國立臺灣大學理學院心理學研究所

碩士論文

Graduate Institute of Psychology

College of Science

National Taiwan University

Master Thesis

化約論者的瘋狂—以貝氏取向探討 AKT1 在精神分裂症多巴胺假說中的角色

Reductionist's madness — a Bayesian approach to investigate the possible role of AKT1 in the dopamine hypothesis of schizophrenia

陳耀主

Yao-Chu Chen

指導教授：賴文崧 博士

Advisor: Wen-Sung Lai, Ph.D.

中華民國 99 年 6 月

June, 2010

Acknowledgement

This thesis is dedicated to John Chen and Janet Lo,

for their unconditional love and understanding.

摘要

本研究旨在探討精神分裂症候選基因 *Akt1* 對增強學習模型中酬賞預測誤差之影響，並進一步探討 AKT1 在精神分裂症多巴胺假說中所扮演的角色。在近幾年的研究中，AKT 被發現是多巴胺 D2 受體下游細胞內訊息傳遞因子之一，且與抗精神病藥物之藥效作用有密切關連。同時由於多巴胺假說認為紋狀體中過剩的多巴胺僅與精神分裂症中的正性症狀有關，且治療正性症狀的抗精神病藥物主要作用在 D2 受體之上。因此透過這些機制 AKT1 極可能參與正性症狀之產生或調節。近期一些關於精神症(即正性症狀之總稱)的研究嘗試採取貝氏推論之觀點，認為精神症之生成與多巴胺系統所產生不正常之酬賞預測誤訊號有關。奠基於這些理論及發現，本研究的假設為：若 *Akt1* 涉入精神症的生成，則 *Akt1* 缺損之小鼠的酬賞預測誤強度會與正常控制組小鼠有所不同。本研究採用 *Akt1* 異型合子之小鼠在動態搜索 T 形迷津之行為資料推算增強學習模型之參數，由此推估酬賞預測誤訊號之強度。實驗一結果顯示在所有三個測試狀況下 *Akt1* 異型合子之小鼠較正常控制組有更強的酬賞預測誤訊號。進一步透過西方墨點法分析腹側及背側紋狀體 AKT1 的磷酸化程度，異型合子小鼠較無法反應甲基安非他命所引發的磷酸化，顯示 AKT1 的活性與多巴胺有關。實驗二如同較強的酬賞預測誤所預期的，此小鼠在其他酬賞關聯學習作業的習得階段中有較迅速的學習表現。進一步以線性擬合法分析配合律，也發現 *Akt1* 異型合子之小鼠對於酬賞有較高的敏感度。最後，實驗三隱性抑制作業證明由實驗一所得較高酬賞預測誤之結果並無法預測以嫌惡學習為基礎的行為結果。本研究的結果顯示，AKT1 可能透過調控腦中多巴胺系統的酬賞預測誤強度參與精神分裂症正性症狀之產生。

關鍵詞：精神分裂症、*Akt1*、多巴胺、酬賞預測誤差、貝氏推論。

Reductionist's Madness — A Bayesian Approach to Investigate the Possible Role of AKT1 in the Dopamine Hypothesis of Schizophrenia

Yao-Chu Chen

Abstract

The current research aims to investigate how the schizophrenia candidate gene *Akt1* (protein kinase Bα) participates in functions of the dopamine system, and further to identify its role in the dopamine hypothesis of schizophrenia. According to recent findings, AKT is found to be a downstream regulator under dopamine D2 receptor, and participates in antipsychotics remedy by acting as an intermediate in the antipsychotics-induced signaling cascade. The dopamine hypothesis of schizophrenia emphasized that psychosis is resulted from excessive dopamine concentration in the striatum, furthermore, antipsychotics mitigates psychosis by acting on dopamine D2 receptors. Recent theory, based on the view that brain is a Bayesian inference machine, regards psychosis is related to disruptions in the reward prediction error (RPE) signal produced by the midbrain dopamine system. Based on these findings, we hypothesized that if AKT1 is involved in the pathogenesis of psychosis, the RPE signal should be different in magnitude between *Akt1* deficient and normal mice. In experiment 1, we estimated parameters in the reinforcement learning model by utilizing the behavioral data collected from a dynamic foraging T maze task perform

by male *Akt1* heterozygous (HET) and wildtype (WT) mice in order to infer the reward prediction error magnitude. The results showed that, compared with WT littermates, *Akt1* HET exhibits higher reward prediction error magnitude among all three testing sections. In experiment2, consistent with the prediction of higher RPE magnitude, *Akt1* HET mice learned more rapidly than WT mice in reward-related tasks. Revealed by the Western blots analysis, a reduction of methamphetamine-induced phosphorylated AKT1 was found in the ventral and dorsal striatum of *Akt1* HET mice but not in WT controls, indicating the activity of AKT1 is indeed related to dopamine. Matching law analysis further revealed that *Akt1* HET mice have higher reward sensitivity compared with WT controls. Finally, in experiment 3, revealed by latent inhibition paradigm, we showed that higher RPE signal cannot predict performance in an aversive-based behavioral paradigm. Our study suggests that, AKT1 might participate in the pathogenesis of psychosis by regulating the RPE magnitude in the dopamine system.

Keywords: schizophrenia, *Akt1*, dopamine, reward prediction error, Bayesian inference.

Table of Contents

Tables and Figures

Chapter 1

General Introduction

"*Though this be madness, yet there is method in't.*"

– William Shakespeare, "Hamlet", Act 2 scene 2

I. Overview

This thesis, in a more restrictive sense, is aimed to investigate the role of a schizophrenia susceptibility gene, *Akt1*, in reinforcement learning model. In a broader sense, it attempts to illustrate how *Akt1* participates in the pathogenesis, at least in parts, in schizophrenia. This is saying that our ultimate concern is to bridge the most fundamental element of an organism to higher level mental function.

This goal is approached by implicating animal model, namely *Akt1* heterozygous (HET) mice which is insufficient in protein AKT1. It is a reductionistic approach to study schizophrenia because we investigate this complex illness by focusing on one particular aspect of schizophrenia—psychosis, and we sought to explain this disease by the abnormality of the most fundamental elements of organism—the genes.

This chapter concerns the development of the concept. We'll discuss the genetic factor of schizophrenia, and the discovery of *Akt1* as a candidate gene. Then the connection between *Akt1* and dopamine (DA) will be addressed, especially its interaction with DA D2 receptor (DRD2). This then leads to our attempt to associate

Akt1 to the DA hypothesis of schizophrenia. Finally, we try to put all these into the context of reward prediction error (RPE) produced by the dopaminergic system, malfunction of which is proposed to be the biological nature of delusion and hallucination by Bayesian view of psychosis.

Chapter 2 provides an empirical study showing that *Akt1* HET mice did exhibit different magnitude of RPE signal implied by changes in estimated parameter in reinforcement learning (RL) model. The last chapter is a general discussion. The idea that *Akt1* participates in the pathogenesis of schizophrenia by aiding the development of psychosis is evaluated.

II. Schizophrenia

Schizophrenia is a psychiatric disorder characterized by disturbances in thought, emotion, disordered thinking; faulty perception and attention; inappropriate or flat affect; and bizarre disturbances in motor activity. People who were diagnosed as schizophrenia have extensive range of problems, although patients typically have only some of these problems at any given time. Unlike most of the diagnostic categories of other diseases, no essential symptom must be presented for a diagnosis of schizophrenia. Thus the heterogeneity within schizophrenic patients is larger than do patients with other disorders. The heterogeneity of schizophrenia implies that it may

be appropriate to subdivide patients into different subtypes that manifest particular constellations of problems. Although there are studies proposed that symptoms of schizophrenia should be divided into five categories (Lindenmayer, Grochowski, & Hyman, 1995), the most common classifications remain: positive, negative, and disorganized. Positive symptoms comprise excesses and distortions, such as hallucinations and delusion. They are defined, for the most part, as the acute episode of schizophrenia. The negative symptoms of schizophrenia consist of behavioral deficits, such as avolition, alogia, anhedonia, flat affect, and asociality. Disorganized symptoms include disorganized speech, often characterized by incoherence and loose association in conversation, and bizarre behavior.

III. The genetics of schizophrenia

i. Schizophrenia as a genetic disorder

Ever since the classical case report of the Genain quadruplets (Rosenthal, 1964), evidence shown unequivocally that schizophrenia is predominantly a genetic disorder. Studies indicate that the heritability of schizophrenia is around 80% (Cardno & Gottesman, 2000; Sullivan, Kendler, & Neale, 2003). Therefore, identification of genes that are responsible for this high heritability will be critical to the understanding of this disorder. Genome-wide linkage studies have revealed probable loci for

susceptibility genes, for example, 1q21-22, 6p24-2, 8p12-21, 13q32-34, and 22q11-13 (Lewis, et al., 2003; Owen, Williams, & O'Donovan, 2003). However, research also indicates that finding a major gene responsible for the disorder is very unlikely (Crow, 2007; Sullivan, 2008). Instead, the additive and/or multiplicative effects of multiple risk genes may provide a better explanatory fit (Gottesman & Shields, 1967). To date, several hundreds of genes have been shown to be associated with schizophrenia (Schwab & Wildenauer, 2009), how these genes, or in what way they contribute to the pathology of schizophrenia have become a serious challenge.

ii. *Akt1*— a schizophrenia candidate gene

In 2004, Emamian and colleagues first identified that the genetic variants of *Akt1* are associated with schizophrenia (Emamian, Hall, Birnbaum, Karayiorgou, & Gogos, 2004). Further evidence from European sib-pair families (Schwab, et al., 2005), Irish families (Thiselton, et al., 2008), the Iranian population (Bajestan, et al., 2006), the Japanese population (Ikeda, et al., 2004; Ikeda, et al., 2006), the Chinese population (Xu, et al., 2007), and the British population (Mathur, Law, Megson, Shaw, & Wei, 2010) have also supported the discovery. Emamian and colleagues conducted a combination of experiments, in addition to association study, to implicate *Akt1* as a susceptibility gene. Their study started with the general concept that kinases and phosphatase are functional candidate genes; they quantified the profusion amount of

several such proteins in lymphocytes from schizophrenic patients and matched controls. After discovering AKT1 protein is consistently reduced in patients with schizophrenia, they demonstrated that such phenomenon was also evident in the hippocampus and frontal cortex, and was accompanied with decreased phosphorylation of glycogen synthase kinase 3 (GSK-3), a downstream target of AKT1 and a molecule that is also of interest in schizophrenia research (Mao, et al., 2009). The genetic association was then found between an *Akt1* haplotype and schizophrenia in 268 affected families, with the risk haplotype being associated with lower AKT1 protein expression in the lymphocytes. Finally, the research found that *Akt1* knockout mice were shown to be more sensitive to amphetamine-induced prepulse inhibition (PPI) disruption.

iii. AKT

The *Akt* gene is the cellular homolog of the v-akt oncogene transduced by AKT8, an acute transforming murine leukemia virus retrovirus (Staal, Hartley, & Rowe, 1977). This gene encodes AKT, a serine-threonine protein kinase (Bellacosa, et al., 1993) also known as protein kinase B (PKB) (Coffer & Woodgett, 1991). AKT plays a central role in regulation of metabolism, cell survival, motility, transcription and cell-cycle progression (Fayard, Tintignac, Baudry, & Hemmings, 2005; Franke, 2008). The AKT subfamily comprises three mammalian isoforms, AKT1, AKT2 and AKT3,

which are products of distinct genes (located at the human chromosome 14q32.32, 19q13.1, and 1q43 respectively), they share a conserved structure that includes three functional domains: an N-terminal pleckstrin homology (PH) domain, a central kinase domain, and a C-terminal regulatory domain containing the hydrophobic motif (HM) phosphorylation site (Hanada, Feng, & Hemmings, 2004).

AKT is a downstream component of phosphoinositide (PI) 3-kinase, which is activated upon autophosphorylation of receptor tyrosine kinases induced by ligands (such as insulin or other growth factors), stimulation of G-protein-coupled receptors, or activation of integrin signaling (Figure 1-1). PI 3-kinase is the key enzyme in the generation of the second messenger PIP3 from PIP2. This results in the translocation of PKB from the cytoplasm to the plasma membrane, which involves its PH domain. Once recruited to the plasma membrane, PKB is activated by a multi-step process that requires phosphorylation of both Thr308 in the activation loop of the kinase domain and Ser473 within the HM of the regulatory domain.

The serine/threonine kinase phosphoinositide-dependent kinase 1 (PDK1) is the kinase responsible for the phosphorylation of Thr308 (Alessi, et al., 1997). Kinases that phosphorylate Ser473 includes AKT autophosphorylation (Toker & Newton, 2000), PDK1 (Balendran, et al., 1999), integrin-linked kinase 1 (ILK1) (Persad, et al., 2001), mitogen activated protein kinase activated protein kinase 2 (MAPKAP-K2)

(Alessi, et al., 1996), protein kinase C ßII (PKCßII) (Kawakami, et al., 2004), and the members of the atypical PI3-kinase related protein kinase (PIKK) family: DNA-dependent protein kinase (DNA-PK) (Feng, Park, Cron, Hess, & Hemmings, 2004), ataxia telangiectasia mutant (ATM) (Viniegra, 2004), and the rapamycin-insensitive mTOR complex TORC2 (Sarbassov, Guertin, Ali, & Sabatini, 2005). Phosphorylation of Ser473 is the key step in the activation of AKT because it stabilizes the active conformation state (Yang, et al., 2002). Once activated at the plasma membrane, phosphorylated AKT can translocate to the cytosol or the nucleus (Andjelkovic, Maira, Cron, Parker, & Hemmings, 1999). The tumor suppressor phosphatase and tensin homology deleted on chromosome ten (PTEN) (Stambolic, et al., 1998) and the SH2-domain-containing inositol polyphosphate 5-phosphatase (SHIP) (Huber, et al., 1999) inhibit AKT activity indirectly by converting PIP3 to PP2. Protein phosphatase 2A (PP2A) and PH domain leucine-rich repeat protein phosphatase (PHLPP) do so directly by dephosphorylating Ser473 and/or Thr308 on PKB (Andjelkovi , et al., 1996; Gao, Furnari, & Newton, 2005).

iv. AKT & Dopamine

In 2004, Beaulieu et al. (Beaulieu, et al., 2004) show that DA transporter (DAT) knockout mice exhibit reduced AKT phosphorylation along with activation of both GSK3α and GSK3β. This is the first evidence to bridge the signaling connection

between DA and AKT, because DAT knockout mice have persistently high level of DA concentration in the brain. Chronic administration of amphetamine, methamphetamine, or apomorphine also lead to decreased AKT phosphorylation (Beaulieu, et al., 2005; Chen, Lao, & Chen, 2006). Further pharmacological experiments applying D1 and D2 antagonist revealed that AKT, GSK3α and GSK3β are regulated by D2-class receptors in DAT knockout mice (Beaulieu, et al., 2004) and haloperidol, a D2 blocker and antipsychotic, has been shown to increase AKT activation and concomitantly inhibit GSK3 in normal animals (Emamian, et al., 2004; Roh, et al., 2007). Research implies DRD2 knockout mice also support the notion that D2 receptors are essential for the inhibition of striatal AKT by dopaminergic drugs (Beaulieu, et al., 2007).

As described earlier, AKT is initially considered a downstream component of PI3 kinase, it also responses to G protein-coupled receptor (GPCR) signaling. Behavioral and biochemical evidence suggests that β-Arrestin 2 (βArr2) is involved in the regulation of the AKT/GSK3 pathway by DRD2 receptors (Beaulieu, et al., 2008; Beaulieu, et al., 2005), which is independent to the canonical cyclic adenosine monophosphate (cAMP) and protein kinase A (PKA) through G protein-dependent signaling.

Once DA D2 receptor was stimulated (Figure 1-2), it leads to receptor

phosphorylation by G protein coupled receptor kinases (GRKs), and subsequently recruit βArr2 to form a signaling complex comprise of PP2A and AKT. PP2A, a phosphatase, then deactivate AKT and initiate GSK3-mediated signaling, since the phosphorylation by AKT is an inhibitory effect to GSK3. In addition to AKT/GSK3 signaling, the recruitment of βArr2 can also result in the formation of an internalization complex comprised of βArr1 and/or βArr2, AP2, clatherin, and other intermediates. Such complex leads to a termination of G protein mediated signaling and to receptor internalization.

IV. Dopamine hypothesis of schizophrenia

i. Dopaminergic system

The mesotelencephalic dopaminergic system arises from three main groups of neurons designated as areas A8, A9 and A10 (Dahlstrom & Fuxe, 1964). The A8 group corresponds to the retrorubral area (RRA), whereas the A9 and A10 groups designate the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA), respectively.

With different innervation areas, the mesotelencephalic dopaminergic system forms three distinct pathways: the mesolimbic pathway, mesocortical pathway, and nigrostriatal pathways (Figure 1-3). The mesolimbic pathway arises from VTA and

terminates at the nucleus accumbens. As indicated by its name, the mesocortical pathway projects DA axons from SNc and VTA to the Isocortex (medial frontal, anterior cingulate, entorhinal, perirhinal) and Allocortex (olfactory bulb, anterior olfactory nucleus, olfactory tubercle, piriform cortex, septal area, and amygdaloid complex). Finally the nigrostriatal pathway is the DA neurons located at the SNc projects to the caudate-putamen and globus pallldus (Moore & Bloom, 1978; Smith & Kieval, 2000).

Schizophrenic patients taking typical antipsychotics, besides the amelioration of psychosis symptoms, often suffer from side effects such as increased negative symptoms, extrapyramidal syndromes, and elevated prolactin level. This is due to the global effect of DRD2 blocking. While antipsychotics acting on mesolimbic pathway can relief psychosis (Anden, 1972; Crow, Deakin, & Longden, 2009), extrapyramidal syndromes were caused by which acting on nigrostriatal pathway (Casey, 1991; Muscettola, Barbato, Pampallona, Casiello, & Bollini, 1999).

ii. The dopamine hypothesis: version I

The theory that schizophrenia is related to DA has evolved concomitantly with the evidence revealed by renovated technologies. The original version of DA hypothesis is developed principally from the knowledge that drugs for the treatment of schizophrenia reduce DA activity.

a. The dopamine hypothesis of antipsychotics action

In 1952, the first antipsychotic chlorpromazine was introduced (Hamon, Paraire, & Velluz, 1952), and since then dozens of antipsychotics have been developed and tested. It was until the 1970s the pharmacological mechanisms of antipsychotics have been confirmed to acting on DRD2. In the 1980s and 1990s the role of DA in psychosis, and that antipsychotics block the DRD2 were firmly established with neuroimaging studies (Farde, Wiesel, Halldin, & Sedvall, 1988; Kapur & Mamo, 2003; Sanger, 2004). Nevertheless, DA receptor in the treatment of psychosis does not by itself constitute proof of the involvement of DA in psychosis (Kapur & Remington, 2001).

b. The dopamine transmission hypothesis of psychosis

Further support for the DA theory of schizophrenia came from the research on psychostimulant psychosis. Psychostimulants can produce a state that closely resembles paranoid schizophrenia (Angrist & Gershon, 1970; Angrist, Sathananthan, Wilk, & Gershon, 1974; Batki, 2000), and cause the worsening of psychotic symptoms in patients with partial remissions (Angrist, Rotrosen, & Gershon, 1980). Furthermore, postmortem studies showed abnormalities in dopaminergic indexes in the brains of schizophrenic patients (Davis, Kahn, Ko, & Davidson, 1991; Seeman, 1987). Perhaps the most compelling evidence is those applying neuroimaging studies.

During the psychosis episode, patients with schizophrenia show a heightened synthesis of DA (Dao-Castellana, et al., 1997; Hietala, et al., 1995; Lindström, et al., 1999; Reith, et al., 1994), and an elevated level of synaptic DA (Abi-Dargham, et al., 2000; Gjedde & Wong, 2001). There are some indications of a change in the number of receptors (Wong, et al., 1997; Wong, et al., 1986), which remain controversial (Farde, et al., 1990; Nordstrom, Farde, Eriksson, & Halldin, 1995).

iii. The dopamine hypothesis: version II

The main progress in version II was drawing the regional specificity into the hypothesis to explain the postmortem and metabolite findings, imaging data, and new insights from animal studies into cortical-subcortical interactions. The fuels that drive the renovations are evidences showing that: (1) in schizophrenic patients, DA metabolites were not universally elevated in the cerebrospinal fluid (CSF) or serum; (2) clozapine has rather low affinity and occupancy for D2 receptors; (3) the postmortem studies of D2 receptors in patients with schizophrenia could not exclude the confounds of previous antipsychotic treatments; and (4) the positron emission tomography (PET) studies of D2/3 receptors in drug-naive patients showed conflicting results.

According to the abovementioned inconsistencies and the emerging evidence that DA receptors have different distributions in the brain, that is, D1 receptors are

predominantly in the cortices and D2 receptors are predominantly in subcortical areas, suggesting that the effects of abnormalities on DA function could vary by brain regions (Davis, et al., 1991). Therefore, the major renovation in version II was the improvement from a one-sided DA hypothesis explaining all aspects of schizophrenia to a regionally specific prefrontal hypodopaminergia and a subcortical hyperdopaminergia.

iv. The dopamine hypothesis version III at a glance

After the second version of DA hypothesis proposed by Davis et al., two decades later, increasing knowledge about DA and schizophrenia unavoidably called up another innovation to this theory. Kapur and his colleague proposed a third version of DA hypothesis of schizophrenia that includes 4 distinct components (Howes & Kapur, 2009). (1) Multiple "hits", including fronto-temporal cortex dysregulation, genes, stress, and drug abuse interact to result in DA dysregulation. (2) Instead of the DRD2 receptor level, the source of DA dysregulation moves to the presynaptic dopaminergic control level. (3) DA dysregulation is linked to "psychosis" rather than schizophrenia. Schizophrenia reflects the nature of the hits coupled with sociocultural factors and not the DA dysfunction itself. And finally, (4) the DA dysregulation results in aberrant salience, that is, an abnormal assignment of salience to stimuli and internal representations, and subsequently develop into psychosis (Kapur, 2003).

Perhaps the most important contribution of version III is that it provides a heuristic framework filling the gap between physiological observed abnormality (subcortical excess DA) and phenomenology (psychosis). Adopting the ideas from incentive salience (Berridge & Robinson, 1998), it was proposed that the central role of DA is to mediate the "salience" of environmental events and internal representations. It is proposed that a dysregulated, hyperdopaminergic state, at a biological level of description and analysis, leads to an aberrant assignment of salience to the elements of one's experience at a mind level. Kapur (2003) addressed that delusions are a cognitive effort by the patient to make sense of these aberrantly salient experiences, whereas hallucinations reflect a direct experience of the aberrant salience of internal representations.

V. Explaining psychosis in a Bayesian approach

In light of the evolution of DA hypothesis of schizophrenia, it appears that knowing “What does dopamine do?” is central to the success of the theory. Hypothesizing the role of DA in schizophrenia can not be successful if the perspective on DA was not precise.

i. The role of dopamine: starting from hedonic hypothesis

Several decades ago, around 1970s and 1980s, it had become clearer that DA

antagonists, primarily antipsychotics, ameliorated many dramatic symptoms of schizophrenia but also resulted in hedonic blunting. This is saying that patients taking antipsychotics appeared to not derive pleasure from stimuli and behavioral acts that should have caused pleasure. Receptor blockers of DA were shown to have a baneful effect on reward learning in laboratory animals. These observations lead to the first hypothesis about the function of DA, the hedonia hypothesis, suggesting DA as the "pleasure neurotransmitter". This was proposed chiefly by Roy Wise and colleagues and such idea had become a very influential view (Wise, Spindler, DeWit, & Gerberg, 1978; Wise, Spindler, & Legault, 1978). According to Wise, the function of DA is to mediate the rewarding or primary motivational characteristics of natural stimuli such as food, water, and sex, as well as those drugs of abuse (Wise, 1982, 2004). The Wise's theory suggested that DA is equal to reward; it is saying that there is equivalence between DA level in the brain and the value of reward. The hedonia hypothesis initiated abundance of research into the effect of antipsychotics on reward-mediated learning, and the results indicated that blocking DA is like removing the reward contingent on an animal's action. A number of more recent reviews have attacked this thesis (Berridge & Robinson, 1998; Ikemoto & Panksepp, 1999; Salamone, Cousins, & Snyder, 1997). These reviewers, however, have proposed refinements that preserved the broad idea that DA is involved in some processes by

which rewards or reward expectations influence behavior.

Early research into the response properties of DA neurons seems to have been driven more by the idea that DA is important for pure motor function. This follows from the most obvious effects of brain DA depletion (as in Parkinson's disease, in which the DA neurons degenerated), which causes gross motor deficits, such as slow movements, difficulty initiating movements, and tremor, paralysis. All of these symptoms are ameliorated by treatment with L-Dopa, a DA precursor which is rendered into DA in the brain. But early attempts, inspired by these findings, to correlate the recorded activity of primate DA neurons with specific motor actions or muscle activations were largely unsuccessful (DeLong, Crutcher, & Georgopoulos, 1983; Schultz, Ruffieux, & Aebischer, 1983).

Instead, the reward hypothesis was proved to be a more useful guide to experimenters. A series of recording studies (Schultz, 1998), revealed that large percentages of DA neurons across both the VTA and SNc respond with a phasic burst of spikes to unexpected primary rewards (such as juice dripped in the mouths of thirsty monkeys) (Figure 1-4). Surprisingly, however, if food delivery was consistently preceded by a tone or light, the dopaminergic response to the reward disappeared after a number of trials. Contrary to the hedonic hypothesis, the lack of measurable dopaminergic response to reward delivery did not accompany extinction,

but rather acquisition. The monkey began showing conditioned responses of anticipatory licking and arm movements to the reward-predictive stimulus. Indeed, not only the monkeys' responses to the tone, but also their dopaminergic neurons began responding to the tone, exhibiting distinct phasic bursts of activity whenever the tone came on. On the other hand, when cued rewards fail to arrive, many DA neurons exhibit a momentary pause in their background firing, timed to the moment reward was expected (Schultz, Dayan, & Montague, 1997).

ii. Reinforcement learning model

In this section, we'll see a very interesting interaction between disciplines—how theories in psychological and computational science together inspired neurophysiologist about the role of dopamine. It appears that the function of DA system is corresponded to the crucial component in learning theories in the psychological field and computational science— an error-correcting term. Finally, we'll address the role of DA in reinforcement learning, the reward prediction error.

The term "reinforcement learning" used in this thesis is not, or should say not only, the well-known concept of stimulus-response learning in psychological literature (Skinner, 1963). As psychologists should know, the reinforcement learning proposed by Skinner focused on how direct associations can be learned between stimuli and responses, neglecting the possible internal states intervene between the stimulus and

its associated response. However, animals are very likely to have covert internal states that affect overt, measurable behavior. Reinforcement learning theory, originally proposed by computer scientist (Barto, Sutton, & Anderson, 1983; Sutton & Barto, 1998) explicitly models such intervening states, and asks how do natural or artificial agents learn to achieve desired states and avoid undesirable ones as efficiently as possible.

a. Rescorla-Wagner model for classical conditioning

The Rescorla-Wagner model (Rescorla & Wagner, 1972), which was developed from the Bush and Mosteller stochastic model of learning (Bush & Mosteller, 1955), postulated that learning occurs only when events violate expectations. For example, in a conditioning experiments in which conditional stimuli CS1 and CS2 (say, a tone and a light) were presented, as well as an affective stimulus such as food or a foot-shock (the unconditional stimuli; US). This model proposed that the associative strength of each of the conditional stimuli $V(CS_i)$ will change according to

$$V_{new}(CS_i) = V_{old}(CS_i) + \eta(CS_i, US) \times \left[\lambda(US) - \sum_i V_{old}(CS_i) \right]$$

(Equation 1, the Rescorla-Wagner model)

In which learning is driven by the mismatch between what was predicted ($\Sigma_i V(CS_i)$, i indicated all the CSs present in the trial) and what actually occurs $\lambda(US)$, which quantifies the maximal associative strength that the unconditional stimulus can

support, and $\eta(CS_i, US)$ is a learning rate that can depend on the significant properties of both the conditional and the unconditional stimuli being associated.

b. Temporal difference learning model

Along with the learning theory developed in psychological field, computational science also quests for algorithm and theories for controlling artificial agents such as robots. It leads to the development of temporal difference (TD) learning. TD learning is a combination of two ideas from reinforcement learning theory, the Monte Carlo idea and the dynamic programming (DP) idea (Sutton, 1988; Sutton & Barto, 1990, 1998; Watkins, 1989). TD methods can learn directly from raw experience without a model of the environment's dynamics like the Monte Carlo method. TD methods update estimates based in part on other learned estimates without waiting for a final outcome as in DP.

The difference between TD and Rescorla-Wagner model is that TD model allows higher-order conditioning and shows sensitivity to the temporal relationships within a learning trial (Sutton & Barto, 1990). In TD learning, the time within a trial is explicitly represented, and learning take place at every time point within a trial, according to

$$V_{new}(S_i,t) = V_{old}(S_i,t) + \eta\left[r(t) + \gamma \sum_{S_k,t+1} V_{old}(S_k,t+1) - \sum_{S_k,t+1} V_{old}(S_j,t) \right]$$

(Equation 2)

According to Equation 2, a stimulus created long-lasting memory traces, and a separate value $V(S,t)$ is learned for every time point of this trace. η is still the learning rate as in Rescorla-Wagner model, so as the learning is driven by the discrepancy between available and expected outcome. However, unlike Rescorla-Wagner model, the associative strength of the stimuli at time t is not only taken to predict the immediately forthcoming reward $r(t)$, but also the future predictions due to those stimuli that will still be used in the next time step $\Sigma_{Sk \text{ at } t+1} V(S,t+1)$ along with the discounting factor $\gamma (0 \le \gamma \le 1)$ these future delayed predictions.

1. The Markov Decision Process

In the computational science version of reinforcement learning, it is important to introduce a concept called Markov decision process (MDP); because it is simplified enough to admit formal analysis but still embodies many features of real-world decision. In MDP, the counterpart to a stimulus in conditioning is a "state". At each time point t, the model environment takes on a discrete state S_t. The S_t follow one another according to some predefined probability distribution $P(s_{t+1} \mid s_t)$, and rewards are observed at each state with probability $P(r_t \mid s_t)$.

2. Temporal difference learning

Given the above setup, the goal of an agent in reinforcement learning is to find an optimal policy that maximizes the expected sum of future rewards. A useful

quantity to predict in such situation is the expected sum of all future rewards:

$$V(s_t) = \mathrm{E}\left[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots \mid s_t \right] = \mathrm{E}\left[\sum_{i=t}^{\infty} \gamma^{i-t} r_i \mid s_t \right]$$

(Equation 3)

Here, the parameter $\gamma (0 \le \gamma \le 1)$ is the discounting factor, which assigns lesser weight on the reward expected farther in the future. The value function can guide the agent's behavior by signaling how well or bad the agent is in based on the prediction of future reward. From Equation 3, it followed directly that

$$V(s_t) = P(r_t \mid s_t) + \gamma \sum_{s_t} P(s_{t+1} \mid s_t) V(s_{t+1})$$

(Equation 4)

Equation 4 holds only if the agent's predictions of the expected discounted sum of the future value are correct. If the values are not correct, there will be a mismatch between the two sides of the equation

$$\delta_t = P(r_t \mid s_t) + \gamma \sum_{s_t} P(s_{t+1} \mid s_t) V(s_{t+1}) - V(s_t)$$

(Equation 5)

Equation 5 represented the temporal-difference prediction error that is to improve the estimates of the function $V(s_t)$, therefore, equation 2 can be reduced to

$$V(S_i)_{new} = V(S_t)_{old} + \eta \cdot \delta_t$$

(Equation 6)

However, in a real-world situation, the dynamic programming updating scheme

requires knowledge of the environment dynamics, that is, $P(r_t \mid s_t)$ and $P(s_{t+1} \mid s_t)$, to compute the prediction error δ_t in equation 5. This prediction is clearly unreasonable. Computational scientist then suggest a "model free" solution (Barto, Sutton, & Watkins, 1989; Werbos, 1977), that is, the environment itself can supply the sufficient information stochastically and incrementally. Which means an agent can sample the reward probabilities in each state to another, as it experiences the task. Updating according to the samples from the environment, it will eventually lead to the correct predictive value. Therefore, the stochastic prediction error can be expressed as

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t),$$

(Equation 7)

Where r_t is the reward observed at time t in state St, and St+1 is the next observed state of the environment. This prediction error can be used as a Monte Carlo approximation to dynamic programming, in order to learn the optimal predictive state value. Finally, combine equation 6 and 7, the TD learning rule can be expressed as

$$V(S_t)_{new} = V(S_t)_{old} + \eta\left[r_t + \gamma V(s_{t+1}) - V(s_t)\right]$$

(Equation 8)

By implementing TD learning rule, a natural or artificial agent can learn the true predictive value of different events in the environment, even when the environment is stochastic and its dynamics are unknown. However, this TD learning only holds in

Pavlovian conditioning, which means the probability of transitioning between different states of situations in the environment are stationary in time.

3. Optimal action selection

The whole purpose of prediction learning is to help selecting actions. In other words, it should help selecting actions in order to obtain as much reward as an agent can. This is the case in "reinforcement learning" (RL) scenario; the environment rewards us for our actions instead for our predictions. To solve this problem, we need to take "action" into the Markov decision process. In MDP, an agent observes the state $s_t \in \{s_1, ..., s_n\}$ and takes an action $a_t \in \{a_1, ..., a_n\}$ according to its policy, which is given stochastically as $P(a \mid s)$. In response to the agent's action a_t, the state of the environment changes stochastically according to a Markov transition matrix $P(s_{t+1} \mid s_t, a_t)$ for each action a_t. The reward $r_{t+1} \in R$ is given stochastically according to $P(r_{t+1} \mid s_t, a_t)$.

A commonly used architecture for optimal action selection is the actor/critic method, which consists of two parts: (1) The critic, which learns to predict future rewards in the form of a state value function $V(s_t)$ for the current policy, and (2) the actor, which improves the policy $P(a \mid s)$ in reference to the future reward predicted by the critic. The TD error δ_t is used as the error signal for the learning of the critic and the reinforcement signal is used for the learning of the actor.

An alternative to Actor/Critic methods for model free RL is to learn explicitly the predictive value (the expected future reward), of taking a specific action at a certain state. Instead of learning the value $V(s_t)$ of each state S_t, Watkins proposed *Q*-learning, a modified TD method in which the agent learns the value $Q(S,a)$ of each state-action pair (*S*,*a*) (Watkins, 1989).

$$Q(S_t,a_t)_{new} = Q(S_t,a_t)_{old} + \eta \cdot \delta_t$$

(Equation 9)

Noted the TD prediction error is slightly different in *Q*-learning

$$\delta_t = r_t + \max_a \gamma Q(s_{t+1},a) - Q(s_t,a_t)$$

(Equation 10)

where the $\max_a$ operator indicates that the TD is computed with respect to what is believe to be the best available action at the subsequent state s_{t+1}. Because *Q*-learning method takes into account the best future action, it is considered an "off-policy" method, even if this will not be the actual action taken at s_{t+1}. There is also an "on-policy" variant SARSA (state-action-reward-state-action), in which the prediction error takes into account the actual chosen action:

$$\delta_t = r_t + \gamma Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)$$

(Equation 11)

In addition to that, in *Q*-learning, action selection is simply taking the

highest $Q(S,a)$ value. However, in a real world scenario, action selection is also stochastically dependent. For a given state *s*, the action value $Q(S,a_i)$ for the candidate action $a_i(i=1,...,m)$ are compared and the one with a higher action value $Q(S,a_i)$ is selected with a higher probability. A common way is so called Boltzmann selection, in which the policy is given by:

$$P(a_i \mid s) = \frac{\exp[\beta Q(s,a_i)]}{\sum_{j=1}^{m} \exp[\beta Q(s,a_j)]}$$

(Equation 12)

Here, the parameter β, which is called the inverse temperature, controls the stochasticity of the policy.

4. Dopamine and temporal difference error

As we have discussed earlier, dopaminergic neurons of the ventral tegmental area and substantia nigra show phasic changes in spike activity that correlate with the history of reward delivery (Bayer & Glimcher, 2005; Hollerman & Schultz, 1998; Schultz, 1998; Schultz, Apicella, & Ljungberg, 1993; Waelti, Dickinson, & Schultz, 2001). This pattern of dopaminergic responding throughout the course of learning conforms exactly to the characteristics of a TD prediction error (Montague, Dayan, & Sejnowski, 1996; Montague, Hyman, & Cohen, 2004; Schultz, et al., 1997), which means they only occur when events are not predicted. Within this framework, it is now interpretable why dopaminergic neurons fire to unexpected rewards but not to

those that are predicted by previous stimuli, and why DA is necessary for reward-mediated learning in the basal ganglia. What's more, the model provides an explanation why, after training, DA neurons did not fire above baseline in the time period between a predictive cue and the reward delivery — in the absence of new information, there are no prediction errors at these intermediate time.

iii. A Bayesian view of psychosis

Taking prediction error into account, Fletcher (Fletcher & Frith, 2009) proposed that disruption in prediction error and its consequences may explain the emergence of psychosis. This idea is grounded in a well-known notion in cognitive psychology — perception is an inference, that is, what we already know influence what we perceive. This echoes to an idea that brain is function as a Bayesian inference machine (Friston, Kilner, & Harrison, 2006; Lee & Mumford, 2003; Summerfield & Koechlin, 2008). For example, first, when the brain is inferring causes from noisy sense data, instead of representing it in "all-or-nothing", it represented in a conditioned probability distribution function. The probability that "given the sense that was the cause" is proportion to "that cause would lead to that sensation" and times the "prior probability of that causes is really out there" (Knill & Pouget, 2004).

$$p(cause \mid sense) \propto p(sense \mid cause) \cdot p(cause)$$

This explains the well-known phenomenon that it is easier to recognize your advisor

in the campus than in your hometown, because your prior belief (no way my advisor would show up here) is against the actual presence of your professor.

The notion that perception is inference was implicated to robustness and efficiency, which means we are able to perceive noisy data very well if we have prior expectations. However, a system would be inflexible if it only saw what was expected to be seen of the world, therefore, the brain needs prediction errors to adapt to the changing world, that is, to update inference. In another words, inferences (prior believes) are updated by perceptions when the sense data are sufficiently improbable. This implies that, experiences that do not challenge a belief system are predicted, and therefore ignore. A well-characterized behavioral paradigm, latent inhibition, captured this idea. Through repeated exposure to a stimulus, one increases the possibility that the stimulus does not predict anything, and therefore decrease the probability that will associates with another stimulus in the future.

Suggested by Fletcher et al. and Baye's theorem, if prediction error persistent and erroneous, inferences would be bizarre, the world does not fit together. What's more, uncertainty is raised, as well as demanding attention, stimuli accompanied by a large prediction error would become more readily associable, perhaps accounting for the bizarre but compelling coincidences that patients with schizophrenia frequently report (Chapman, 1966).Other phenomenon includes: things feel novel, important,

and salient; the distinction between internally- and externally-generated changes may be blurred, e.g., something that I did might not produce the expected sensory consequences, and thus it may not feel like I doing it (Corlett, Frith, & Fletcher, 2009; Fletcher & Frith, 2009). Taken together, a relatively small prediction error might be given undue weight (if the uncertainty is underestimated), leading to a false inference. Alternatively, excessive noise might dilute the effects of even a large prediction-error signal, leading to reluctance to accept an inference as adequately explaining the perceiving input. A noisy prediction-error signal could therefore lead to patients' strange experiences, together with their readiness to accept incidental stimuli and events as important and meaningful and to link them in unusual ways.

Evidence showed drugs that increase DA function in healthy participants strengthened error-dependent reward learning, whereas a DA-blocking drug reduced such learning (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). Patients with schizophrenia also showed disrupted prediction error system, that is, the ventral striatum, the terminal of mesolimbic DA system, where it was shown to be under-active in response to rewarding experiences (Juckel, et al., 2006). Furthermore, abnormal reward-based prediction error activity has also been observed in the ventral striatum in schizophrenic patients undergo psychosis episode (Murray, et al., 2008). These patients had a relatively suppressed response to stimuli that should be relevant

and important and a relatively augmented response to stimuli that should be neutral or unimportant. This observation was confirmed in the same patients using a causal inferential task that led to prediction error (Corlett, et al., 2007). Moreover, the baseline measure of prediction error signals in healthy controls predicts their vulnerability to positive symptoms when subsequently administered ketamine, a psychosis-promoting drug (Corlett, et al., 2006).

VI. Taken together

From the literature reviews, we know that the schizophrenia candidate gene *Akt1* encodes AKT1, an important signaling component, along with its AKT family members carry various function in cellular physiology. In addition, we learn that besides PI3K, AKT also plays a regulatory role downstream to DA D2 receptor, which is the central focus in the DA hypothesis of schizophrenia. While the evolution of the hypothesis shifts its focus from schizophrenia to psychosis as a whole, we pondered whether AKT1 would participate in the development of psychosis.

In the next chapter, we tackled this problem by assessing the RPE magnitude of *Akt1* HET mice along with several other behavioral assessments (Table 1-1). We postulated that, if AKT1 is involved in the pathogenesis of psychosis, according to Bayesian's view, it should somehow affect the RPE magnitude estimated in a

reinforcement learning model. That is, the estimated RPE magnitude should be different in a way between *Akt1* HET mice and WT controls. Chapter 2 comprised three major experiments. In experiment 1, we designed a simple T-maze value-learning task. Mice will learn to choose an arm with higher reward rate by trials and error. The trial-by-trial choice data obtained from the T-maze task will then be fitted into a standard reinforcement learning model to obtain parameter that describe RPE magnitude. The parameter estimation was done by hierarchical Bayesian inference. The major advantage of Bayesian inference is that we can describe the parameter by a probability distribution instead of point estimation. In addition, a probability distribution offers much more information, e.g., the mean, the median, and the standard deviation, than traditional point estimation. Therefore, the difference between two groups can be described in terms of probability instead of statistical significance. In experiment 1, we also concern about whether the choice behavior in the maze will be affected by reduced AKT1 expression, therefore the choice perseveration parameter and matching law analysis were both estimated and conducted in addition. In experiment 2, we ask whether the different RPE observed in *Akt1* HET mice will also change their behaviors that are related to reward learning. We considered both natural reward and drug reward to be comprehensive. In experiment 3, the latent inhibition function was tested in a conditioned emotional

response paradigm. The latent inhibition is distinct from the reward-related behaviors in experiment 2, but is also proposed sensitive to the RPE magnitude. Therefore, we applied this task to further confirm whether the RPE is really different in HET mice. The general prediction about the behavioral task is that if the estimated parameter "learning rate" is large (it represents large RPE magnitude), the speed of updating the value of reward will be fast, and therefore, a subject should learn the reward-related behaviors more rapidly than those possessed lower "learning rate". For latent inhibition, an augmented RPE will decrease the prediction of the probability that a previously exposed stimulus will further associated with a noxious unconditioned stimulus, therefore the inhibition effect should be manifested.

Figure legends

Figure 1-1. Signaling cascade of AKT. This figure represents the canonical PI3K-mediated AKT signal transduction, and the various biological functions, such as angiogenesis, cell survival, metabolism, proliferation, and translation, upon activation. Adapted from Franke et al. (2008)

Figure 1-2. AKT as a downstream regulator of DRD2. (1) When DA activates DRD2, GRKs were phosphorylated and recruit *β*-Arrestin 2. (2) Forming a signaling complex consist of *β*-Arrestin 2, AKT, and PP2A. (3) Recruitment of *β*-Arrestin 2 can also result in DRD2 internalization. Adapted from Beaulieu et al. (2009).

Figure 1-3. DA pathways in the brain, including mesolimbic, mesocortical, nigrostriatal, and tuberoinfundibular pathway. Adapted from (Crocker, 1994)

Figure 1-4. Firing patterns of dopaminergic neurons code for an error in the prediction in monkey performing an analogous instrumental conditioning task. Each raster plot shows action potentials (dots) with different rows representing different trials, aligned on the time of the cue (or the reward). Peri-event time histograms show activity summed over the trials plotted below. (Top) When a reward unexpectedly obtained,

dopaminergic neurons respond with a phasic burst of firing. (Middle) After conditioning with a predictive cue (which, in this task, predicted a food reward if the animal quickly performed the correct response), the reward occurred as predicted, that is, no error in the prediction of reward. Therefore, the reward no longer elicits a burst of activity, and the phasic burst now occurs at the presentation of the predictive cue. (Bottom) When the food reward was unexpectedly omitted, because of a mistake in the behavioral response of the monkey, dopaminergic neurons showed a pause in firing exactly at the time when the reward would have occurred, below their standard background firing rate. This figure is adapted from Schultz (1993).

Table 1-1.
Summary of experiments in chapter 2.

	Purpose	**Experiment**
Exp. 1	Estimating RPE in a value-learning task	Dynamic foraging T-Maze RL model parameters estimation
	Choice behavior analysis	Matching law analysis
Exp. 2	Reward-related behavior—natural reward	Two bottle preference test Incentive licking test
	Reward-related behavior—drug reward	Conditioned place preference
Exp. 3	Latent inhibition	Conditioned emotional response paradigm

Figure 1-1

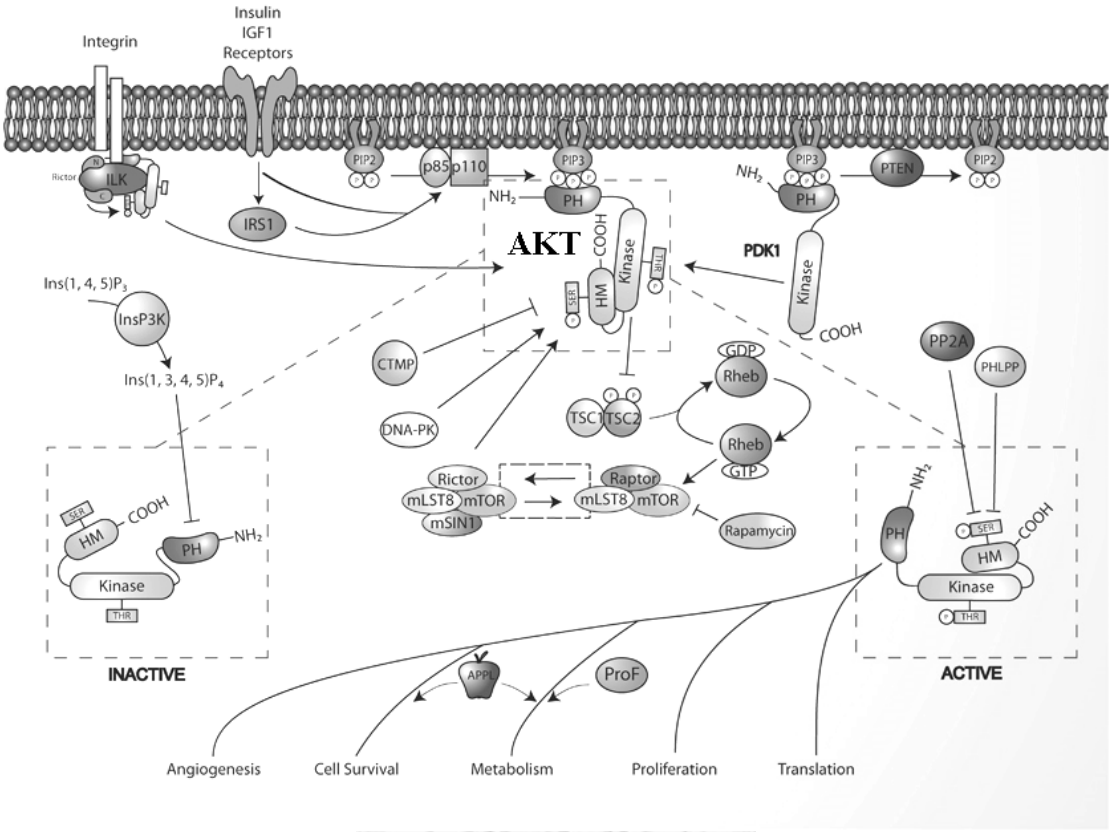

Integrin
Insulin
IGF1
Receptors
Rictor
ILK
IRS1
PIP2
p85
p110
PIP3
NH2
PH
AKT
COOH
Kinase
HM
SER
THR
PDK1
PTEN
Ins(1, 4, 5)P3
InsP3K
Ins(1, 3, 4, 5)P4
CTMP
DNA-PK
GDP
Rheb
GTP
TSC1
TSC2
Rictor
mLST8
mTOR
mSIN1
Raptor
Rapamycin
PP2A
PHLPP
INACTIVE
ACTIVE
APPL
ProF
Angiogenesis
Cell Survival
Metabolism
Proliferation
Translation

Figure 1-2

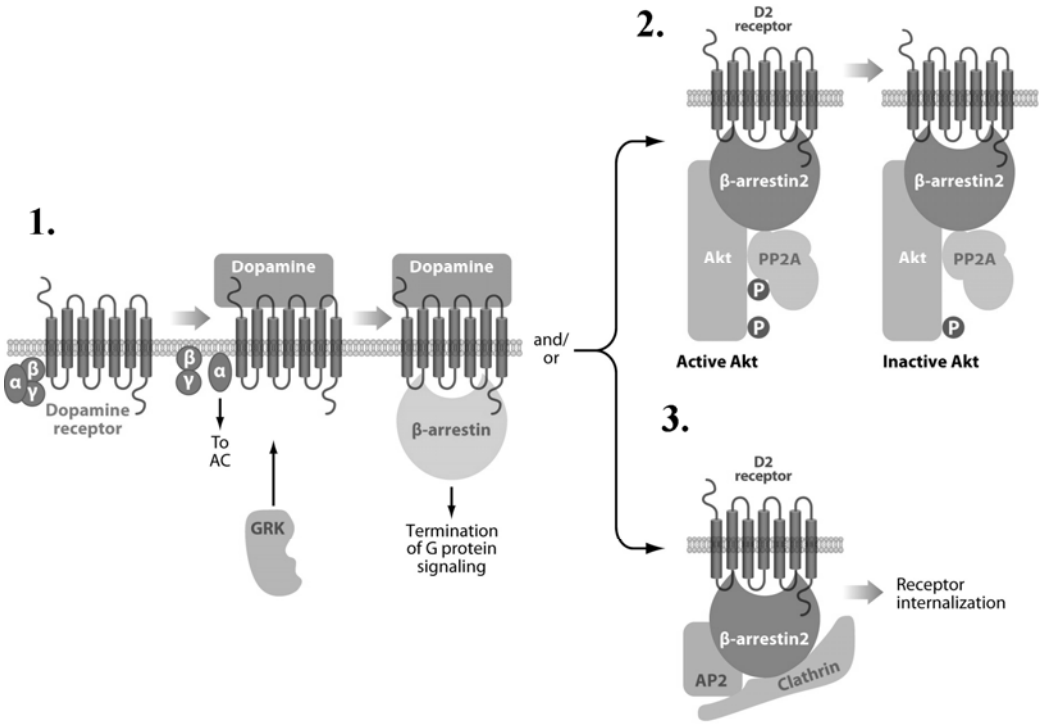

1.
Dopamine
Dopamine
α
β
γ
Dopamine receptor
To AC
GRK
β-arrestin
Termination of G protein signaling
and/ or
2.
D2 receptor
β-arrestin2
Akt
PP2A
P
P
Active Akt
β-arrestin2
Akt
PP2A
P
Inactive Akt
3.
D2 receptor
β-arrestin2
AP2
Clathrin
Receptor internalization

Figure 1-3

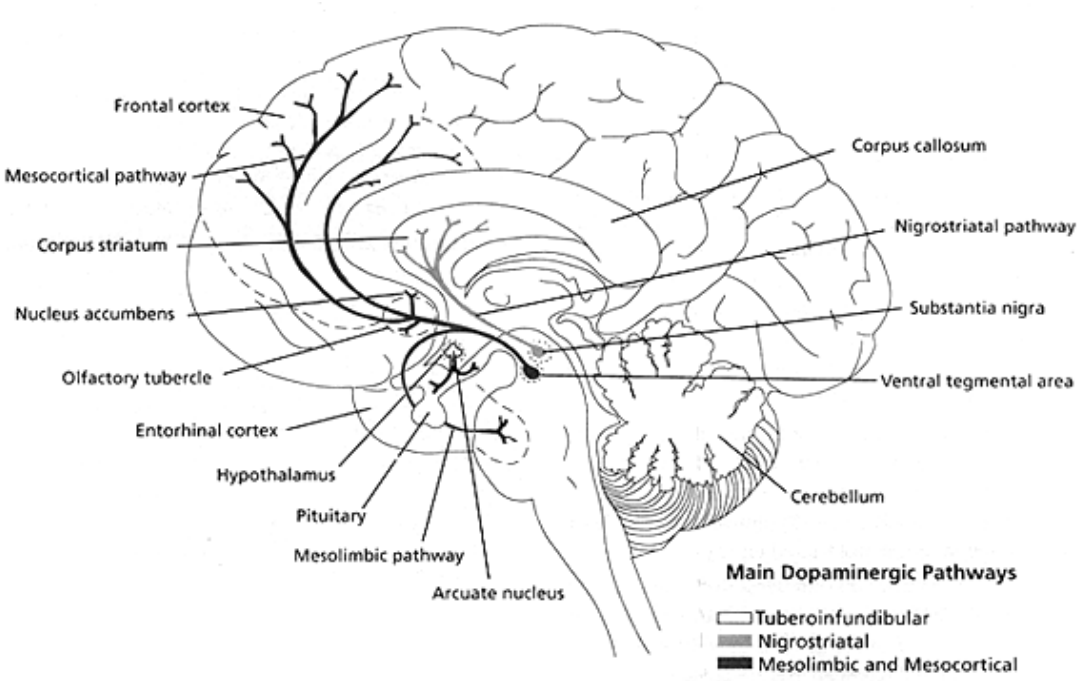
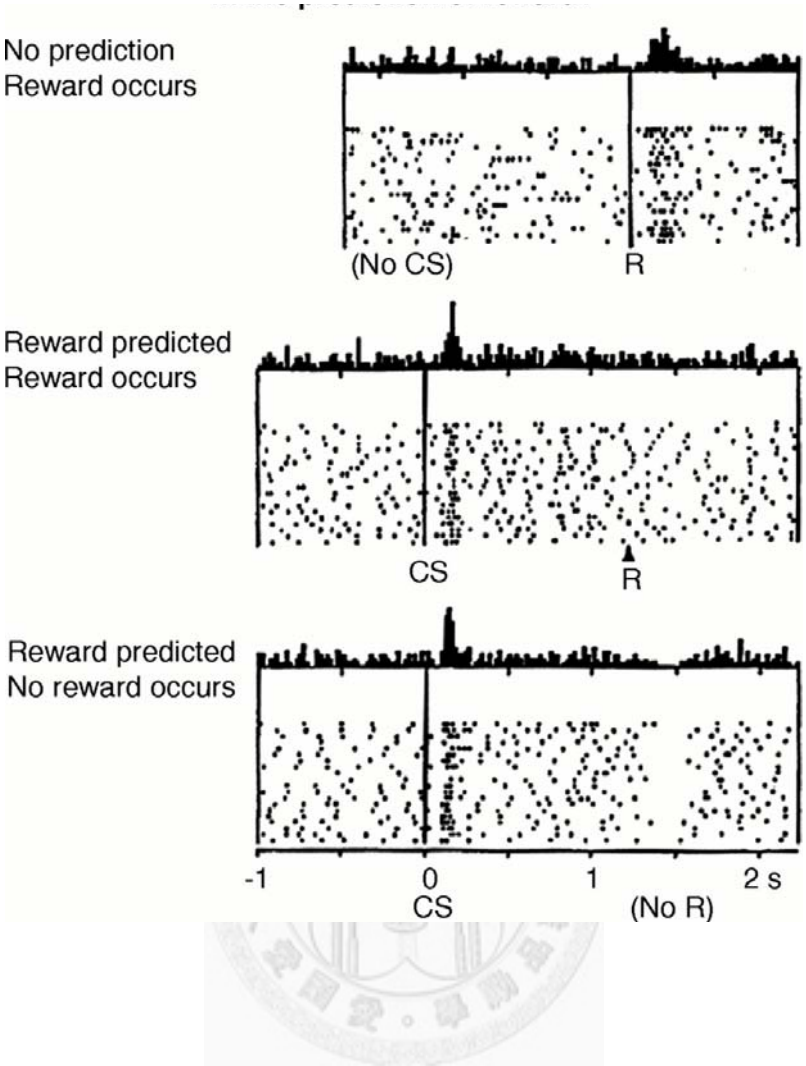

Frontal cortex
Mesocortical pathway
Corpus striatum
Nucleus accumbens
Olfactory tubercle
Entorhinal cortex
Hypothalamus
Pituitary
Mesolimbic pathway
Arcuate nucleus
Corpus callosum
Nigrostriatal pathway
Substantia nigra
Ventral tegmental area
Cerebellum
Main Dopaminergic Pathways
Tuberoinfundibular
Nigrostriatal
Mesolimbic and Mesocortical

Figure 1-4


No prediction
Reward occurs
(No CS)
R
Reward predicted
Reward occurs
CS
R
Reward predicted
No reward occurs
-1
0
1
2 s
CS
(No R)

Reference

Abi-Dargham, A., Rodenhiser, J., Printz, D., Zea-Ponce, Y., Gil, R., Kegeles, L., et al. (2000). Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America, 97*(14), 8104.

Alessi, D., Andjelkovic, M., Caudwell, B., Cron, P., Morrice, N., Cohen, P., et al. (1996). Mechanism of activation of protein kinase B by insulin and IGF-1. *The EMBO Journal, 15*(23), 6541.

Alessi, D., James, S., Downes, C., Holmes, A., Gaffney, P., Reese, C., et al. (1997). Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase B alpha. *Current Biology, 7*(4), 261-269.

Anden, N. (1972). Dopamine turnover in the corpus striatum and the lumbic system after treatment with neuroleptic and anti-acetylcholine drugs. *The Journal of pharmacy and pharmacology, 24*(11), 905.

Andjelkovi , M., Jakubowicz, T., Cron, P., Ming, X., Han, J., & Hemmings, B. (1996). Activation and phosphorylation of a pleckstrin homology domain containing protein kinase (RAC-PK/PKB) promoted by serum and protein phosphatase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America, 93*(12), 5699.

Andjelkovic, M., Maira, S., Cron, P., Parker, P., & Hemmings, B. (1999). Domain swapping used to investigate the mechanism of protein kinase B regulation by 3-phosphoinositide-dependent protein kinase 1 and Ser473 kinase. *Molecular and Cellular Biology, 19*(7), 5061.

Angrist, B., & Gershon, S. (1970). The phenomenology of experimentally induced amphetamine psychosis: Preliminary observations. *Biological Psychiatry, 2*(2), 95-107.

Angrist, B., Rotrosen, J., & Gershon, S. (1980). Responses to apomorphine, amphetamine, and neuroleptics in schizophrenic subjects. *Psychopharmacology, 67*(1), 31-38.

Angrist, B., Sathananthan, G., Wilk, S., & Gershon, S. (1974). Amphetamine psychosis: behavioral and biochemical aspects. *Journal of Psychiatric Research, 11*, 13-23.

Bajestan, S. N., Sabouri, A. H., Nakamura, M., Takashima, H., Keikhaee, M. R., Behdani, F., et al. (2006). Association of AKT1 haplotype with the risk of

schizophrenia in Iranian population. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 141B*(4), 383-386.

Balendran, A., Casamayor, A., Deak, M., Paterson, A., Gaffney, P., Currie, R., et al. (1999). PDK1 acquires PDK2 activity in the presence of a synthetic peptide derived from the carboxyl terminus of PRK2. *Current Biology, 9*(8), 393-404.

Barto, A., Sutton, R., & Anderson, C. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on systems, man, and cybernetics, 13*(5), 834-846.

Barto, A., Sutton, R., & Watkins, C. (1989). Sequential decision problems and neural networks. In D. Touretzky (Ed.), *Advances in Neural Information Processing Systems 2* (pp. 686-693). Cambridge, MA: MIT Press.

Batki, D. (2000). Stimulant psychosis: symptom profile and acute clinical course. *American Journal on Addictions, 9*(1), 28-37.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron, 47*(1), 129-141.

Beaulieu, J. M., Marion, S., Rodriguiz, R. M., Medvedev, I. O., Sotnikova, T. D., Ghisi, V., et al. (2008). A beta-arrestin 2 signaling complex mediates lithium action on Behavior. *Cell, 132*(1), 125-136.

Beaulieu, J. M., Sotnikova, T. D., Marion, S., Lefkowitz, R. J., Gainetdinov, R. R., & Caron, M. G. (2005). An Akt/beta-arrestin 2/PP2A signaling complex mediates dopaminergic neurotransmission and behavior. *Cell, 122*(2), 261-273.

Beaulieu, J. M., Sotnikova, T. D., Yao, W. D., Kockeritz, L., Woodgett, J. R., Gainetdinov, R. R., et al. (2004). Lithium antagonizes dopamine-dependent behaviors mediated by an AKT/glycogen synthase kinase 3 signaling cascade. *Proceedings of the National Academy of Sciences of the United States of America, 101*(14), 5099-5104.

Beaulieu, J. M., Tirotta, E., Sotnikova, T. D., Masri, B., Salahpour, A., Gainetdinov, R. R., et al. (2007). Regulation of Akt signaling by D-2 and D-3 dopamine receptors in vivo. *Journal of Neuroscience, 27*(4), 881-885.

Bellacosa, A., Franke, T., Gonzalez-Portal, M., Datta, K., Taguchi, T., Gardner, J., et al. (1993). Structure, expression and chromosomal mapping of c-akt: relationship to v-akt and its implications. *Oncogene, 8*(3), 745.

Berridge, K., & Robinson, T. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28*(3), 309-369.

Bush, R., & Mosteller, F. (1955). *Stochastic models for learning*. New York: John Wiley and Sons, Inc.

Cardno, A. G., & Gottesman, II. (2000). Twin studies of schizophrenia: From

bow-and-arrow concordances to star wars mx and functional genomics. *American Journal of Medical Genetics, 97*(1), 12-17.

Casey, D. (1991). Neuroleptic drug-induced extrapyramidal syndromes and tardive dyskinesia. *Schizophrenia Research, 4*(2), 109-120.

Chapman, J. (1966). The early symptoms of schizophrenia. *The British Journal of Psychiatry, 112*(484), 225.

Chen, P., Lao, C., & Chen, J. (2006). Dual alteration of limbic dopamine D1 receptor-mediated signalling and the Akt/GSK3 pathway in dopamine D3 receptor mutants during the development of methamphetamine sensitization. *Journal of Neurochemistry, 100*(1), 225-241.

Coffer, P., & Woodgett, J. (1991). Molecular cloning and characterisation of a novel putative protein-serine kinase related to the cAMP-dependent and protein kinase C families. *European Journal of Biochemistry, 201*(2), 475-481.

Corlett, P., Frith, C., & Fletcher, P. (2009). From drugs to deprivation: a Bayesian framework for understanding models of psychosis. *Psychopharmacology, 206*(4), 515-530.

Corlett, P., Honey, G., Aitken, M., Dickinson, A., Shanks, D., Absalom, A., et al. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. *Archives of General Psychiatry, 63*(6), 611.

Corlett, P., Murray, G., Honey, G., Aitken, M., Shanks, D., Robbins, T., et al. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*.

Crocker, A. (1994). Experimental and clinical pharmacology: Dopamine-mechanisms of action. *Australian Prescriber, 17*, 17-17.

Crow, T. (2007). How and why genetic linkage has not solved the problem of psychosis: review and hypothesis. *American Journal of Psychiatry, 164*(1), 13.

Crow, T., Deakin, J., & Longden, A. (2009). The nucleus accumbens possible site of antipsychotic action of neuroleptic drugs? *Psychological Medicine, 7*(02), 213-221.

Dahlstrom, A., & Fuxe, K. (1964). Evidence for the existence of monoamine-containing neurons in the central nervous system. I. Demonstration of monoamines in the cell bodies of brain stem neurons. *Acta physiologica Scandinavica. Supplementum*.

Dao-Castellana, M., Paill re-Martinot, M., Hantraye, P., Attar-Levy, D., Remy, P., Crouzel, C., et al. (1997). Presynaptic dopaminergic function in the striatum of schizophrenic patients. *Schizophrenia Research, 23*(2), 167-174.

Davis, K., Kahn, R., Ko, G., & Davidson, M. (1991). Dopamine in schizophrenia: a

review and reconceptualization. *American Journal of Psychiatry, 148*(11), 1474.

DeLong, M., Crutcher, M., & Georgopoulos, A. (1983). Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *Journal of Neuroscience, 3*(8), 1599.

Emamian, E. S., Hall, D., Birnbaum, M. J., Karayiorgou, M., & Gogos, J. A. (2004). Convergent evidence for impaired AKT1-GSK3 beta signaling in schizophrenia. *Nature Genetics, 36*(2), 131-137.

Farde, L., Wiesel, F., Halldin, C., & Sedvall, G. (1988). Central D2-dopamine receptor occupancy in schizophrenic patients treated with antipsychotic drugs. *Archives of General Psychiatry, 45*(1), 71.

Farde, L., Wiesel, F., Stone-Elander, S., Halldin, C., Nordstrom, A., Hall, H., et al. (1990). D2 dopamine receptors in neuroleptic-naive schizophrenic patients: a positron emission tomography study with [11C] raclopride. *Archives of General Psychiatry, 47*(3), 213.

Fayard, E., Tintignac, L. A., Baudry, A., & Hemmings, B. A. (2005). Protein kinase B/Akt at a glance. *Journal of Cell Science, 118*(24), 5675-5678.

Feng, J., Park, J., Cron, P., Hess, D., & Hemmings, B. (2004). Identification of a PKB/Akt hydrophobic motif Ser-473 kinase as DNA-dependent protein kinase. *Journal of Biological Chemistry, 279*(39), 41189.

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience, 10*(1), 48-58.

Franke, T. (2008). PI3K/Akt: getting it right matters. *Oncogene, 27*(50), 6473-6488.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris, 100*(1-3), 70-87.

Gao, T., Furnari, F., & Newton, A. (2005). PHLPP: a phosphatase that directly dephosphorylates Akt, promotes apoptosis, and suppresses tumor growth. *Molecular Cell, 18*(1), 13-24.

Gjedde, A., & Wong, D. (2001). Quantification of neuroreceptors in living human brain. V. Endogenous neurotransmitter inhibition of haloperidol binding in psychosis. *Journal of Cerebral Blood Flow & Metabolism, 21*(8), 982-994.

Gottesman, I., & Shields, J. (1967). A polygenic theory of schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America, 58*(1), 199-&.

Hamon, J., Paraire, J., & Velluz, J. (1952). *Remarques sur l'action du 4560 RP sur l'agitation maniaque*.

Hanada, M., Feng, J., & Hemmings, B. (2004). Structure, regulation and function of

PKB/AKT--a major therapeutic target. *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics, 1697*(1-2), 3-16.

Hietala, J., Syvalahti, E., Vuorio, K., Rakkolainen, V., Bergman, J., Haaparanta, M., et al. (1995). Presynaptic dopamine function in striatum of neuroleptic-naive schizophrenic patients. *Lancet, 346*(8983), 1130.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience, 1*(4), 304-309.

Howes, O. D., & Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III--The Final Common Pathway. *Schizophrenia Bulletin, 35*(3), 549-562.

Huber, M., Helgason, C., Damen, J., Scheid, M., Duronio, V., Liu, L., et al. (1999). The role of SHIP in growth factor induced signalling. *Progress in Biophysics and Molecular Biology, 71*(3-4), 423-434.

Ikeda, M., Iwata, N., Suzuki, T., Kitajima, T., Yamanouchi, Y., Kinoshita, Y., et al. (2004). Association of AKT1 with schizophrenia confirmed in a Japanese population. *Biological Psychiatry, 56*(9), 698-700.

Ikeda, M., Iwata, N., Suzuki, T., Kitajima, T., Yamanouchi, Y., Kinoshiya, Y., et al. (2006). Positive association of AKT1 haplotype to Japanese methamphetamine use disorder. *International Journal of Neuropsychopharmacology, 9*(1), 77-81.

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research Reviews, 31*(1), 6-41.

Juckel, G., Schlagenhauf, F., Koslowski, M., Wustenberg, T., Villringer, A., Knutson, B., et al. (2006). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage, 29*(2), 409-416.

Kapur, S. (2003). Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. *American Journal of Psychiatry, 160*(1), 13-23.

Kapur, S., & Mamo, D. (2003). Half a century of antipsychotics and still a central role for dopamine D2 receptors. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 27*(7), 1081-1090.

Kapur, S., & Remington, G. (2001). Dopamine D2 receptors and their role in atypical antipsychotic action: still necessary and may even be sufficient. *Biological Psychiatry, 50*(11), 873-883.

Kawakami, Y., Nishimoto, H., Kitaura, J., Maeda-Yamamoto, M., Kato, R., Littman, D., et al. (2004). Protein kinase C beta II regulates Akt phosphorylation on Ser-473 in a cell type-and stimulus-specific fashion. *Journal of Biological*

Chemistry, M408797200v408797201.

Knill, D., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27*(12), 712-719.

Lee, T., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A, 20*(7), 1434-1448.

Lewis, C. M., Levinson, D. F., Wise, L. H., DeLisi, L. E., Straub, R. E., Hovatta, I., et al. (2003). Genome Scan Meta-Analysis of Schizophrenia and Bipolar Disorder, Part II: Schizophrenia. *The American Journal of Human Genetics, 73*(1), 34-48.

Lindenmayer, J., Grochowski, S., & Hyman, R. (1995). Five factor model of schizophrenia: replication across samples. *Schizophrenia Research, 14*(3), 229-234.

Lindström, L. H., Gefvert, O., Hagberg, G., Lundberg, T., Bergström, M., Hartvig, P., et al. (1999). Increased dopamine synthesis rate in medial prefrontal cortex and striatum in schizophrenia indicated by -([beta]-11C) DOPA and PET. *Biological Psychiatry, 46*(5), 681-688.

Mao, Y., Ge, X., Frank, C. L., Madison, J. M., Koehler, A. N., Doud, M. K., et al. (2009). Disrupted in Schizophrenia 1 Regulates Neuronal Progenitor Proliferation via Modulation of GSK3beta/beta-Catenin Signaling. *Cell, 136*(6), 1017-1031.

Mathur, A., Law, M., Megson, I., Shaw, D., & Wei, J. (2010). Genetic association of the AKT1 gene with schizophrenia in a British population. *Psychiatric Genetics, 20*(3), 118.

Montague, P., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*(5), 1936.

Montague, P., Hyman, S., & Cohen, J. (2004). Computational roles for dopamine in behavioural control. *Nature, 431*(7010), 760-767.

Moore, R., & Bloom, F. (1978). Central catecholamine neuron systems: anatomy and physiology of the dopamine systems. *Annual Review of Neuroscience, 1*(1), 129-169.

Murray, G., Corlett, P., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., et al. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry, 13*(3), 267-276.

Muscettola, G., Barbato, G., Pampallona, S., Casiello, M., & Bollini, P. (1999). Extrapyramidal syndromes in neuroleptic-treated patients: prevalence, risk factors, and association with tardive dyskinesia. *Journal of clinical psychopharmacology, 19*(3), 203.

Nordstrom, A., Farde, L., Eriksson, L., & Halldin, C. (1995). No elevated D2 dopamine receptors in neuroleptic-naive schizophrenic patients revealed by positron emission tomography and [11C] N-methylspiperone. *Psychiatry Research: Neuroimaging, 61*(2), 67-83.

Owen, M. J., Williams, N. M., & O'Donovan, M. C. (2003). The molecular genetics of schizophrenia: new findings promise new insights. *Molecular Psychiatry, 9*(1), 14-27.

Persad, S., Attwell, S., Gray, V., Mawji, N., Deng, J., Leung, D., et al. (2001). Regulation of protein kinase B/Akt-serine473phosphorylationbyintegrin-limkedkinase: Critical roles for kinase activity and amino acids arginine 211 and serine 343. *The Journal of Biological Chemistry, 276*, 29.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R., & Frith, C. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042-1045.

Reith, J., Benkelfat, C., Sherwin, A., Yasuhara, Y., Kuwabara, H., Andermann, F., et al. (1994). Elevated dopa decarboxylase activity in living brain of patients with psychosis. *Proceedings of the National Academy of Sciences of the United States of America, 91*(24), 11651.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical Conditioning. II: Current Research and Theory*. New York: Appleton-Century-Crofts.

Roh, M., SEA, M., Kim, Y., Kim, S., Jeon, W., Ahn, Y., et al. (2007). Haloperidol and clozapine differentially regulate signals upstream of glycogen synthase kinase 3 in the rat frontal cortex. *Experimental and molecular medicine, 39*(3), 353-360.

Rosenthal, D. (1964). *The Genain quadruplets: a case study and theoretical analysis of heredity and environment in schizophrenia*: Basic Books.

Salamone, J., Cousins, M., & Snyder, B. (1997). Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. *Neuroscience & Biobehavioral Reviews, 21*(3), 341-359.

Sanger, D. (2004). The search for novel antipsychotics: pharmacological and molecular targets. *Expert Opinion on Therapeutic Targets, 8*(6), 631-641.

Sarbassov, D., Guertin, D., Ali, S., & Sabatini, D. (2005). Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science, 307*(5712), 1098.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of*

Neurophysiology, 80(1), 1.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience, 13*(3), 900.

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593.

Schultz, W., Ruffieux, A., & Aebischer, P. (1983). The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Experimental Brain Research, 51*(3), 377-387.

Schwab, S., Hoefgen, B., Hanses, C., Hassenbach, M. B., Albus, M., Lerer, B., et al. (2005). Further evidence for association of variants in the AKT1 gene with schizophrenia in a sample of European sib-pair families. *Biological Psychiatry, 58*(6), 446-450.

Schwab, S., & Wildenauer, D. (2009). Update on key previously proposed candidate genes for schizophrenia. *Current Opinion in Psychiatry, 22*(2), 147-153

Seeman, P. (1987). Dopamine receptors and the dopamine hypothesis of schizophrenia. *Synapse, 1*(2), 133-152.

Skinner, B. F. (1963). Behaviorism at 50. *Science, 140*(357), 951-958.

Smith, Y., & Kieval, J. (2000). Anatomy of the dopamine system in the basal ganglia. *Trends in Neurosciences, 23*, S28-S33.

Staal, S., Hartley, J., & Rowe, W. (1977). Isolation of transforming murine leukemia viruses from mice with a high incidence of spontaneous lymphoma. *Proceedings of the National Academy of Sciences of the United States of America, 74*(7), 3065.

Stambolic, V., Suzuki, A., De la Pompa, J., Brothers, G., Mirtsos, C., Sasaki, T., et al. (1998). Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. *CELL, 95*, 29-39.

Sullivan, P. F. (2008). Schizophrenia genetics: the search for a hard lead. *Current Opinion in Psychiatry, 21*(2), 157.

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait - Evidence from a meta-analysis of twin studies. *Archives of General Psychiatry, 60*(12), 1187-1192.

Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron, 59*(2), 336-347.

Sutton, R. (1988). Learning to predict by the methods of temporal differences. *Machine learning, 3*(1), 9-44.

Sutton, R., & Barto, A. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience:*

Foundations of adaptive networks (pp. 497-537). Cambridge, MA: MIT Press.

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

Thiselton, D. L., Vladimirov, V. I., Kuo, P. H., McClay, J., Wormley, B., Fanous, A., et al. (2008). AKT1 is associated with schizophrenia across multiple symptom dimensions in the Irish study of high density schizophrenia families. *Biological Psychiatry, 63*(5), 449-457.

Toker, A., & Newton, A. (2000). Akt/protein kinase B is regulated by autophosphorylation at the hypothetical PDK-2 site. *Journal of Biological Chemistry, 275*(12), 8271.

Viniegra, G. (2004). Full activation of PKB/Akt in response to insulin or ionizing radiation is mediated through ATM. *Journal of Biological Chemistry, 280*, 4029-4036.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature, 412*(6842), 43-48.

Watkins, C. (1989). *Learning from delayed rewards.* Cambridge University, Cambridge.

Werbos, P. (1977). Advanced forecasting methods for global crisis warning and models of intelligence. *General Systems, 22*, 25-38.

Wise, R. (1982). Neuroleptics and operant behavior: The anhedonia hypothesis. *Behavioral and Brain Sciences, 5*(1), 39-87.

Wise, R. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience, 5*(6), 483-494.

Wise, R., Spindler, J., DeWit, H., & Gerberg, G. (1978). Neuroleptic-induced" anhedonia" in rats: pimozide blocks reward quality of food. *Science, 201*(4352), 262.

Wise, R., Spindler, J., & Legault, L. (1978). Major attenuation of food reward with performance-sparing doses of pimozide in the rat. *Canadian journal of psychology, 32*(2), 77-85.

Wong, D., Pearlson, G., Tune, L., Young, L., Meltzer, C., Dannals, R., et al. (1997). Quantification of neuroreceptors in the living human brain: IV. Effect of aging and elevations of D2-like receptors in schizophrenia and bipolar illness. *Journal of Cerebral Blood Flow & Metabolism, 17*(3), 331-342.

Wong, D., Wagner Jr, H., Tune, L., Dannals, R., Pearlson, G., Links, J., et al. (1986). Positron emission tomography reveals elevated D2 dopamine receptors in drug-naive schizophrenics. *Science, 234*(4783), 1558.

Xu, M. Q., Xing, Q. H., Zheng, Y. L., Li, S., Gao, J. J., He, G., et al. (2007). Association of AKT1 gene Polymorphisms with risk of schizophrenia and with response to Antipsychotics in the chinese population. *Journal of Clinical*

Psychiatry, 68, 1358-1367.

Yang, J., Cron, P., Good, V., Thompson, V., Hemmings, B., & Barford, D. (2002). Crystal structure of an activated Akt/protein kinase B ternary complex with GSK3-peptide and AMP-PNP. *Nature Structural & Molecular Biology, 9*(12), 940-944.

Chapter 2

Investigation of the Role of AKT1 in Reinforcement Learning Model and Its Possible Implications in Schizophrenia Psychosis

Introduction

Schizophrenia is a severe mental illness which can be characterized mainly by 'positive symptoms', 'negative symptoms', and cognitive deficits. Increasing evidence has suggested that genetic variants of *Akt1* are implicated in the pathogenesis of schizophrenia (Emamian, Hall, Birnbaum, Karayiorgou, & Gogos, 2004; Schwab, et al., 2005; Schwab & Wildenauer, 2009). AKT1 (also known as PKBα), a serine/threonine kinase of the AKT family, is involved in multiple biological processes and diverse signal transduction pathways (Franke, 2008). Accumulating evidence from studies of postmortem brain of schizophrenic patients (Emamian, et al., 2004; Zhao, Ksiezakreding, Riggio, Haroutunian, & Pasinetti, 2006), AKT1 deficient mice (Lai, et al., 2006), and functional neuroimaging of humans (Tan, et al., 2008), all suggest epistatic effects of *Akt1* variations (or its protein) in the regulation of dopamine (DA)-associated functions and DA system (i.e., mesolimbic, mesocortical, and nigrostriatal pathways) in the brain.

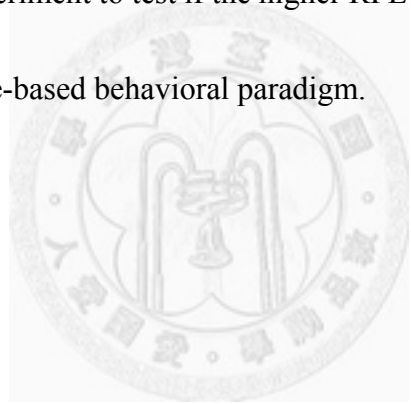
The DA hypothesis of schizophrenia has been one of the most durable hypotheses in schizophrenia, especially for the positive symptoms. After the discovery of Dopamine D2 receptors (DRD2) as the major target for antipsychotic drugs, abnormalities in DA system have long been implicated in the explanatory context of schizophrenia, especially psychosis (Kapur & Mamo, 2003; Kapur, Zipursky, Jones, Remington, & Houle, 2000). Convergent evidence indicates that AKT is a key signaling intermediate downstream from the DRD2, which interacts with β-arrestin2 and PP2A in the regulation of DA signaling cascades and the expression of DA-dependent behaviors (Beaulieu, Gainetdinov, & Caron, 2007, 2009).

Emerging evidence indicates that the phasic DA activity arising from the ventral tegmental area and substantia nigra of the midbrain appears to correlate with the history of reward delivery and reward prediction error (RPE) signal (Bayer & Glimcher, 2005; Hollerman & Schultz, 1998; Schultz, 1998; Schultz, Dayan, & Montague, 1997; Waelti, Dickinson, & Schultz, 2001). Throughout the course of learning, this pattern of dopaminergic firing tallies exactly with the characteristics of a temporal-difference prediction error in the reinforcement learning model (Montague, Dayan, & Sejnowski, 1996; Montague, Hyman, & Cohen, 2004; Schultz, et al., 1997), which means DA neurons fire only when things are not happened as expected. Neuroimaging data further suggested that this RPE system in the ventral striatum was

disrupted in patients with schizophrenia (Corlett, et al., 2004; Corlett, et al., 2007; Juckel, et al., 2006). Furthermore, the RPE signals arise from the ventral striatum in schizophrenic patients during psychosis episode appeared to be relatively suppressed to stimuli that were supposed to be salient, on the contrary, neutral stimuli elicit relatively augmented responses (Murray, et al., 2008). In addition, carriers with different DRD2 polymorphisms have been found to affect striatal RPE-related signal. For instance, carriers of the DRD2 Taq 1A (ANKK1-Taq 1A) polymorphism A1 allele showed reduced RPE signal (Cohen, Krohn-Grimberghe, Elger, & Weber, 2007; Cohen, Young, Baek, Kessler, & Ranganath, 2005; Klein, et al., 2007), whereas the 141C *Ins/Del* polymorphism carriers showed stronger response (Forbes, et al., 2009), suggesting the involvement of DRD2 in RPE-related learning. Disruption in prediction error and its consequences may explain the emergence of psychosis, a DA-related clinical hallmark of schizophrenia (Fletcher & Frith, 2009).

Based on the findings above, it is of particular interest to ask whether AKT1 is involved in the regulation of RPE signal by the DA system. In order to verify this hypothesis and establish a causal relationship, a mutant mouse model is a simplified and alternative approach for determining the basic function of AKT1 in the RPE and related task. In this study, *Akt1* heterozygous (HET) mice, which have similar decreased AKT1 expression level as schizophrenic patients, and wild-type (WT)

littermates were used in a series of experiments. In experiment 1, choice behaviors from a dynamic foraging T-maze task were fit to the *Q* learning model to examine whether AKT1 modulates RPE signal. The parameter estimation was done by Bayesian hierarchical parameter estimation, in which we are able to describe the difference of parameters between experiment and control groups in probability. In experiment 2, a series of reward-related tasks were conducted to examine whether the prediction by the simulation can be revealed behaviorally. Latent inhibition was performed as the third experiment to test if the higher RPE signal can also predict performance in an aversive-based behavioral paradigm.



Material and Method

Subjects

All *Akt1* heterozygous (HET) mice and their wild-type (WT) littermates used in this study were generated from *Akt1*-heterozygous breeding pairs in a C57BL/6J genetic background and genotyped using PCR analysis of mouse tail DNA as described previously (Cho, Thorvaldsen, Chu, Feng, & Birnbaum, 2001). *Akt1* heterozygous mice express only 23% of AKT1 protein in the ventral striatum compared with WT controls (Figure 2-1a and 2-1b). Similar reduction was also evident in other brain areas, including the prefrontal cortex, whole cortex, hippocampus, and cerebellum (Figure 2-8). The expressions of other AKT isoforms (i.e., AKT2 and AKT3) did not differ between genotypes (Figure 2-1c and 2-1d). All subjects were 12 to 16 weeks old at the time of experiments. They were housed in Polysulfone individually ventilated cages (Alternative design Inc., U.S.A.) within the animal rooms of the Psychology Department, National Taiwan University, and the housing condition (either single housing or group (3-5 per cage) housing depended on experimental requirements) was settled at least one week prior to any experiment. Behavioral experiments were conducted at least half an hour after dark/light (depends on the experimental requirements) cycle begins. All animal procedures were performed according to the protocols approved by the Animal Care and Use

Committees established by National Taiwan University.

Immunoblotting

Samples of mouse brain were quickly dissected, frozen by liquid nitrogen, and stored at -80 °C until protein extraction. Tissue samples were homogenized in lysis buffer, containing 25 mM Tris, pH 8.0, 125 mM NaCl, 1% Triton, Protease Inhibitor Cocktail tablets (Roche) and Phosphatase Inhibitor Cocktail 1 (Sigma), centrifuged at 14,000 rpm at 4 °C for 20 min, and the supernatant was collected. Protein concentration was measured by Bradford protein assay (Bio-Rad) and spectrometry at 620 nm. An equal amount of protein was separated on 4%-10% SDS/PAGE and transferred onto nitrocellulose membrane (Millipore). Following transfer, the membranes were washed in Tris-buffered saline with 0.1% Tween-20 (TBST), and blocked in 5% w/v skimmed milk for 1 hr at 25 °C. Membrane was then incubated in 1% BSA solution with the appropriate primary antibody overnight at 4 °C. The blot was probed with the following antibodies: AKT1 (1:2000, Cell Signaling Technology, #2967), AKT2 (1:2000, Cell Signaling Technology, #5239), AKT3 (1:1000, Cell Signaling Technology, #3788), phospho-AKT1 Ser473 (1:1000, Millipore, #05-669), and GAPDH (1:5000, Cell Signaling Technology, #2118). Immune complexes were revealed by using appropriate peroxidase-conjugated secondary antibodies (Cell

Signaling Technology). Membranes were stripped and reprobed while necessary. Antibody binding was detected with ECL chemiluminescence kit (Millipore) and densitometric analysis was performed using the Dolphin-View image software (Wealtec Corp, Sparks, NV, USA).

Behavioral procedures and analysis

Dynamic foraging T-maze task – behavioral procedure: Subjects were deprived for food and kept in 80~85% of their original body weight. A white Acrylic T-maze, consisted of one start arm and two choice-arms (50 X 10 X 15 cm), was used. A red square food cup was placed at the terminal of the choice-arm, where a small piece of chocolate cereal (Kellogg's) reward could be obtained. After two days of shaping, each mouse was trained to reach the food cup in either arm within 10 sec for 10 continuous trials. The dynamic foraging T-maze was a two-alternative forced-choice task. One of the alternative arms was assigned to present a reward at the rate of 75%, while independently, the probability of receiving reward in the other arm was 25%. Each subject discovered this rule and chose the high reward rate arm by trial and error. Every day, a subject underwent testing with minimum 3 to maximum 6 blocks, where a block consisted of 10 trials. There were three sequential testing sections for the T-maze task, including acquisition, reverses learning, and methamphetamine (METH,

1 mg/kg, i.p.) challenge. METH was administrated 30 min before daily training. The criteria of accomplishing each section was chosen with the high reward rate arm above 70% accuracy consecutively in 3 blocks, along with an average of above 80%. Once the criterion was achieved, each mouse moved on to the next section on the next testing days and the reward rates of the two choice arms were switched. Choice result and latency to reach the food cup were recorded trials by trials during daily training. The trial-by-trial data for the dynamic foraging T-maze task were further analyzed and elaborated by the reinforcement learning model and matching law analysis.

Dynamic foraging T-maze - Reinforcement learning model: Standard *Q* learning model (Sutton & Barto, 1998; Watkins, 1989) was applied to fit trial-by-trial choice data from all subjects in the T-maze task. In this model, series of choice and choice result were utilized to estimate the action value of the two options for every trial (Samejima, Ueda, Doya, & Kimura, 2005). At the beginning of the T-maze experiment, the expected values are set to 0, the value of the chosen arm was updated according to the following rule (e.g., $Q_{high}(t)$ for the value of choosing the high reward rate arm on trial *t*):

$$Q_{high}(t+1) = Q_{high}(t) + \alpha\delta(t)$$

$$\delta(t) = R_{high}(t) - Q_{high}(t)$$

δ(t) is the reward prediction error (RPE), the discrepancy between experienced and

the reward just received. $R_{high}(t)$ features the outcome received from the high reward rate arm on trial *t* with a value of 5 for a reward and 0 for the absent of reward. α denotes the learning rate, which determines how rapidly the estimate of expected value is updated. The expected value of both high and low rate arm was calculated, the probability of choosing the high reward arm $P_{high}(t)$ is determined by softmax selection or Boltzmann exploration (Kaelbling, Littman, & Moore, 1996; Luce, 1959):

$$P_{high}(t) = \frac{e^{\beta Q_{high}}}{e^{\beta Q_{high}} + e^{\beta Q_{low}}}$$

The parameter β represents the choice perseveration. When β is 0, the chances of a subject to chose high reward rate arm is 0.5, indicates the choice is random between two options. Therefore, the closer the β is to 0, the less perseverated the choice is. The values of α and β were estimated using Bayesian hierarchical model as depicted in Figure 2-2a. In Bayesian parameter estimation, the uncertainty about parameters is quantified by probability distributions. Prior parameter distributions are updated by incoming data to yield posterior distributions. The posterior distributions are the quantification of our uncertainty about the parameter after seen the data. In Bayesian hierarchical model, parameter of an individual is assumed to be drawn from a group-level distribution. One of the benefits of using hierarchical model is that it deals

with the individual differences problem. Because the multi-level structures naturally incorporate both the differences and the commonalities between subjects. As shown in Figure 2-2a, α and β of subject *i* (α_i and β_i) was assigned as normal distribution, and its mean and standard deviation is drawn from the group level (μ_α σ_α and μ_β σ_β respectively). Therefore μ_α and μ_β of the WT and HET groups in each testing section are what we actually interested in and monitored.

This model was computed and programmed in the WinBUGS (the MS Windows operating system version of BUGS: Bayesian inference Using Gibbs Sampling (Casella & George, 1992)) and WinBUGS Development Interface (Lunn, Thomas, Best, & Spiegelhalter, 2000), that have been developed to approximate distributions by sampling value from them using Markov Chain Monte Carlo techniques. Three Markov chains were constructed to estimate the two parameters μ_α and μ_β. By using three chains, the $\hat{R}$ value was obtained to ensure the Markov Chains converged properly (all $1 \leq \hat{R}$ <1.1). A chain is consisted of 16000 iterations. The first 6000 (burn in) points were discarded to ensure that we only use samples come from the stationary distribution, and the data were hence unaffected by the starting value. Thus we obtained 30000 points of estimations from 3 chains, and took samples at the interval 5, which gave us 6000 points. All interpretations and tests were made based on these 6000 samples. To test whether the assigned prior distribution can

bias the posterior distribution, Beta distribution (2, 5) and (2, 8) prior distribution was also assign to μ_α ; and normal distribution with mean 0 and standard deviation 1was assign to μ_β .

Dynamic foraging T-maze - Matching law analysis: Matching law was used to describe how subjects make choices among two arms that differed in the expected value. Steady-state (last 30 trials of each testing stage) choice behavior in the dynamic foraging T-maze task was fitted by least-square regression in the logarithmic form of the generalized matching law (Baum, 1974):

$$\log_2\left(\frac{C_{high}}{C_{low}}\right) = a\log_2\left(\frac{R_{high}}{R_{low}}\right) + \log_2 c$$

Numbers of choices to the high and low reward probability arms are denoted as C_{high} and C_{low} respectively, R_{high} and R_{low} represent the number of rewards received from the high and low arm. When the slope (a) of this line is 1, it represents the maximal discrimination between two choices. In another words, it is a measure of the sensitivity of choice allocation to reward frequency. When c differs from 1, its magnitude represents the degree of bias in choice. A part of data (WT: 5 out of 24 blocks, HET: 5 out of 27 blocks) in which their C_{low} or R_{low} were equal to 0 (i.e., completely chose the high reward arm) was excluded from analysis in order to fit the formula above.

Reward related behaviors - Two-bottle sucrose preference test: Another batch of subjects was single-housed for at least a week before experiment began. The test was conducted inside each subject's home cage and sucrose preference was measured by the two-bottle free-choice paradigm as a test for nature reward. The two bottles were first filled with diluents (drinking water) in day 1 for 24-h acclimation, and drinking baseline was measured to make sure subjects showed no preference for either one of the bottles. In day 2~4, the two bottles were refilled with 1% and 2% sucrose solution respectively and daily fluid intake was measured by weighting the bottles for three days. The positions of the bottles were interchanged every time after the measurements. The preference was calculated as: 100 x (weight of 2% fluid intake / weight of 1% fluid intake).

Reward related behaviors - Incentive licking behavior: Incentive licking behavior was conducted to measure the needs for natural reward under differential levels of deprivation and reward intensity. Subjects were group-housed by 4 and deprived for water. Mice were kept in 80-85 % of their original body weight. Incentive licking behavior was performed in a mouse shuttle cage (Coulbourn Instrument, Whitehall, PA, USA) during the dark cycle, only one box was equipped with an optical lickometer which provided fluid as a reward and recorded licking frequency. The shuttle cage was placed in a sound proofed room with a dim light

while the experiment proceeded. On each trial, each subject was placed into the other box for a 30-sec acclimation before a guillotine door opened. The door was closed once the subject entered the other box. Each subject was given three minutes to stay in the box and lick the lickometer to obtain fluid. The whole task consisted of 7 blocks that varied from different water-deprivation or sucrose concentration conditions, and each block consisted of 3 daily training processes. The first 3 blocks were the training period which subjects learned and obtained water in the shuttle cage. In block 4, drinking water was *ad lib* in their home cages while 4% of sucrose solution was provided in the testing shuttle box. In block 5, water was still freely available in their home cages, whereas the concentration of the sucrose solution was raised to 32%. In block 6, the sucrose solution was maintained in 32% but subjects were deprived for water 8 hours before the experiment. In block 7, 32% sucrose solution remained used but the water deprivation was prolonged to 23 hr. On each testing day, latency from the guillotine door open to the first lick and the licking frequency were recorded automatically.

Reward related behaviors - Conditioned place preference (CPP): CPP paradigm was adapted from the Current Protocols in Neuroscience (Roux, Froger, Porsolt, Valverde, & Maldonado, 2003) and it was used to measure the acquisition of drug reward. Briefly, the procedure was conducted during light cycle, and another batch of

subjects was group housed by four. CPP was conducted in a PVC chamber (48 × 24 × 25 cm) that was equally divided into a pairing and a non-pairing compartments connected by a square hole (5×6 cm). The two compartments were distinct in their olfactory (vanilla and orange scents), tactile (wood or corn cobs bedding), and wall context (stripes or geometric pattern) properties. The CPP procedure consisted of 8 daily sessions, including a 18-min free exploration of the pairing and non-pairing compartments in day 1, a 36-min of METH- (1.0 or 2.0 mg/kg, i.p.) context conditioning in days 2, 4 and 6, a 36-min of saline-context conditioning in days 3, 5, and 7, and a 18-min testing in day 8. According to the preference recorded in day1, subjects were conditioned to the compartment that showed less preference. Throughout the task, time spent in each compartments and locomotor activity were recorded simultaneously and analyzed automatically by EthoVision tracking system (Noldus Information Technology, the Netherlands).

Latent inhibition: Latent inhibition was assessed in a conditioned emotional response paradigm to measure the reduction of learning to a prior exposure of irrelevant distraction. It was conducted in a conditioning chamber (Coulbourn Instrument) during the dark cycle. Another batch of mice were placed on a 23-h water restriction schedule before pre-training began and maintained 80~85% of their original body weight throughout this experiment. Both male HET and WT mice were

randomly assigned to either the non pre-exposed (NPE) group or pre-exposed (PE) group (n = 7 or 8). The procedure consisted of a 6-day pre-training session, a pre-exposure session in day 7, a conditioning session in day 8, a re-baseline session in day 9, and a test session in day 10. During the pre-training session, each subject was placed in the conditioning chamber and allowed to drink freely from a water sipper for 20 min in day 1 and for 15 min in days 2-6. During the pre-exposure session (day 7), mice were placed in the conditioning chamber without access to water. Males in the PE group were given 40 presentations of an 80-dB (2000 Hz) 50s tone with a 15-s interstimulus interval. NPE control mice were place in the chambers for the same amount of time but received no pre-exposures to the tone. During the conditioning session (day 8), mice were placed in the chamber without access to water. After 5 min, each male received two tone-shock pairings. Each tone (conditioned stimulus, CS) was of 5-s duration and was followed by a 1-s 0.4-mA footshock (unconditioned stimulus, US). There was a 5 min interval between pairings and mice remained in the chamber for 5 min following the second tone-shock presentation. During the re-baseline session (day 9), mice were placed in the conditioning chamber for 15 min and given free access to the water sipper to re-establish licking in the chamber prior to testing. In the test session (day 10), mice were placed in the conditioning chamber with access to the water sipper. The number of licks, time to first lick, time to

complete licks 1-80, 81-90 (A), 91-100 (B), and latency to first lick after the CS tone presentation were recorded automatically. After the first 90 licks, the CS tone was presented until the mouse reached lick 100. The standard measure of conditioned suppression was the time taken to complete licks 90-100 in the presence of the CS. A suppression ratio (SR) was calculated according to the formula A/(A+B) yielding a scale of 0-0.5. Low SR indicated good learning while high SR indicated poor learning of the association between the tone and footshock. Latent inhibition is demonstrated as a higher SR in the PE group compared to the NPE group.

Statistical Analysis and software:

Behavioral data were analyzed with one or two ANOVA, or student's t-tests where appropriate. *Post-hoc* analysis was performed with Fisher's LSD test when *F* values reaching significant difference. Priori t-test and Bonferroni adjustments was conducted to answer specific hypotheses. Statistic analysis was done by SPSS 13.0 (SPSS, Inc., Chicago, IL, USA). *P* values of $< .05$ were considered statistically significant. Kolmogorov–Smirnov test and least square regression curve fitting were performed by Matlab (The MathWorks Inc., Natick, MA, USA) statistics toolbox and curve fitting toolbox. Kolmogorov-Smirnov Test was performed to test whether any two of the distrubution were identical.

Results

Experiment 1

Dynamic foraging T-maze - Behavioral restuls: For total number of cumulated trials, both WT and HET groups took about 90 trials to acquire this task in the acquisition section and there is no significant difference. After the first transition, HET mice spent significantly fewer trials to achieve the criteria compared with WT mice in the reverse section ($t(15) = -2.54$, $p < .05$; Figure 2-2b). WT mice spent more trials to switch to the new role compared to what they did in the acquisition section ($t(7) = 2.79$, paired, $p < .05$). No genotypic difference was found in the METH section. For average latency from the start box to food cup, no significant difference was found in the 3 sections ($F(1, 15) = 1.171$, $p > .05$).

Dynamic foraging T-maze - Bayesian parameter estimation: Posterior distrubutions of learning rate (α) in the three testing sections were shown in Figure 2-3a, and choice perserverations (β) were shown in Figure 2-3b. For both μ_α and μ_β, Kolmogorov-Smirnov Test indicated that none of any two posterior distributions between WT and HET groups in the three testing sections were identical (Table 2-1). For the measurement of learning rate (α), μ_α for HET group (HET(μ_α , σ_α)) is significantly higher than WT group (WT (μ_α , σ_α)) in the acquisition section ((0.345, 0.12) VS. (0.229, 0.11)), reverse section (0.143, 0.10) VS. (0.049, 0.03), and METH

section ((0.018, 0.007) VS. (0.011, 0.007)) respectively. For choice perseveration (β), μ_β for HET group (HET(μ_β, σ_β)) is significantly lower than WT group (WT(μ_β, σ_β)) in the acquisition section ((0.22, .05) VS. (0.296, 0.08)), reverse section ((0.246, 0.15) VS. (0.491, 0.28)), and METH section ((1.991, 0.66) VS. (3.377, 0.79)). For bias controls, non-informative prior distribution (uniform distribution) and Beta distribution assigned to μ_α all yielded similar distribution patterns, indicating that our model is not sensitive to, or easily biased by manually assigned prior distribution (Figure 2-9, 2-10). Uniform distribution and normal distribution assigned to μ_β also did not affect the patterns of the posterior distribution (Figure 2-8). Thus, these data indicate that AKT1 deficit significantly enhanced learning rate and reduced choice perserveration in all testing sections. METH treatment, however, made a differential and great impact on reducing the learning rate and enhancing choice perserveration in both groups, suggesting that METH-induced synaptic alteration might interact with postsynaptic AKT1 activity in the regulation of reinforcement learning.

AKT1 activity analysis: The effect of METH and AKT1 activity were further evaluated by measuring phosphorylation of AKT1 at the serine473 position (pAKT1) in striatum using Western blot. In the dorsal striatum (Figure 2-4a), 30 minutes after the injection of METH, pAKT1 level was elevated in the WT mice compared with

saline injection ($t(9) = 1.775$, $p = .05$) whereas such treatment effect did not find between the two HET groups ($t(10) = 0.443$, $p > .05$). A genotypic difference of pAKT1 level was also observed under the METH treatment condition ($t(9) = 2.915$, $p < .05$) but not in the saline condition. Similar patterns were also found in the ventral striatum (Figure 2-4b). Again, compared with saline treatment, the injection of METH significantly enhanced pAKT1 expression in the WT group but had no effect on the HET group. The METH injection also resulted in a marginal increase of pAKT1 expression in WT males compared with HET males ($t(10) = 1.54$, $p = .07$). These data support the involvement of AKT1 in the regulation of DA-related striatal responses.

Dynamic foraging T-maze - Matching law analysis: Behavioral data of steady-state choices generally obeyed the matching law (Figure 2-5), that is, allocating choices as a function of reward ratios. Log choice ratios (ratio of high to low arms choices) were plotted as a function of log reward ratios for the last 30 trials of each section. Reward sensitivity (i.e., the slope of the fitted line) for HET is 0.509 (with 95% confident bound 0.305 to 0.713) and it is relatively higher than the one for WT (0.176 with 95% confident bound -0.187 to 0.54). The R^2, goodness of fit, for WT group is 0.537, and HET is 0.739.

Experiment 2

Reward related behaviors - Two bottles preference test: As depicted in Figure 2-6a, after establishing water baseline in day 1, HET mice showed an immediate preference to the 2% sucrose solution in day 2, that is, there is a significant difference between day 1 and day 2 (Bonferroni , $p < .05$). In contrast, WT mice didn't exhibit such preference. In day 2, HET males also showed a significantly higher magnitude of preference to 2% sucrose solution than WT males did (Bonferroni, $p < .05$); indicating HET males quickly noticed the change and showed their preference. No significant differences were found in day 3 and 4 between the two genotypes.

Reward related behaviors - incentive licking test: As shown in Figure 2-6b, a two-way ANOVA revealed a blocks effect ($F(6, 78) = 29.122$, $p < .05$) and genotype-blocks interaction ($F(6, 78) = 2.901$, $p < .05$). During the second and third blocks (acquisition blocks), the latency to first lick of HET was significantly lower than WT (both $p < .05$), suggesting HET learned faster and had higher learning rate than WT controls. After 3-day acquisition, no significant difference between groups was found in the following blocks or under different testing conditions. Total licking frequency, total drinking volume, and licking motor did not differ between genotypes in each testing block (all $p > .05$) (Figure 2-11).

Reward related behaviors - conditioned place preference (CPP):

For evaluating METH-induced reward in CPP, a two-way ANOVA (conditioning × genotypes) revealed that only the conditioning effect was significant in both 1mg/kg ($F(1, 14) = 30.43$, $p < .05$), and 2mg/kg ($F(1, 18) = 24.735$, $p < .05$). As depicted in Figure 2-6c, the conditioning was established successfully at both doses of METH (all $p < .05$) and no significant difference was found between the two genotypes. For average locomotion during conditioning, the injections of both 1 and 2 mg/kg METH significantly induced hyperlocomotion compared with saline injections. A genotypic difference was revealed at the dose of 2 mg/kg and HET males moved significantly less than the WT males ($p < .05$, Figure 2-6d: 1 mg/kg on the left; 2 mg/kg on the right).

Experiment 3

Latent inhibition: Both WT and HET groups showed differences between pre-exposed and non pre-exposed group (Bonferonni, all $p < .05$). Also as indicated by two-way ANOVA, there is no significant interaction ($F(1, 36) = 0.104$, $p > .05$), suggesting neither pre-exposed or non pre-exposed group was affected by the genotype. (Figure 2-7).

Discussion

In the present study, a series of behavioral tasks was conducted to test *Akt1* HET mice and their WT littermates. A dynamic foraging T-Maze task and a standard *Q* learning model were applied to examine whether AKT1 plays a role in regulating RPE magnitude. By implementing Bayesian hierarchical parameter estimation, *Akt1* deficient mice were found to have higher learning rate (α) than WT controls among all three testing sections, meaning that *Akt1* mutants have higher RPE magnitudes and action values were updated faster in these mice. Similar findings were also demonstrated during acquisition phase of two bottles preference test and incentive licking test, in which *Akt1* mutants display relatively efficient way to update reward information from the environment. Both genotypes showed normal latent inhibition, it suggested that reward prediction error obtained from an appetitive-based behavioral task, which showed genotypic differences, may not be able to explain or predict aversive-based learning. Meanwhile, analyses on choice behaviors (β) and matching law in the T-maze task further reveal that both WT and HET allocate more choices to the arm with higher reward rate but HET mice show lower choice perseveration than WT. HET mice also have a relatively higher reward sensitivity (i.e., the slope of the fitted line) to discrimination between two choices than WT mice. These metadata might provide fundamental parameters to describe the genotypic differences in reward

learning and a probable explanation for the behavioral difference, especially during the reverse learning in the T-maze.

In this study we compared the behavior of *Akt1* HET mice in two different states; one was the basal state without any pharmacological challenge, and the other was the METH state, in which we tried to determine whether the behavioral output of the mice could be disturbed by increased extracellular DA concentration with METH injection. In the T-maze task, the behavioral data indicates differences between genotypes in the reverse section but not in the METH challenge section. The dosage of the METH used in this section was relatively mild (1 mg/kg) in order to ensure the subjects are still capable for performing the task. The dosage, on the other hand, could also diminish the expected effect of METH on gross behavioral index, the total number of cumulated trials. However, by implementing trial-by-trial data, the parameter yield from the model revealed that learning rate was decelerated, and the choice perseveration was increased in the METH section. Furthermore, this finding was further supported by biochemical analysis of striatal samples using Western blots. The phosphorylated AKT1 of the HET group wasn't significantly different from the WT group in the basal condition, while the METH challenge increased the phosphorylation in the WT controls but not in the HET mice. This implies that the METH-induced DA efflux can interact with postsynaptic DRD2 in the striatum and

affect DA signaling cascade. The deficiency of AKT1, a downstream protein of DA cascade, could potentially diminish DA-related actions and the expression of DA-dependent responses as demonstrated in the mutant mice of this study. In contrast, in METH-induced conditioned place preference test, both HET and WT mice readily prefer METH-paired chamber more than saline-paired one, suggesting AKT1 may not involve in conditioning between drug reward and environmental cues. However, independent to rewarding effect of METH, 2 mg/kg of METH induced genotypic differences in hyperlocomotor activity, suggesting that Akt1 also plays a role in the regulation of METH-induced locomotion instead of METH-induced rewarding in this specific addiction task.

The striatum is a major target of DA innervations and DA-related reinforcement learning. Human functional magnetic resonance imaging studies have consistently show correlations between striatum and reward prediction error (Abler, Walter, Erk, Kammerer, & Spitzer, 2006; Klein, et al., 2007; Knutson & Cooper, 2005; McClure, Berns, & Montague, 2003; O'Doherty, Hampton, & Kim, 2007). Several other brain areas receiving dopaminergic projections are thought to further process RPE signals, including posterior medial frontal cortex (pMFC) (Frank, Woroch, & Curran, 2005), and rostral cingulate zone (Holroyd & Coles, 2002). For examples, it was reported that rostral cingulate zone receiving error signal produced by

nigrostriatal and mesocortical DA neurons to adjust action selection, and optimize action outcome prediction in the striatum (Holroyd & Coles, 2002), where pMFC is responsible for performance monitoring, integrating action outcome over multiple trials (Behrens, Woolrich, Walton, & Rushworth, 2007; Holroyd & Coles, 2002; Jocham, Neumann, Klein, Danielmeier, & Ullsperger, 2009; Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006; Rushworth, 2008; Rushworth & Behrens, 2008). Within the striatum, striatal DA D2 receptors (DRD2) were thought to play a prominent role in reward prediction and reinforcement learning (Frank, Moustafa, Haughey, Curran, & Hutchison, 2007; Jocham, Klein, et al., 2009; Kirsch, et al., 2006; Klein, et al., 2007). Carriers with DRD2 Taq 1A (ANKK1-Taq 1A) polymorphism, which was thought to possess lower DRD2 density in the striatum (Jonsson, et al., 1999), were found to be associated with lower reward-related, and increased error-related signals (Cohen, et al., 2007; Cohen, et al., 2005; Klein, et al., 2007). These findings highlight the probable roles of striatal DRD2 in the process of RPE. On the same line, in Parkinson's patients, it was reported that patients on L-dopa medication exhibit higher learning rate and lower choice perseveration in RL model (Rutledge, et al., 2009), suggesting that the RPE magnitude is subsequently intensified by promoting DA biosynthesis. Parkinson's patients on dopaminergic medication also display higher reward sensitivity in matching law analysis. These

findings in Parkinson's patients are somewhat similar to our current findings in *Akt1* deficient mice. Although it is plausible to speculate mutant mice might have higher striatal DA level as Parkinson's patients do on dopaminergic medication, previous study using *in vivo* microdialysis did not reveal higher basal DA release in the striatum and prefrontal cortex in AKT1 knockout mice (Lai, et al., 2006). Instead of speculation around presynaptic activity, it is more probable that postsynaptic events and signaling cascades are responsible for processing RPE signal because AKT has been proven downstream to DRD2 signaling cascades in striatum (Beaulieu, et al., 2007; Beaulieu, et al., 2005). Further research is needed to confirm this speculation.

In addition to the appetitive-rewarding tasks, based on a higher RPE signal we observed in the first two experiments, the performances of *Akt1* HET mice in an aversive-based learning task (i.e., the latent inhibition) were also predicted. We postulated that higher RPE signal will decrease the prediction of the likelihood that a previously exposed stimulus will further be associated with an unconditioned stimulus, therefore the inhibitory effect should be increased. However, both genotypes readily showed normal function of latent inhibition. The pre-exposed group of the HET does not show higher suppression ratio compared with WT group. It is possible that there is a ceiling effect that HET cannot have suppression ratio score any higher (i.e., less suppressed on licking). However, what is more likely is that there are two distinct

brain systems processing positive and negative motivation (Matsumoto & Hikosaka, 2008, 2009). Therefore, simulation data acquired from reward-based learning task may not be able to make prediction on an aversive-based learning one. It further implies that AKT1 might involve in the positive motivation circuit rather than circuit processing the negative motivation. There are indeed more works need to be done in order to elucidate this issue.

RL model and hierarchical Bayesian parameters estimation has been successfully applied in a number of human and non-human primate studies. However, to the best of our knowledge, it hasn't been done in mice or any genetically modified animal. A unique attempt of the present study is to combine computational methods with genetically engineered mice to investigate the role of AKT1, a schizophrenia susceptibility gene, in a DA-related, trial-by-trial reward learning task using AKT1 deficient mice as a model. Mouse models, especially genetically modified mice, offer numerous of advantages which cannot be done in other animal models or human. These *Akt1* HET mice used in this study have similar decreased AKT1 expression level as reported in schizophrenic patients (Emamian, et al., 2004; Zhao, et al., 2006), which provides a feasible model to examine the causal relationships between *Akt1* gene and psychosis-like symptoms. In the past decade, in the search for the biological functions of each schizophrenia susceptibility candidate gene or their combinations in

the pathogenesis of schizophrenia, genetically modified mice constitute a feasible model in which to study whether such manipulation affects the behavioral and physiological alterations. An oncoming challenge will be to link biological abnormalities to the clinical phenomena. For example, it remains unclear how to link dopaminergic abnormality in the brain to the psychosis reported in schizophrenic patients. As proposed by Kapur and his colleagues in the DA hypothesis of psychosis, multiple hits interact to result in striatal DA dysregulation is hypothesized to alter the appraisal of stimuli, perhaps through a process of aberrant salience, and eventually lead to psychosis (Howes & Kapur, 2009; Kapur & Mamo, 2003).

In a seminal attempt to link clinical phenomena to basic biological mechanism, a renovated theory of DA hypothesis of schizophrenia was proposed in which psychosis is considered as a state of aberrant salience (Kapur, 2003). A dysregulated hyperdopaminergic state at the biology level results in an aberrant assignment of salience to the elements that one's experiences at the "mind" level. This theory has been further elaborated by taking in a more sophisticated view of DA's role (Fletcher & Frith, 2009), the reward prediction error (RPE). Accordingly, a renovated theory of DA hypothesis of psychosis was further proposed that psychosis is a result of disturbances in the RPE signal generated by DA system, where inferences and beliefs about the world can not be properly updated, or corrected (Fletcher & Frith, 2009).

Such idea grounded in the theory that the brain is a Bayesian inference machine (Friston, Kilner, & Harrison, 2006; Lee & Mumford, 2003), which processes information with a hierarchical Bayesian framework (Bayes, 1783). Patients with schizophrenia also showed disrupted prediction error system stem from the ventral striatum, where it was shown to be under-active in response to rewarding experiences (Juckel, et al., 2006). Furthermore, patients undergo psychotic episode had a relatively suppressed response to stimuli that should be relevant and important and a relatively augmented response to stimuli that should be neutral or unimportant (Murray, et al., 2007). In addition, the vulnerability of ketamine psychosis was predicted by the baseline measure of prediction error signals in healthy controls (Corlett, et al., 2006).

Our data demonstrated *Akt1,* a schizophrenia candidate gene, is involved in regulating reward prediction error produced by the dopamine system: decreased AKT1 quantity resulted in increased RPE magnitude. This suggested that AKT1 might involve in the pathogenesis or vulnerability to psychosis. Using car driving as a metaphor, in reward learning, the RPE-processing circuit with reduced AKT1 level might produce augmented prediction-error signals, which corresponds with a cruise control that failed to take over the throttle to maintain a steady speed or strategy to gain a known reward. Accordingly, in addition to other genetic deficits and environmental insults, we hypothesize that lack of fully-functional AKT1 in the brain

could lead to an overreacting prediction-error signals that failed to maintain a steady winning strategy, and result disturbances in updating information from the current environment and eventually lead to disturbances in forming new/posterior believes, which could be one of the causes for psychosis in schizophrenia. In contrast to the final common pathway of presynaptic striatal hyperdopaminergia hypothesis proposed by Howes and Kapur, our data suggest that AKT1 in the postsynaptic striatum might be involved in the regulation of DA-related RPE or motivational salience. Nonetheless, the compensatory and modulatory effects of other downstream proteins in the DA signaling cascades cannot be underestimated. Future work will be needed to further confirm these speculations, for example, electrophysiological or voltametry recording directly from midbrain DA neurons or terminals to investigate phasic neuron firing or DA release during reinforcement learning in Akt1 deficient mice.

Figure legends

Figure 2-1. Protein expression analysis. (a) Quantitative Western blots result against AKT1 for *Akt1* wildtype (WT), heterozygous (HET), and homozygous (HOM) in the ventral striatum. The figure is represented as mean + s.e.m. (b) Representative blots of AKT1 expression in the ventral striatum, HET mice only expressed 23% AKT1 level comparing with normal WT level. Expressional level of other AKT isoforms AKT2 (c) and AKT3 (d) are not different between genotypes.

Figure 2-2. (a) Graphical Bayesian hierarchical model of reinforcement learning model in the dynamic foraging T-maze task. In this graphical model, nodes are the variable of interest, and the arrows indicate dependencies between the variables. For nodes having double borders mean that the variables are deterministic rather than stochastic. Whereas circular nodes represent continuous variables, square nodes represent discrete variables. Shaded nodes are the observed variables, nodes that are not shaded indicates variables unobserved. $R_{i,j-1}$ indicates the reward subject i received in trial $j-1$. $Ch_{i,j}$ represents the observed choice of subject i in trial j. $i = 1, \ldots, N$ represents the number of subjects; $j = 1, \ldots, TRIALS$ corresponds to the number of choice in the T-maze task. According to the reinforcement learning

model, α is between 1 and 0, therefore we assigned a non-informatics continuous uniform prior distribution to μ_α with upper bond 1 and lower bond 0. The σ_α was also assigned a uniform (1, 0) prior distribution. For μ_β, consider β can be either positive of negative, we assigned a continuous uniform prior distribution to it with upper bond 5 and lower bond -5, and the prior of σ_β is as well assigned to be a continuous uniform distribution (1, 0). (b) Behavioral result of dynamic foraging T-maze task, HET spent fewer trials to meet criteria in the reverse block. The figure is represented as mean + s.e.m. * $p < .05$.

Figure 2-3. Bayesian parameter estimation of learning rate and choice perseveration. (a) learning rate α, in terms of posterior mean, HET is consistently higher than WT through out blocks. (b) The posterior mean of choice perseveration β is lower in acquisition, reverse, and METH blocks. According to Kolmogorov-Smirnov test, no any of the two distributions are identical. By Gelman-Rubin convergence statistic (all $1 \leq \hat{R} < 1.1$) and visual inspection, the Markov chains converged properly. Numbers in the figures indicates means and the 95% credible interval of the posterior distribution.

Figure 2-4. AKT1 activity analysis as indicated by serine 473 phosphorylation. After

METH administration (2mg/kg, i.p., 30 min) WT showed an increased phosphorylation of AKT1 in dorsal striatum (a) and ventral striatum (b), indicating higher AKT1 activity induced by METH. Such increments are not observed in HET. The figure is represented as mean + s.e.m. *$p < .05$.

Figure 2-5. Steady state choice behavior of both WT and HET obeys the matching law. Log choice ratios are plotted as a function of log reward ratio. Data are plotted for all subjects of WT and HET groups. The reward sensitivity represented by the slope of the fitted lines is higher for HET than WT.

Figure 2-6. Reward-related behaviors. (a) HET mice showed an immediate preference toward 2% sucrose solution. In contrast, no difference was detected in WT between Day1 diluents water and Day2 2% sucrose preference. (b) HET mice learned the incentive licking behavior more rapidly than WT. Significant differences were detected in the second and third blocks of acquisition. There are no differences between genotypes in licking frequency, total solution intake, and licking motor (Figure 2-11). (c) Both WT and HET showed preference to METH paired chamber under the dosage of 1 mg/kg and 2mg/kg. (d) Locomotor activity monitored simultaneously during CPP test. Under the dosage of 2 mg/kg, HET didn't show

comparable increased locomotor activity as WT control. The figure is represented as mean + s.e.m. * $p < .05$.

Figure 2-7. Latent inhibition. There is a significant difference of suppression ratio between pre-exposed and non pre-exposed group in WT mice, indicating normal latent inhibition. Where suppression ratio of HET did not significantly different from pre-exposed and non pre-exposed group, suggesting impaired latent inhibition. The figure is represented as mean + s.e.m. Asterisk: $p < .05$.

Figure 2-8. AKT1 expression in other brain areas. *Akt1* HET mice showed similar reduced expression level in the prefrontal cortex, hippocampus, whole cortex, and the cerebellum.

Figure 2-9. Sensitivity test for assigned prior distribution to μ_α. Beta distribution (2, 5) and (2, 8) were applied to test whether an assigned non-informative uniform distribution can bias the posterior distribution. Beta (2, 5) and (2, 8) showed similar posterior distribution patterns as uniform prior.

Figure 2-10. Sensitivity test for assigned prior distribution to μ_β. Normal distribution (0, 1) was assign to μ_β for sensitivity test. The posterior distribution yields similar distribution patterns as compared with uniform distribution prior. The sensitivity test showed our model and data are robust enough to resist bias from the prior distributions.

Figure 2-11. Licking parameters. No significant differences were detected between genotype in licking frequency (a), total solution consumed (b), and licking motor (d). All $p > .05$ in repeated two-way ANOVA test.

Table 2-1.

Summary of the parameters and the Kolmogorov–Smirnov test results of the estimated distributions of α and β. ***$p < .0001$

α		**mean**	**median**	**0.025**	**0.975**	**sd**	**KS statistic**
Acquisition	WT	0.229	0.213	0.061	0.491	0.112	0.393***
	HET	0.345	0.333	0.135	0.632	0.128	
Reverse	WT	0.049	0.037	0.012	0.150	0.039	0.592***
	HET	0.143	0.114	0.031	0.421	0.102	
METH	WT	0.011	0.010	0.002	0.029	0.007	0.445***
	HET	0.018	0.017	0.006	0.035	0.008	

β		**mean**	**median**	**0.025**	**0.975**	**sd**	**KS statistic**
Acquisition	WT	0.296	0.290	0.132	0.493	0.090	0.485***
	HET	0.220	0.216	0.124	0.343	.054	
Reverse	WT	0.491	0.455	0.021	1.158	0.284	0.46***
	HET	0.246	0.225	-0.001	0.608	0.154	
METH	WT	3.377	3.408	1.795	4.750	0.794	0.653***
	HET	1.991	1.921	0.897	3.461	0.661	

Figure 2-1

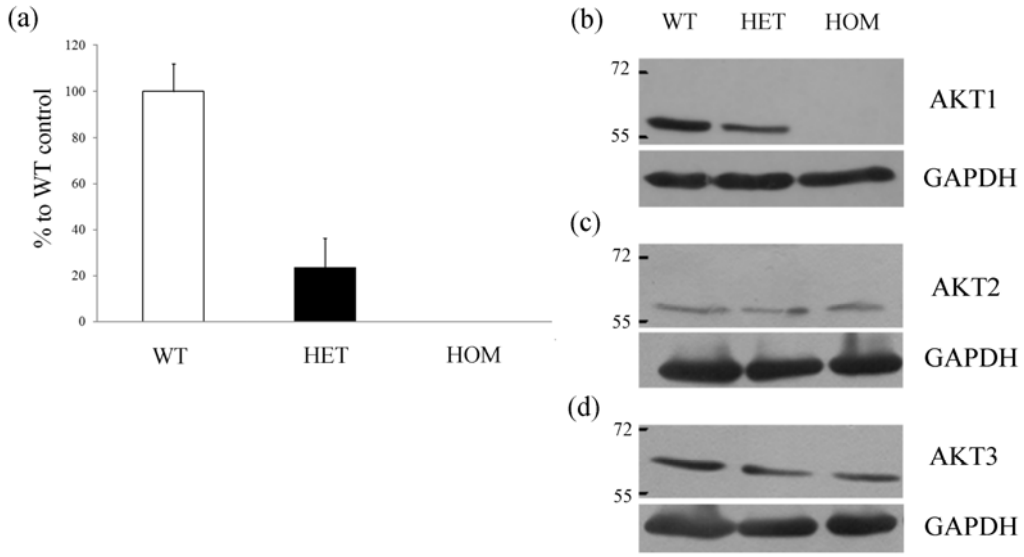

(a)
% to WT control
120
100
80
60
40
20
0
WT
HET
HOM
(b)
WT
HET
HOM
72
55
AKT1
GAPDH
(c)
72
55
AKT2
GAPDH
(d)
72
55
AKT3
GAPDH

Figure 2-2

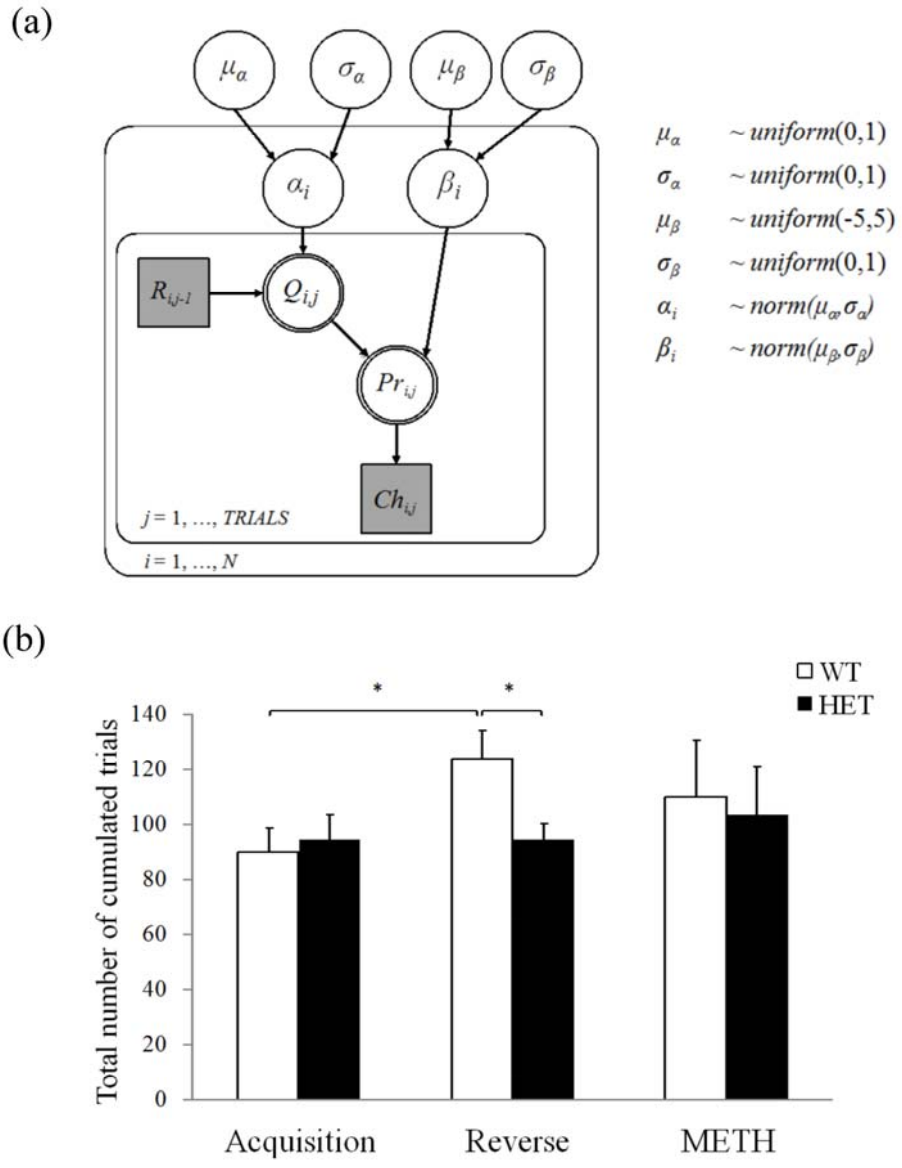

(a)
μ_α
σ_α
μ_β
σ_β
α_i
β_i
$R_{i,j-1}$
$Q_{i,j}$
$Pr_{i,j}$
$Ch_{i,j}$
$j = 1, \ldots, TRIALS$
$i = 1, \ldots, N$
$\mu_\alpha \sim uniform(0,1)$
$\sigma_\alpha \sim uniform(0,1)$
$\mu_\beta \sim uniform(-5,5)$
$\sigma_\beta \sim uniform(0,1)$
$\alpha_i \sim norm(\mu_\alpha, \sigma_\alpha)$
$\beta_i \sim norm(\mu_\beta, \sigma_\beta)$
(b)
WT
HET
*
*
Total number of cumulated trials
140
120
100
80
60
40
20
0
Acquisition
Reverse
METH

Figure 2-3

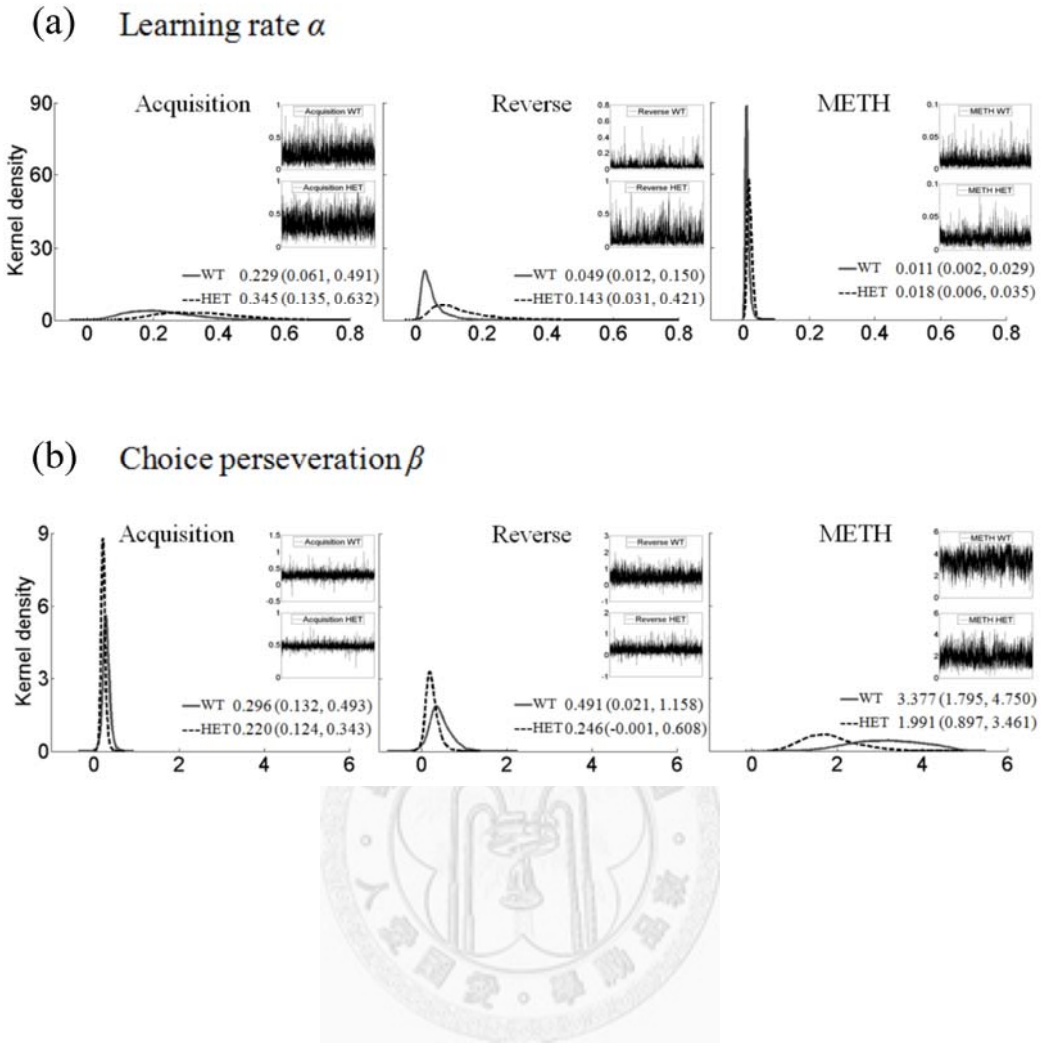

(a) Learning rate α
Acquisition
Reverse
METH
Kernel density
—WT 0.229 (0.061, 0.491)
---HET 0.345 (0.135, 0.632)
—WT 0.049 (0.012, 0.150)
---HET 0.143 (0.031, 0.421)
—WT 0.011 (0.002, 0.029)
---HET 0.018 (0.006, 0.035)
(b) Choice perseveration β
Acquisition
Reverse
METH
Kernel density
—WT 0.296 (0.132, 0.493)
---HET 0.220 (0.124, 0.343)
—WT 0.491 (0.021, 1.158)
---HET 0.246 (-0.001, 0.608)
—WT 3.377 (1.795, 4.750)
---HET 1.991 (0.897, 3.461)

Figure 2-4

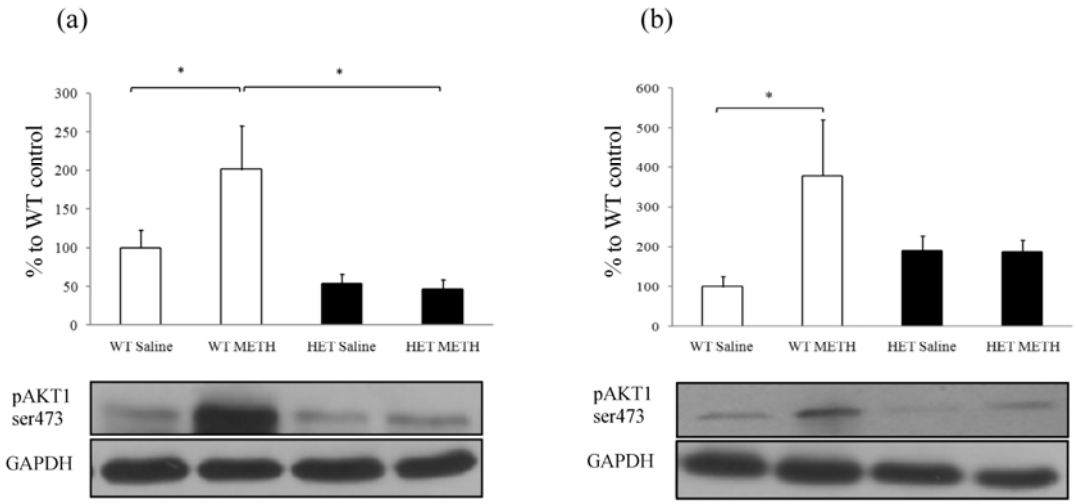

(a)
(b)
% to WT control
WT Saline
WT METH
HET Saline
HET METH
pAKT1
ser473
GAPDH

Figure 2-5

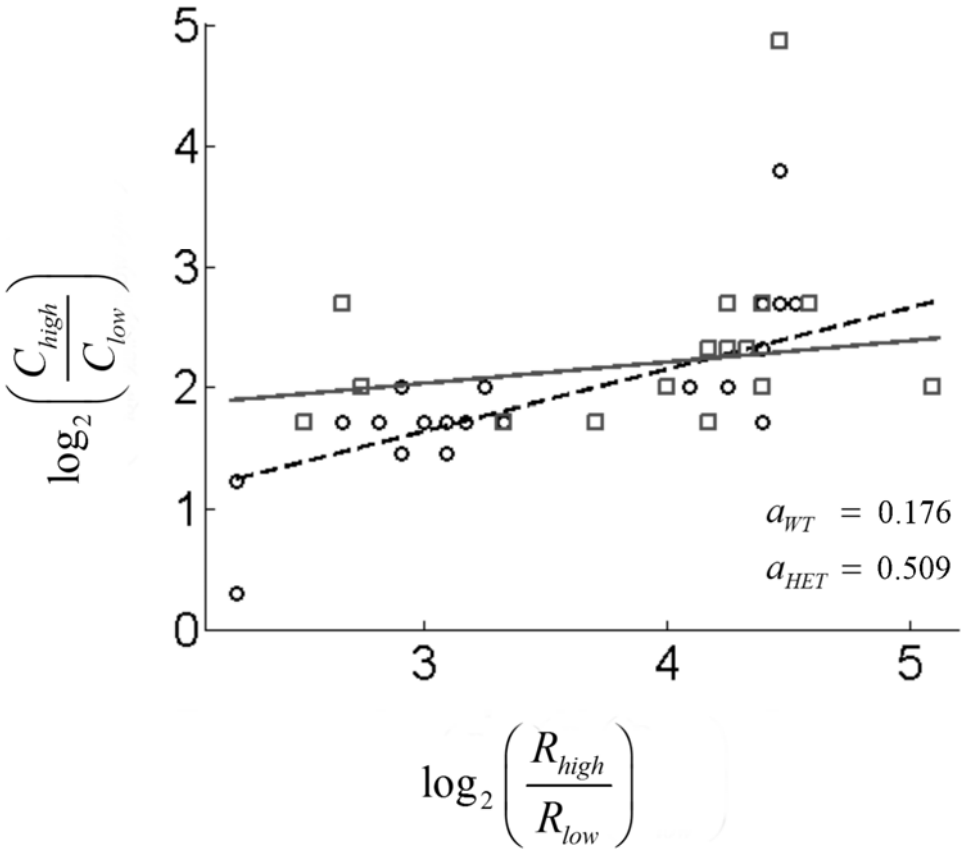
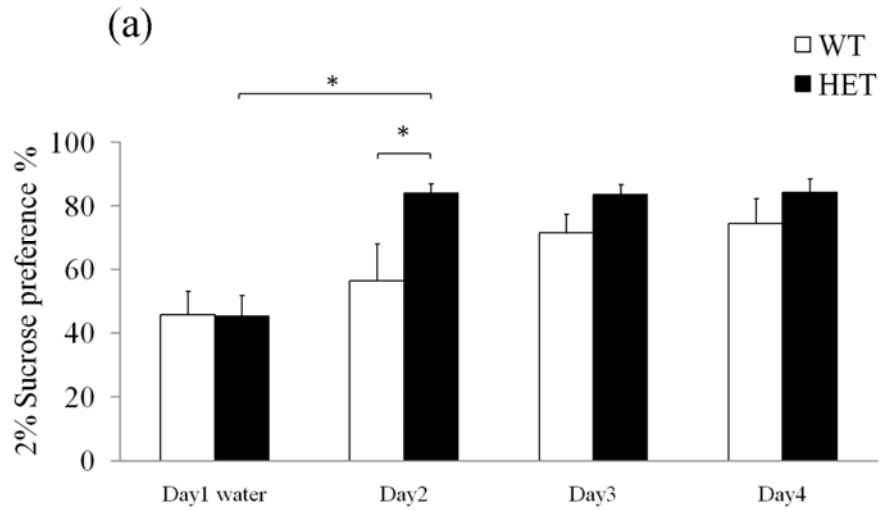

$\log_2\left(\frac{C_{high}}{C_{low}}\right)$
$\log_2\left(\frac{R_{high}}{R_{low}}\right)$
a_{WT} = 0.176
a_{HET} = 0.509
0
1
2
3
4
5
3
4
5

Figure 2-6a, 2-6b



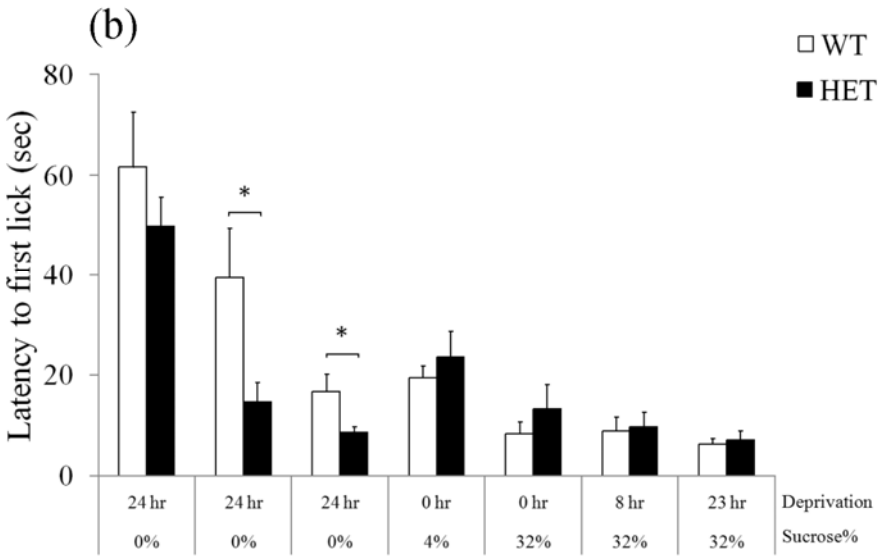


Figure 2-6c, 2-6d

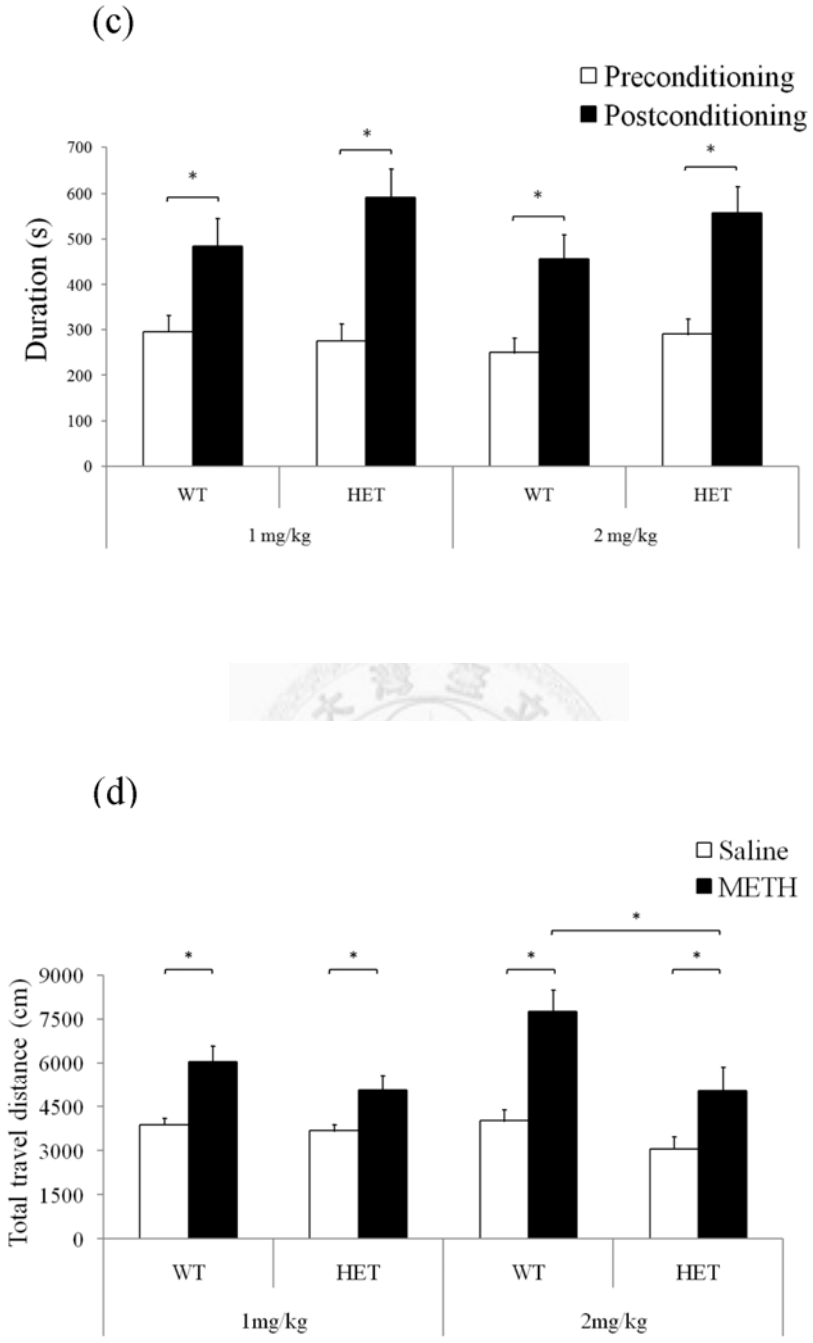


Figure 2-7

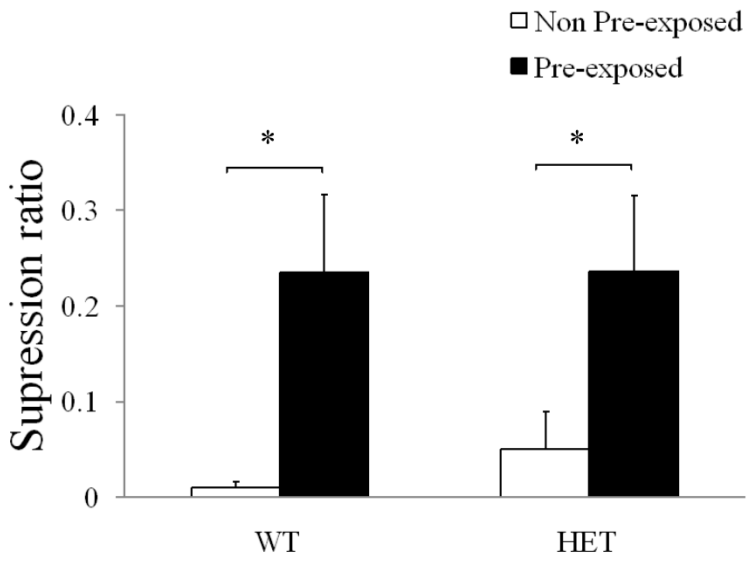

Non Pre-exposed
Pre-exposed
Supression ratio
0.4
0.3
0.2
0.1
0
*
*
WT
HET

Figure 2-8

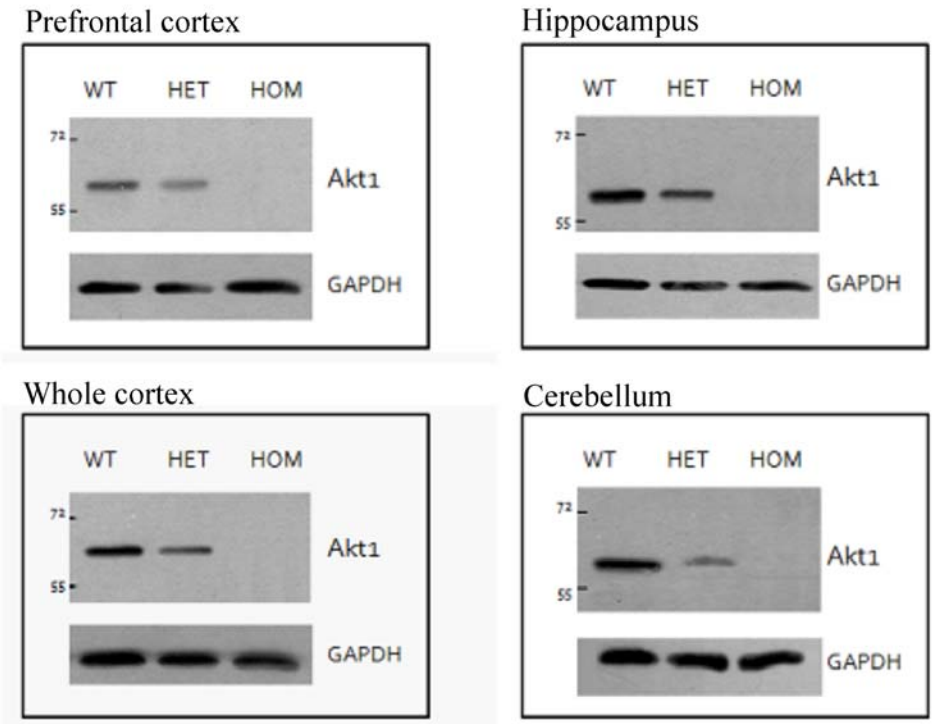

Prefrontal cortex
WT
HET
HOM
72
55
Akt1
GAPDH
Hippocampus
WT
HET
HOM
72
55
Akt1
GAPDH
Whole cortex
WT
HET
HOM
72
55
Akt1
GAPDH
Cerebellum
WT
HET
HOM
72
55
Akt1
GAPDH

Figure 2-9

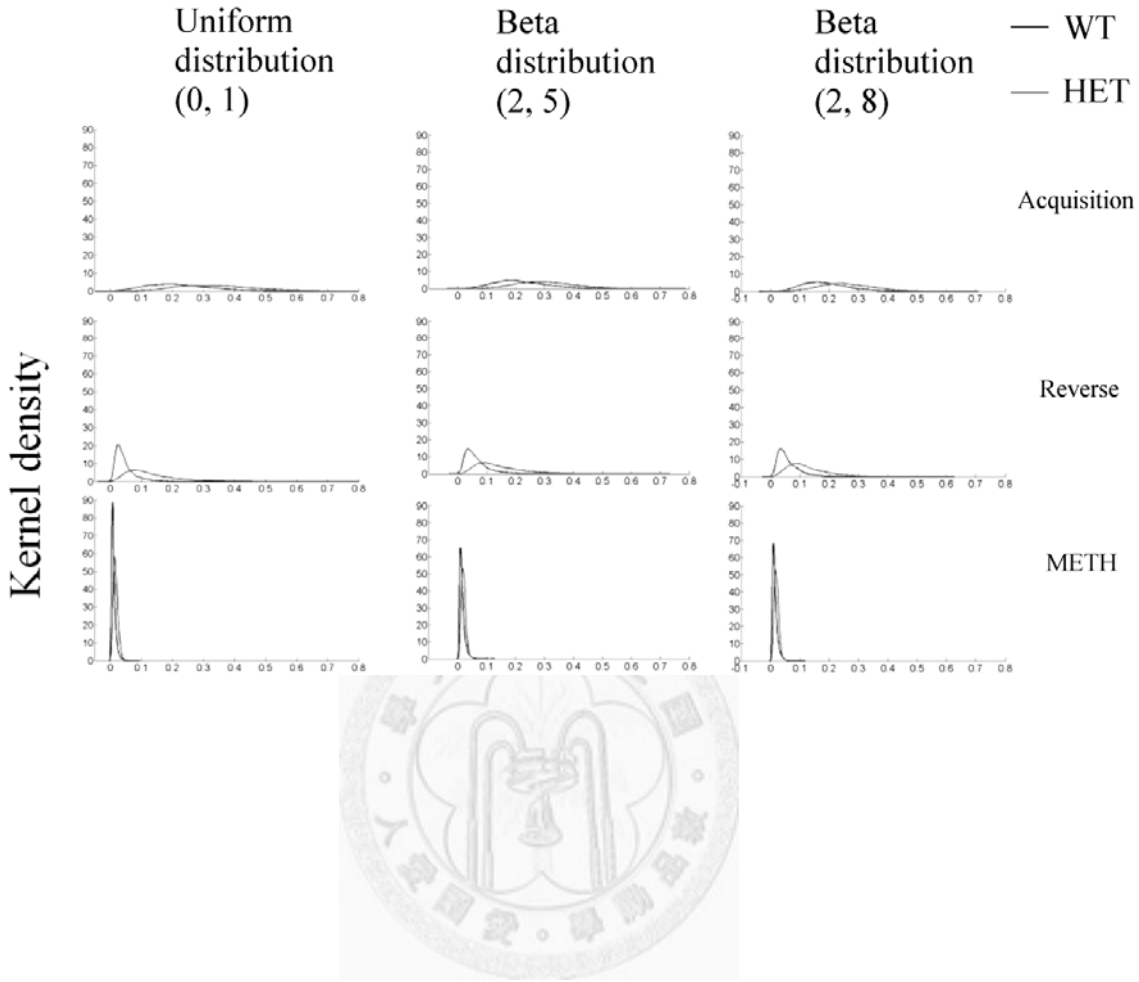

Uniform distribution (0, 1)
Beta distribution (2, 5)
Beta distribution (2, 8)
— WT
— HET
Kernel density
Acquisition
Reverse
METH

Figure 2-10

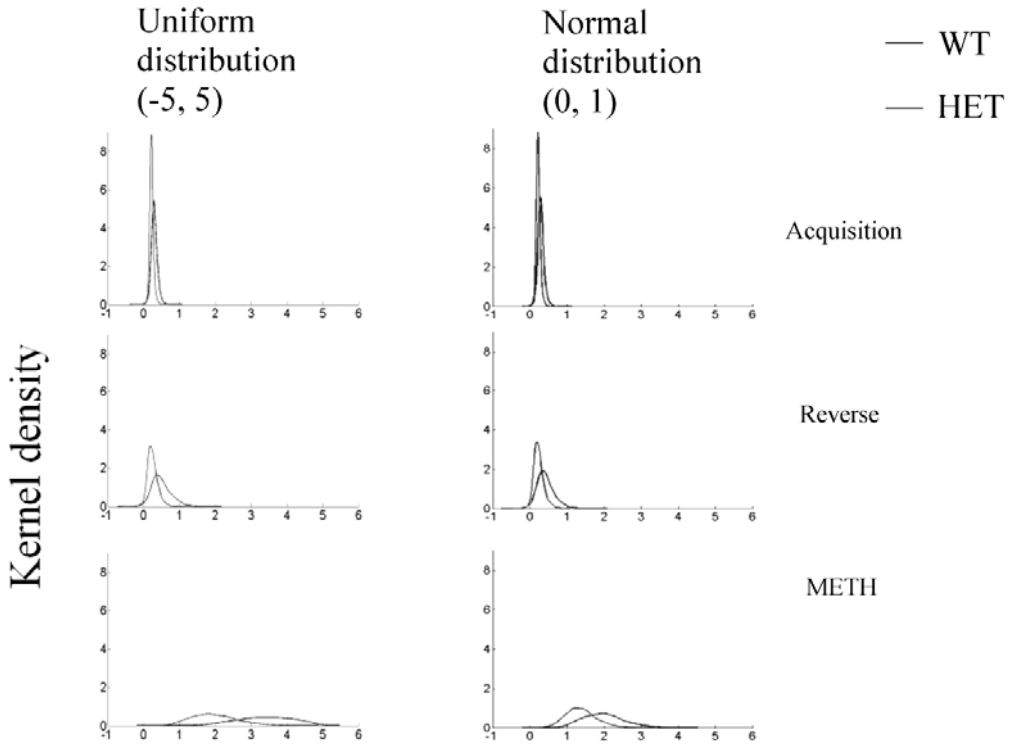
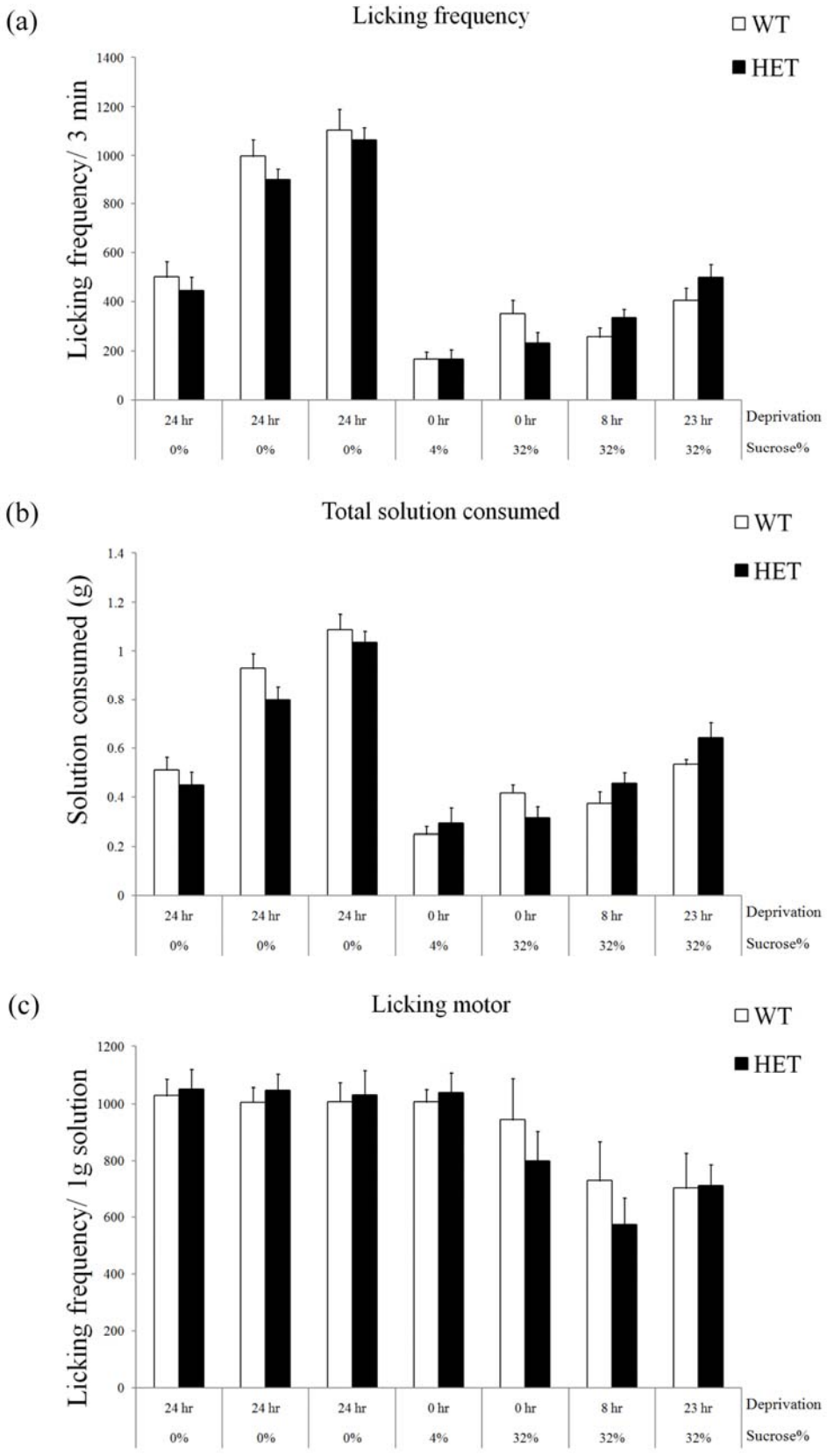

Uniform distribution (-5, 5)
Normal distribution (0, 1)
WT
HET
Kernel density
Acquisition
Reverse
METH

Figure 2-11


(a)
Licking frequency
WT
HET
Licking frequency/ 3 min
1400
1200
1000
800
600
400
200
0
24 hr 24 hr 24 hr 0 hr 0 hr 8 hr 23 hr
Deprivation
0% 0% 0% 4% 32% 32% 32%
Sucrose%
(b)
Total solution consumed
WT
HET
Solution consumed (g)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
24 hr 24 hr 24 hr 0 hr 0 hr 8 hr 23 hr
Deprivation
0% 0% 0% 4% 32% 32% 32%
Sucrose%
(c)
Licking motor
WT
HET
Licking frequency/ 1g solution
1200
1000
800
600
400
200
0
24 hr 24 hr 24 hr 0 hr 0 hr 8 hr 23 hr
Deprivation
0% 0% 0% 4% 32% 32% 32%
Sucrose%

Reference

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage, 31*(2), 790-795.

Baum, W. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*(1), 231.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron, 47*(1), 129-141.

Bayes, T. (1783). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society London 53*, 370-418.

Beaulieu, J. M., Gainetdinov, R. R., & Caron, M. G. (2007). The Akt-GSK-3 signaling cascade in the actions of dopamine. *Trends in Pharmacological Sciences, 28*(4), 166-172.

Beaulieu, J. M., Gainetdinov, R. R., & Caron, M. G. (2009). Akt/GSK3 Signaling in the Action of Psychotropic Drugs. *Annual Review of Pharmacology and Toxicology, 49*, 327-347.

Beaulieu, J. M., Sotnikova, T. D., Marion, S., Lefkowitz, R. J., Gainetdinov, R. R., & Caron, M. G. (2005). An Akt/beta-arrestin 2/PP2A signaling complex mediates dopaminergic neurotransmission and behavior. *Cell, 122*(2), 261-273.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*(9), 1214-1221.

Casella, G., & George, E. (1992). Explaining the Gibbs sampler. *American Statistician, 46*(3), 167-174.

Cho, H., Thorvaldsen, J. L., Chu, Q. W., Feng, F., & Birnbaum, M. J. (2001). Akt1/PKB alpha is required for normal growth but dispensable for maintenance of glucose homeostasis in mice. *journal of Biological Chemistry, 276*(42), 38349-38352.

Cohen, M., Krohn-Grimberghe, A., Elger, C., & Weber, B. (2007). Dopamine gene predicts the brain's response to dopaminergic drug. *European Journal of Neuroscience, 26*(12), 3652-3660.

Cohen, M., Young, J., Baek, J. M., Kessler, C., & Ranganath, C. (2005). Individual differences in extraversion and dopamine genetics predict neural reward responses. *Cognitive Brain Research, 25*(3), 851-861.

Corlett, P., Aitken, M., Dickinson, A., Shanks, D., Honey, G., Honey, R., et al. (2004). Prediction Error during Retrospective Revaluation of Causal Associations in Humans fMRI Evidence in Favor of an Associative Model of Learning.

Neuron, 44(5), 877-888.

Corlett, P., Honey, G., Aitken, M., Dickinson, A., Shanks, D., Absalom, A., et al. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. *Archives of General Psychiatry, 63*(6), 611.

Corlett, P., Murray, G., Honey, G., Aitken, M., Shanks, D., Robbins, T., et al. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*.

Emamian, E. S., Hall, D., Birnbaum, M. J., Karayiorgou, M., & Gogos, J. A. (2004). Convergent evidence for impaired AKT1-GSK3 beta signaling in schizophrenia. *Nature Genetics, 36*(2), 131-137.

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience, 10*(1), 48-58.

Forbes, E., Hariri, A., Martin, S., Silk, J., Moyles, D., Fisher, P., et al. (2009). Altered striatal activation predicting real-world positive affect in adolescent major depressive disorder. *American Journal of Psychiatry, 166*(1), 64.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America, 104*(41), 16311-16316.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47*(4), 495-501.

Franke, T. F. (2008). PI3K/Akt: getting it right matters. *Oncogene, 27*(50), 6473-6488.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris, 100*(1-3), 70-87.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience, 1*(4), 304-309.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis. of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709.

Howes, O. D., & Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III--The Final Common Pathway. *Schizophrenia Bulletin, 35*(3), 549-562.

Jocham, G., Klein, T. A., Neumann, J., von Cramon, D. Y., Reuter, M., & Ullsperger, M. (2009). Dopamine DRD2 Polymorphism Alters Reversal Learning and Associated Neural Activity. *Journal of Neuroscience, 29*(12), 3695-3704.

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive Coding of Action Values in the Human Rostral Cingulate Zone. *Journal of Neuroscience, 29*(23), 7489-7496.

Jonsson, E. G., Nothen, M. M., Grunhage, F., Farde, L., Nakashima, Y., Propping, P., et al. (1999). Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. *Molecular Psychiatry, 4*(3), 290-296.

Juckel, G., Schlagenhauf, F., Koslowski, M., Wustenberg, T., Villringer, A., Knutson, B., et al. (2006). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage, 29*(2), 409-416.

Kaelbling, L., Littman, M., & Moore, A. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research, 4*(237-285), 102-138.

Kapur, S. (2003). Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. *American Journal of Psychiatry, 160*(1), 13-23.

Kapur, S., & Mamo, D. (2003). Half a century of antipsychotics and still a central role for dopamine D-2 receptors. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 27*(7), 1081-1090.

Kapur, S., Zipursky, R., Jones, C., Remington, G., & Houle, S. (2000). Relationship between dopamine D-2 occupancy, clinical response, and side effects: A double-blind PET study of first-episode schizophrenia. *American Journal of Psychiatry, 157*(4), 514-520.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience, 9*(7), 940-947.

Kirsch, P., Reuter, M., Mier, D., Lonsdorf, T., Stark, R., Gallhofer, B., et al. (2006). Imaging gene-substance interactions: The effect of the DRD2 TaqIA polymorphism and the dopamine agonist bromocriptine on the brain activation during the anticipation of reward. *Neuroscience Letters, 405*(3), 196-201.

Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science, 318*(5856), 1642-1645.

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology, 18*(4), 411-417.

Lai, W., Xu, B., Westphal, K., Paterlini, M., Olivier, B., Pavlidis, P., et al. (2006). Akt1 deficiency affects neuronal morphology and predisposes to abnormalities in prefrontal cortex functioning. *Proceedings of the National Academy of Sciences, 103*(45), 16906.

Lee, T., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A, 20*(7), 1434-1448.

Luce, R. (1959). *Individual choice behavior*: Wiley New York.

Lunn, D., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing, 10*(4), 325-337.

Matsumoto, M., & Hikosaka, O. (2008). Representation of negative motivational value in the primate lateral habenula. *Nature Neuroscience, 12*(1), 77-84.

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459*(7248), 837-841.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38*(2), 339-346.

Montague, P., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*(5), 1936.

Montague, P., Hyman, S., & Cohen, J. (2004). Computational roles for dopamine in behavioural control. *Nature, 431*(7010), 760-767.

Murray, G., Corlett, P., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., et al. (2007). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry, 13*(3), 267-276.

Murray, G., Corlett, P., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., et al. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry, 13*(3), 267-276.

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. In B. W. Balleine, K. Doya, J. Odoherty & M. Sakagami (Eds.), *Reward and Decision Making in Corticobasal Ganglia Networks* (Vol. 1104, pp. 35-53). Oxford: Blackwell Publishing.

Roux, S., Froger, C., Porsolt, R. D., Valverde, O., & Maldonado, R. (2003). Place preference test in rodents. *Currrrent Protocol Neuroscience, 9*, 15.

Rushworth, M. F. S. (2008). Intention, choice, and the medial frontal cortex *Year in Cognitive Neuroscience 2008* (Vol. 1124, pp. 181-207). Oxford: Blackwell Publishing.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience, 11*(4), 389-397.

Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic Drugs Modulate Learning Rates and Perseveration

in Parkinson's Patients in a Dynamic Foraging Task. *Journal of Neuroscience, 29*(48), 15104-15114.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science, 310*(5752), 1337.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*(1), 1.

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Schwab, S., Hoefgen, B., Hanses, C., Hassenbach, M. B., Albus, M., Lerer, B., et al. (2005). Further evidence for association of variants in the AKT1 gene with schizophrenia in a sample of European sib-pair families. *Biological Psychiatry, 58*(6), 446-450.

Schwab, S., & Wildenauer, D. (2009). Update on key previously proposed candidate genes for schizophrenia. *Current Opinion in Psychiatry, 22*(2), 147-153

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

Tan, H. Y., Nicodemus, K. K., Chen, Q., Li, Z., Brooke, J. K., Honea, R., et al. (2008). Genetic variation in AKT1 is linked to dopamine-associated prefrontal cortical structure and function in humans. *Journal of Clinical Investigation, 118*(6), 2200-2208.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature, 412*(6842), 43-48.

Watkins, C. (1989). *Learning from delayed rewards.* Cambridge University, Cambridge.

Zhao, Z., Ksiezakreding, H., Riggio, S., Haroutunian, V., & Pasinetti, G. (2006). Insulin receptor deficits in schizophrenia and in cellular and animal models of insulin receptor dysfunction. *Schizophrenia Research, 84*(1), 1-14.

Chapter 3

General Discussion

In this thesis, we applied a dynamic foraging T-Maze task for mice to learn the value of their action by trial-and-error. Mice readily learned to allocate their choice to the higher reward rate arm, and adjust their choice to unsignaled change to reward probability. This action value is updated by the discrepancy between the reward received and the reward expected according to the RL model. Therefore, we fit the choice data to a standard *Q* learning model to obtain estimation about the RPE. Evidences from human and monkey studies show that trial-by-trial action values estimated from choice behavior by reinforcement learning model is correlated with the neuronal activity in the striatum, where it receives dense RPE input encoded by DA neurons (Samejima, Ueda, Doya, & Kimura, 2005; Tanaka, et al., 2004). Accordingly, it is reasonable to infer that AKT1 deficiency leads up to the heightened RPE through its interaction with DA system. Our study also implies that such unusual heightened RPE signals could contribute, at least in part, to the vulnerability to develop psychosis, since AKT1 plays a critical role in the neuroleptic receptor DRD2 and in the action and mood stabilizer (Beaulieu, Gainetdinov, & Caron, 2007, 2009). With overwhelming number of schizophrenia candidate genes in hand, the oncoming

challenge is to elucidate how these genes contribute to the pathogenesis of schizophrenia. It is also critical to link biological functions of each proven gene to the clinical symptoms. The current research can be considered as an example and it provides important insights to the role AKT1 in the psychosis of schizophrenia. It is worth further applying antipsychotic drugs in a value-learning task to see whether this augmented RPE could be rescued.

In addition to the discussion in chapter 2, several issues raised in this research are discussed as following:

I. AKT pathway and RPE

An immediate question follows the present study is: how does an intracellular protein contributes to a computational circuitry in the brain? Because AKT/GSK3 pathway is considered as a relatively slow responding intracellular signaling cascade (Beaulieu, et al., 2007), therefore it may be unthinkable to participate in a rapid computation process. Instead, evoked intracellular signaling often results in changes of cellular physiology. It is probable to assume that neurons with AKT1 deficiency, such as the medium spiny neurons in the striatum, behave differently in response to DA innervations, and thus cause changes in the overall function of the circuitry. Concomitantly with the notion above, after antipsychotics treatment, it usually takes

2-14 days for the positive symptoms to mitigate (Catafau, et al., 2006; Grace, Bunney, Moore, & Todd, 1997), suggesting that there might be some relatively endured physiological changes caused by the DRD2 signaling and its downstream signaling, like AKT1. Future works dealing with this question should apply conditional knockout mice or RNA interference techniques along with electrophysiological recording at the basal ganglia or at the ventral tegmental area and substantial nigra to verify the role of AKT1 in the RPE-processing circuitry.

II. RPE in patients with Parkinson's disease

Rutledge et al. (Rutledge, et al., 2009) demonstrated Parkinsonic patients on L-dopa medication exhibit higher learning rate and RL model. The authors provided evidence showing that promoting DA biosynthesis could lead up to subsequent intensification of the RPE magnitude. This result is of particular interest because nearly half of the Parkinsonic patients on chronic L-dopa medication have reported psychosis symptoms (Moskovitz, Moses, & Klawans, 1978). Such finding provides a hint that perhaps augmented RPE magnitude in Parkinson's patients has something to do with the inclination to develop psychosis. This subtle connection should be investigated in further detail. In addition, Parkinsonic patients on dopaminergic medication also display higher reward sensitivity in matching law analysis. It is also

similar to our findings in which *Akt1* HET mice exhibits higher reward sensitivity, implying that they allocate more choice to the arm with higher reward probability. Nonetheless, this is highly expectable because a large amount of researches indicated that RPE is implicated in decision making in human (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Egelman, Person, & Montague, 1998; Holroyd & Coles, 2002; Morris, Nevet, Arkadir, Vaadia, & Bergman, 2006; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006) and animal (Bromberg-Martin & Hikosaka, 2009; Matsumoto & Hikosaka, 2009; Roesch, Calu, & Schoenbaum, 2007).

III. AKT1 and choice behavior

Our study also raised some interesting findings on choice behaviors. The inverse temperature, in terms of Boltzmann exploration, or the choice perseveration parameter β is lower in HET than WT. This is similar to the Rutledge et al. (2009) study, in addition to higher learning rate, lower choice perseveration was also found in patients with Parkinson's disease on L-dopa medication. This parameter leaves some room for interpretations about the strategy of decision making in a subject. Relatively lower β suggests the choice behavior of HET mice is more flexible than WT controls, which is probably capable to explain the lower number of trials that HET needs to complete the criterion in the reverse section of the T-Maze task. This might also provide hints into

the relationship between AKT1 and the course to develop habits (Jog, Kubota, Connolly, Hillegaart, & Graybiel, 1999). Further works are needed. It is advisable to apply different RL algorithms to differentiate habit learning and goal-directed learning (Balleine, Daw, & O'Doherty, 2008).

IV. Methamphetamine and RPE

We found that instead of increasing learning rate, METH decreases it. This is consistent with the original notion that phasic DA firing encodes the RPE signal, because METH seems to dampen the phasic DA neuron firing. METH exerts its effect by blocking the DA transporters, and then increases the DA concentration in the synaptic cleft. By interfering reuptake, the effect of phasic DA firing is prolonged initially. However, DA is then allowed to diffuse out of the synaptic cleft to act upon presynaptic autoreceptor, where DA synthesis and releasing will be down regulated. Such effect then causes an increase in tonic extrasynaptic DA level (Di Chiara & Imperato, 1988; Hurd & Ungerstedt, 1989). As a result, the increased tonic DA release resulted from stimulants will cause a decrease in phasic DA release (Grace, 1995). Thus, the decreased phasic DA release should manifest itself in lower RPE magnitude, that is, lower learning rate.

V. Animal model of psychosis

Because schizophrenia encompass various aspects of symptoms, including positive, negative, disorganized symptoms, and cognitive deficits, it is difficult to model the full phenotypes in any model system. What's more, according to common disease/common allele hypothesis of schizophrenia, common variants in more than one susceptibility gene, each contributes a small effect and acts in combinations to increase the risk of illness. Therefore, the purpose of building a genetic animal model of schizophrenia is to study the causal-relationship between a genotype and a specific endophenotype (Walters & Owen, 2007). Current strategy in modeling schizophrenia in animals targets separately to positive symptoms (Geyer, Krebs-Thomson, Braff, & Swerdlow, 2001), negative symptoms (Ellenbroek & Cools, 2000), and cognitive deficits (Kellendonk, Simpson, & Kandel, 2009).

There are plenty good behavioral paradigms to model negative symptoms and cognitive deficits in light of the previous researches in the field of behavioral neuroscience. However, the positive symptoms, or psychosis, are the most difficult one to model in animal. It is obvious, because positive symptoms are characterized by hallucination and paranoid delusion. Nevertheless, efforts are made, stress or novelty-induced hyperactivity and hypersensitivity to psychostimulants have been suggested as useful correlates that can be modeled in rodents and have been

extensively used in the validation and assessment of pharmacological models (Geyer & Moghaddam, 2002). The original impetus for this hyperlocomotor activity model was derived from the psychostimulant models of amphetamine psychosis (Snyder, 1973). However, it still falls short of decent explanation to interpret how excess DA leads up to the bizarre experiences in patients.

A better attempt incorporates the idea of sensory motor gating, which assumed that impairments in either filtering or gating lead to sensory overload and cognitive fragmentation. The validity of this gating construct has been assessed most extensively by means of an operational measure based on cross-species homologies in the startle reflex—that is, the prepulse inhibition (PPI) of startle paradigm (Braff & Geyer, 1990; Braff, Swerdlow, & Geyer, 1999; Light & Braff, 2003). In addition, PPI has an analogous measure of event-related potentials component called P50, or N40 in rodents (Stevens, et al., 1996). Schizophrenic patients and their first-degree relatives exhibit less sensory gating (Bramon, Rabe-Hesketh, Sham, Murray, & Frangou, 2004; Clementz, Geyer, & Braff, 1998; Freedman, et al., 1994).

The notion of present study provides an additional possibility to model psychosis in animal. Accumulating evidences from human studies have shown that RPE disruption is evident in patients with psychosis. Therefore, measuring RPE in animal could capture the fundamental nature of the symptoms. Unlike other schizophrenia

model, RPE model is grounded on strong cognitive neuroscience knowledge and clear neural mechanisms, therefore it can provide explanation from neurotransmission level to the phenomenological level. The RPE model is better than the gating theory because it fits into a general description of the brain's function—Bayesian inference, instead of simply explaining psychosis with overloaded sensory input. In the framework that brain functions as a Bayesian machine, the role of the prediction error signal plays as a marker that the existing model or inference of the world has not fully accounted for the input. If this marker is disrupted, the system can not be adjusted, and the internal representation of the world will become unrealistic. Though sounds promising, this Bayesian view of schizophrenia is a relatively new hypothesis about how DA system contributes to the pathogenesis of psychosis, plenty of works need to be done in order to establish the legitimacy of RPE disruption as an animal model of psychosis.

Reference

Balleine, B., Daw, N., & O'Doherty, J. (2008). Multiple forms of value learning and the function of dopamine. In P. Glimcher, E. Fehr, C. Camerer & R. Poldrack (Eds.), *Neuroeconomics: decision-making and the brain* (pp. 367-385): Academic Press.

Beaulieu, J. M., Gainetdinov, R. R., & Caron, M. G. (2007). The Akt-GSK-3 signaling cascade in the actions of dopamine. *Trends in Pharmacological Sciences, 28*(4), 166-172.

Beaulieu, J. M., Gainetdinov, R. R., & Caron, M. G. (2009). Akt/GSK3 Signaling in the Action of Psychotropic Drugs. *Annual Review of Pharmacology and Toxicology, 49*, 327-347.

Braff, D., & Geyer, M. (1990). Sensorimotor gating and schizophrenia: Human and animal model studies. *Archives of General Psychiatry, 47*(2), 181-188.

Braff, D., Swerdlow, N., & Geyer, M. (1999). Symptom correlates of prepulse inhibition deficits in male schizophrenic patients. *American Journal of Psychiatry, 156*(4), 596.

Bramon, E., Rabe-Hesketh, S., Sham, P., Murray, R., & Frangou, S. (2004). Meta-analysis of the P300 and P50 waveforms in schizophrenia. *Schizophrenia Research, 70*(2-3), 315-329.

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. *Neuron, 63*(1), 119-126.

Catafau, A., Corripio, I., Perez, V., Martin, J., Schotte, A., Carrio, I., et al. (2006). Dopamine D2 receptor occupancy by risperidone: Implications for the timing and magnitude of clinical response. *Psychiatry Research: Neuroimaging, 148*(2-3), 175-183.

Clementz, B., Geyer, M., & Braff, D. (1998). Poor P50 suppression among schizophrenia patients and their first-degree biological relatives. *American Journal of Psychiatry, 155*(12), 1691.

Daw, N., O'Doherty, J., Dayan, P., Seymour, B., & Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441*(7095), 876-879.

Di Chiara, G., & Imperato, A. (1988). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences, 85*(14), 5274.

Egelman, D., Person, C., & Montague, P. (1998). A computational role for dopamine

delivery in human decision-making. *Journal of Cognitive Neuroscience, 10*, 623-630.

Ellenbroek, B. A., & Cools, A. R. (2000). Animal models for the negative symptoms of schizophrenia. *Behavioural Pharmacology, 11*(3 & 4), 223-233.

Freedman, R., Adler, L., Bickford, P., Byerley, W., Coon, H., Cullum, C., et al. (1994). Schizophrenia and nicotinic receptors. *Harvard Review of Psychiatry, 2*(4), 179-192.

Geyer, M., Krebs-Thomson, K., Braff, D., & Swerdlow, N. (2001). Pharmacological studies of prepulse inhibition models of sensorimotor gating deficits in schizophrenia: a decade in review. *Psychopharmacology, 156*(2-3), 117-154.

Geyer, M., & Moghaddam, B. (2002). Animal models relevant to schizophrenia disorders *Neuropsychopharmacology: The Fifth Generation of Progress* (pp. 689-701): Lippincott Williams & Wilkins.

Grace, A. (1995). The tonic/phasic model of dopamine system regulation: its relevance for understanding how stimulant abuse can alter basal ganglia function. *Drug and Alcohol Dependence, 37*(2), 111-129.

Grace, A., Bunney, B., Moore, H., & Todd, C. (1997). Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. *Trends in Neurosciences, 20*(1), 31-37.

Holroyd, C., & Coles, M. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709.

Hurd, Y., & Ungerstedt, U. (1989). In vivo neurochemical profile of dopamine uptake inhibitors and releasers in rat caudate-putamen. *European journal of pharmacology, 166*(2), 251-260.

Jog, M., Kubota, Y., Connolly, C., Hillegaart, V., & Graybiel, A. (1999). Building neural representations of habits. *Science, 286*(5445), 1745.

Kellendonk, C., Simpson, E., & Kandel, E. (2009). Modeling cognitive endophenotypes of schizophrenia in mice. *Trends in Neurosciences, 32*(6), 347-358.

Light, G., & Braff, D. (2003). Sensory gating deficits in schizophrenia: can we parse the effects of medication, nicotine use, and changes in clinical status? *Clinical Neuroscience Research, 3*(1-2), 47-54.

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459*(7248), 837-841.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience,*

9(8), 1057-1063.

Moskovitz, C., Moses, H., 3d, & Klawans, H. (1978). Levodopa-induced psychosis: a kindling phenomenon. *American Journal of Psychiatry, 135*(6), 669-675.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R., & Frith, C. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042-1045.

Roesch, M., Calu, D., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience, 10*(12), 1615-1624.

Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic Drugs Modulate Learning Rates and Perseveration in Parkinson's Patients in a Dynamic Foraging Task. *Journal of Neuroscience, 29*(48), 15104-15114.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science, 310*(5752), 1337.

Snyder, S. (1973). Amphetamine psychosis: a" model" schizophrenia mediated by catecholamines. *American Journal of Psychiatry, 130*(1), 61.

Stevens, K., Freedman, R., Collins, A., Hall, M., Leonard, S., Marks, M., et al. (1996). Genetic correlation of hippocampal auditory evoked response and alpha-bungarotoxin binding in inbred mouse strains. *Neuropsychopharmacology, 15*, 152-162.

Tanaka, S., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience, 7*(8), 887-893.

Walters, J., & Owen, M. (2007). Endophenotypes in psychiatric genetics. *Molecular Psychiatry, 12*(10), 886-890.

Appendix

Chen, Y. C., & Lai, W. S. (2010). Behavioural phenotyping of dopamine transporter knockdown mice using local small interference RNA. *Behavioural Brain Research*, (in press).