://doi.org/10.1016/j.cpr.2010.08.004

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nature Reviews Neuroscience*, *16*(6), 317–331. https://doi.org/10.1038/nrn3945

Voon, V., Baek, K., Enander, J., Worbe, Y., Morris, L. S., Harrison, N. A., Robbins, T. W., Rück, C., & Daw, N. (2015). Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Translational Psychiatry*, *5*(11), e670–e670. https://doi.org/10.1038/tp.2015.165

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W., Harrison, N. A., Wood, J., Daw, N. D., Dayan, P., Grant, J. E., & Bullmore, E. T. (2015). Disorders of compulsivity: a common bias towards learning habits. *Molecular Psychiatry*, *20*(3), 345–352. https://doi.org/10.1038/mp.2014.44

Voon, V., Reiter, A., Sebold, M., & Groman, S. (2017). Model-Based Control in Dimensional Psychiatry. *Biological Psychiatry*, *82*(6), 391–400. https://doi.org/10.1016/j.biopsych.2017.04.006

Wise, T., & Dolan, R. J. (2020). Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. *Nature Communications*, *11*(1), 4179. https://doi.org/10.1038/s41467-020-17977-w

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine Enhances Model-Based over Model-Free Choice Behavior. *Neuron*, *75*(3), 418–424. https://doi.org/10.1016/j.neuron.2012.03.042

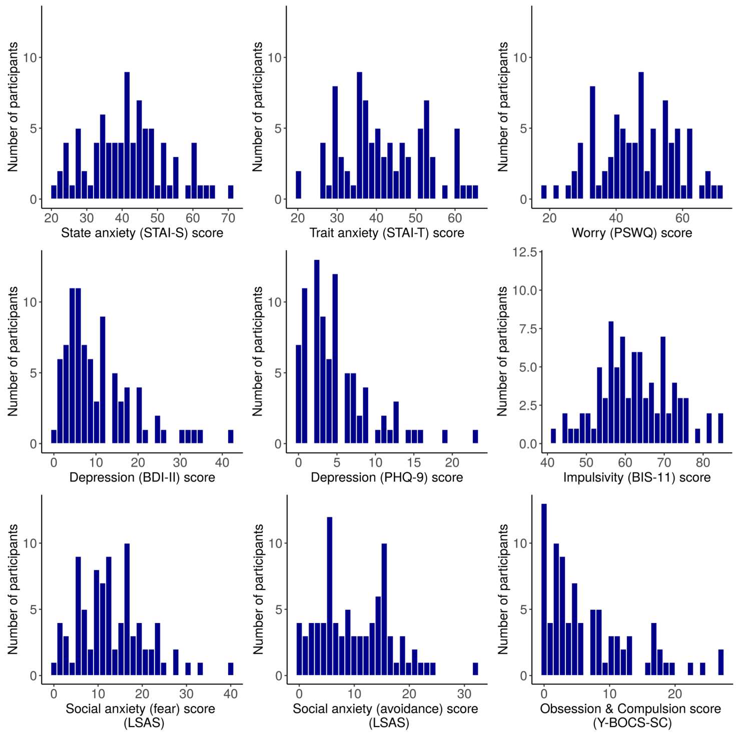
Zebb, B. J., & Beck, J. G. (1998). Worry Versus Anxiety. *Behavior Modification*, *22*(1), 45–61. https://doi.org/10.1177/01454455980221003

Appendix A

Survey measures

Figure A1

Distributions of all survey scores



Appendix B

Supplementary figures

Figure B1

Correlation between survey score and model-based weight

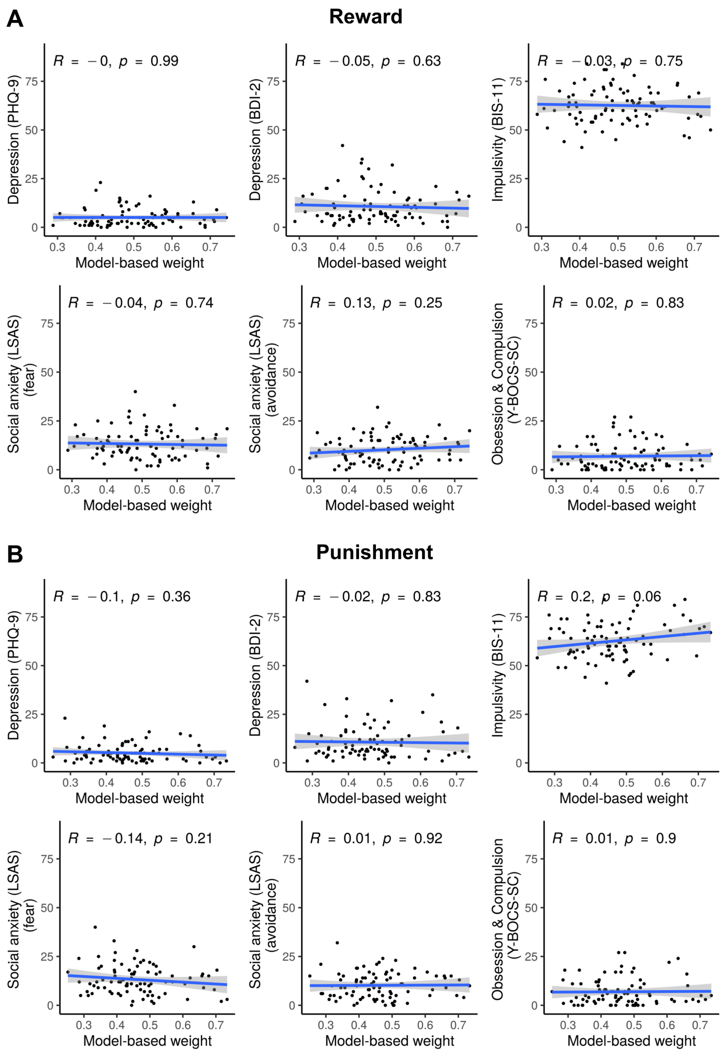
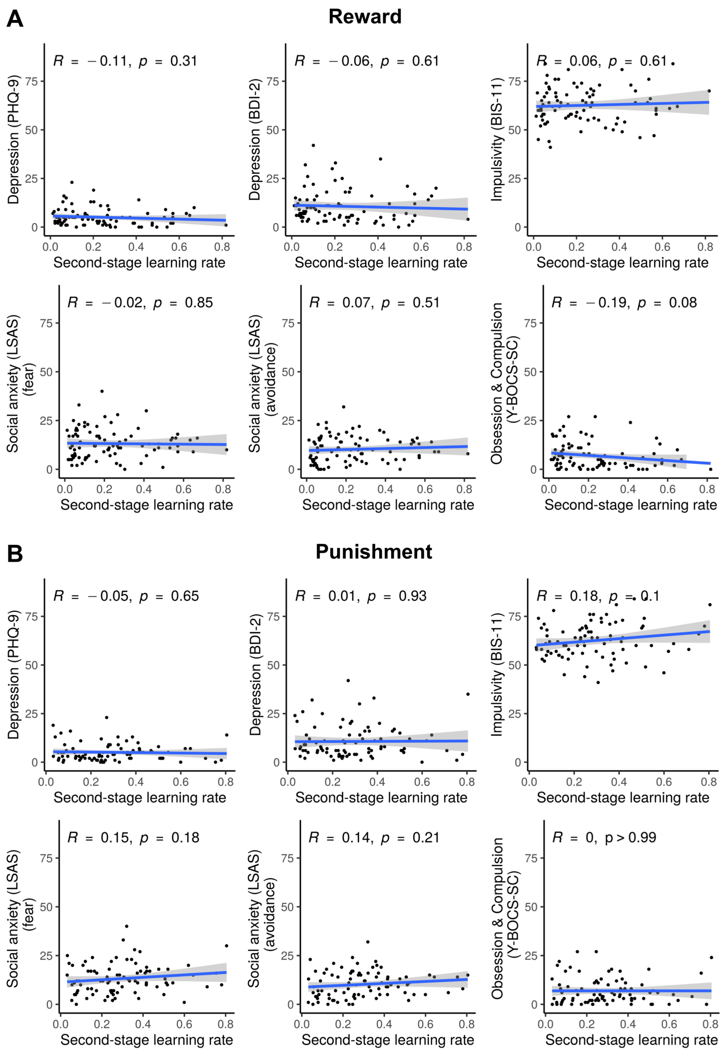


Figure B2

Correlation between survey score and second-stage learning rate



국문초록

불안한 사람들의 의사결정과 학습은 가변적이다. 선행 연구에 따르면 불안 수준이 높은 집단에게서 일정하지 않은 학습 패턴이 관찰되었다. 예를 들어, 처벌 상황에서는 불안 수준이 높을수록 적응적인 학습 지표가 낮아졌던 반면, 중립적이거나 보상 상황에서는 그러한 상관이 유의하지 않았다. 최근 심리학자들은 불안의 의사결정과 학습을 모델프리 (model-free) 그리고 모델기반(model-based) 시스템을 기반으로 연구해왔다. 모델프리 시스템은 보상이 주어졌던 선택을 단순히 강화함으로써 습관적인 행동에 관여하는 반면, 모델기반 시스템은 과제의 구조를 내면화함으로써 보다 목표 지향적인 행동에 기여한다. 선행 연구에서 금전적 보상이 주어졌던 조건에서는 불안 수준과 모델기반 학습 결함이 유의한 상관을 보이지 않았다. 그러나 처벌 조건에서도 불안 수준이 높은 집단에서 모델기반 학습 결함을 보이지 않는지는 아직 충분히 연구되지 않았다. 따라서 본 연구에서는 상태 불안, 특성 불안, 그리고 걱정 수준이 각각 보상과 처벌 조건에서 모델기반 학습 지표와 어떠한 관계를 갖는지 탐구하였다. 이를 위해 비임상군 성인을 모집한 뒤 보상과 처벌 조건에서 같은 의사결정 과제를 실시하도록 하였다. 계산모델링을 통해 개인의 모델기반 학습 지표를 추정한 후 각 불안 지표와 모델기반 학습의 상관이 조건별로 유의하게 달라지는지를 분석하였다. 그 결과 처벌 조건에서는 걱정 수준이 모델기반 학습 지표와 유의한 부적 상관을 보인 반면, 보상 조건에서는 그러한 상관을 보이지 않았다. 또한 걱정 수준이 높은 집단에서만

해결중심사고(refocus on planning)를 더 많이 사용할수록 모델기반 학습 지표가 유의하게 낮아졌다. 본 연구는 불안과 모델기반 학습의 관계를 처벌 조건에서 규명했을 뿐만 아니라, 걱정 수준이 높고 특정 정서조절 전략 사용을 지향할수록 처벌 조건에서 낮은 모델기반 학습 능력을 보일 수 있음을 제시하였다.

주요어: 불안, 계산모델링, 강화학습, 모델기반 의사결정

학번: 2019-21129

감사의 글

석사 과정 및 논문 작성을 잘 마무리할 수 있도록 도와주신 많은 분들께 감사의 인사를 드립니다.

먼저, 2년 8개월의 시간 동안 저를 아낌없이 지도해 주신 안우영 교수님께 진심으로 감사의 말씀을 드립니다. 학부 졸업 후 임상심리학이라는 분야에 막연한 관심이 생겨 찾아뵈었을 때 공부해 볼 수 있도록 기회를 주셔서 감사드립니다. 남들보다 조금은 늦게 걷게 된 연구의 길이었지만 교수님의 지도와 격려 덕분에 지난 석사 과정 동안 지치지 않고 꾸준히 성장할 수 있었습니다. 불확실함으로 가득 찬 학자로서의 여정을 버텨내 오신 선배님이자 저의 지도 교수님이신 안우영 교수님을 만나 제 향후 여정이 더욱 기대가 되고 설렙니다. 인생사 새옹지마이기에 일희일비하지 말고 묵묵히 본인을 믿고 나아가라는 교수님의 말씀대로 학자로서의 큰 목표를 위해 한 발짝씩 자신감 가지고 정진해 나가겠습니다.

또한 바쁘신 와중에도 논문 심사를 맡아주신 김청택, 이훈진 교수님께 진심으로 감사드립니다. 논문 프러포절부터 마지막 심사까지 도움이 되는 질문들과 첨언 주신 덕분에 제 연구가 더욱 발전할 수 있었습니다. 그뿐만 아니라 석사 과정 동안 교수님들의 수업을 들으며 통계와 임상심리학 이론들에 대해 깊게 다루고 학습해보는 기회를 가질 수 있었습니다. 교수님들로부터 심사와 가르침을 받는 것만으로도 학생으로서 큰 성장을 이룰 수 있었습니다. 두 교수님께 다시 한번 감사드립니다.

석사 과정 동안 저와 함께 고민하고 나아갔던 서울대학교 계산임상연구실 동료분들께도 진심으로 감사드립니다. 우선 석사 과정을 함께 하는 것 그 자체만으로도 큰 힘이 되어주셨던 저의 석사 동기, 권민아 선생님께 감사드립니다. 또한 2019년 1월, 같은 날짜에 연구실 처음 들어왔을 때부터 제 석사 졸업까지 연구 관련된 고민들이나 토론을 함께해 주셨던 도호영 선생님께 감사드립니다. 연구실 적응과 더불어 석사 연구 마무리까지 큰 도움 주셨던 김현진 선생님과 양재영 선생님께도 진심으로 감사드립니다. 더불어 제 석사 연구 실험을 도와주신 학부생 연구원 김재연 선생님과 박규리 선생님께도 감사드립니다. 제 석사 연구의 가설이나 설계, 분석 방법 등에 대해 함께 토론하고 고민해 주신 강상훈, 김소연, 박희선, 이정현, 그리고 조윤서 선생님 등을 포함한 모든 계산임상연구실 선생님들께 감사드립니다.

마지막으로 저를 항상 믿고 지지해 주시는 저의 아버지와 어머니께 진심으로 감사드립니다. 즐거움과 동시에 어려움도 많았던 석사 과정 동안 한결같이 제 곁에서 응원해 주시고 저를 저보다 더 믿어주신 부모님 덕분에 석사 과정을 잘 마무리 할 수 있었습니다. '하고 싶은 것 다 해라'라고 말씀해 주시며 늘 아무 걱정 없이 공부에 전념할 수 있도록 도와주신 아버지, 언제나 제 말에 귀 기울여 주시고 무조건적인 사랑을 베풀어주신 어머니께 다시 한번 진심으로 감사드립니다.

그 밖에도 제가 석사 과정을 잘 마칠 수 있도록 곁에서 혹은 멀리서 응원해 주신 모든 분들께 감사드립니다. 그 응원 잊지 않고 앞으로도 열심히 연구하여 임상심리학 분야의 발전을 위해 노력하는 임상심리학자로서 성장해 나가겠습니다. 감사합니다.