Behavioural Brain Research 335 (2017) 136–144

Contents lists available at ScienceDirect

Behavioural Brain Research

journal homepage: www.elsevier.com/locate/bbr

Research report

Acute Δ-9-tetrahydrocannabinol administration in female rats attenuates immediate responses following losses but not multi-trial reinforcement learning from wins

Scott A. Wong, Sienna H. Randolph, Victorita E. Ivan, Aaron J. Gruber*

Canadian Centre for Behavioural Neuroscience, Department of Neuroscience, University of Lethbridge, Lethbridge, AB, Canada

ARTICLE INFO

Keywords:
Dopamine
Cognitive flexibility
THC
Striatum
Basal ganglia
Female

ABSTRACT

Δ-9-Tetrahydrocannabinol (THC) is the main psychoactive component of marijuana and has potent effects on decision-making, including a proposed reduction in cognitive flexibility. We demonstrate here that acute THC administration differentially affects some of the processes that contribute to cognitive flexibility. Specifically, THC reduces lose-shift responding in which female rats tend to immediately shift choice responses away from options that result in reward omission on the previous trial. THC, however, did not impair the ability of rats to flexibly bias responses toward feeders with higher probability of reward in a reversal task. This response adaptation developed over several trials, suggesting that THC did not impair slower forms of reinforcement learning needed to choose among options with unequal utility. This dissociation of THC's effects on innate/rapid and learned/gradual decision-making processes was unexpected, but is supported by emerging evidence that lose-shift responding is mediated by neural mechanisms distinct from those involved in other forms of reinforcement learning. The present data suggest that, at least in some tasks, the apparent reductions in cognitive flexibility by THC may be explained by the immediate effects on loss sensitivity, rather than impairments of all processes used for choice adaptation.

1. Introduction

THC and other components in marijuana are among the most commonly used illicit drugs in the world. A recent report [1] indicates over 181 million users world-wide, while an estimated 13.1 million people exhibit cannabis dependence globally [2]. Marijuana usage is increasing in many countries, [3] and chronic use appears to be a risk factor for several mental illnesses including psychosis [4] and depression [5]. Moreover, chronic use is also linked to reduced performance on decision-making tasks [6,7]. It is therefore important to better understand how THC affects the neural processes involved in cognition and decision-making. Previous studies have indicated deficits in executive function, verbal and visual memory, and visuoperception following chronic cannabis use in humans [8]. Rodent models have likewise shown impairments in working memory [9,10] and spatial memory [11] following acute THC administration. Furthermore, acute THC causes impairments in reversal learning in macaques [12], as well as reversal learning and intradimensional set shifting impairments in rats [13,14]. These findings have led to the proposal that THC impairs cognitive flexibility. This is a nebulous term; for the purpose of this paper we define cognitive flexibility as the ability to change response strategies when it is advantageous to do so. Cognitive flexibility is influenced by distinct neural processes that may compete or cooperate to modulate behaviour [15]. For example, animals can learn to select options based on action value estimates acrued over many trials (reinforcement learning), or could instead use heuristics to guide choice [16]. Animals may also rely on innate strategies such as the propensity to switch choices after reward omission, a widespread phenomenon termed lose-shift responding [17–20]. These and other systems can be used to derive choice, and can all manifest as cognitive flexibility.

A common methodological feature in reversal learning and set shifting paradigms is a sudden and unexpected change in reward contingencies, which is used to assess how animals adapt choice strategy. Both lose-shift and reinforcement learning strategies are likely engaged in these paradigms. A physiological mechanism thought to be important for such reinforcement-driven response adaptations is the reward prediction error (RPE) signaling properties of midbrain dopamine neurons [21–23]. These neurons briefly increase firing rate following unexpectedly good reinforcements (rewards), signaling a positive RPE. Conversely, these neurons decrease firing following unexpectedly poor

* Corresponding author at: University of Lethbridge, Department of Neuroscience, Canadian Centre for Behavioural Neuroscience, Polaris Brain Dynamics Research Group, 4401 University Drive West, Lethbridge, AB, Canada.
E-mail address: aaron.gruber@uleth.ca (A.J. Gruber).

http://dx.doi.org/10.1016/j.bbr.2017.08.009
Received 12 May 2017; Received in revised form 28 July 2017; Accepted 5 August 2017
Available online 12 August 2017

reinforcements such as reward omission, signaling a negative RPE. This RPE signal is sufficient to solve a variety of difficult tasks [24]. Drugs or diseases that affect dopamine transmission are therefore expected to impair this prediction error, and thereby impair reinforcement learning, which has been demonstrated in many species and tasks [25–28].

Similar to other drugs of abuse, THC increases the firing rates of midbrain dopamine neurons [29] and increases dopamine release in downstream targets such as the ventral striatum [30]. THC acts as an agonist at the Cannabinoid Receptor 1 (CB1), which is G_i coupled and acts as an inhibitory presynaptic regulator of neurotransmitter release [31]. CB1 is located on inhibitory afferents to dopamine neurons [32]. It is therefore likely that CB1 agonism reduces inhibitory input to dopamine neurons. We expect that this will suppress the pause in dopamine neuron firing following reward omission and thereby attenuate the negative RPE. This should attenuate processes driven by reward omission, such as lose-shift responding and reinforcement learning from worse-than-expected reinforcements. We have previously shown that lose-shift is attenuated by acute administration of amphetamine, which increases striatal dopamine among other effects [26]. We expected that THC would have a similar attenuating effect on lose-shift due to reduction of negative RPE. Furthermore, we expect that THC would impair reinforcement learning from reward omission for the same reason. Learning from positive reinforcements, however, should remain intact [25].

To test this hypothesis, we systemically injected rats with THC and analyzed specific features of reinforcement-driven response adaptation in two reward-based decision-making tasks. The first was a binary choice task that rewards random choice; rats nonetheless employ a win-stay/lose-shift strategy [33]. The second was an uncued reversal task that can be solved using reinforcement learning. These tasks allow us to differentiate between the potential effects of drug on lose-shift responding that occurs on a trial-by-trial basis, and on the multi-trial learning required to track reversals of asymmetric reward probabilities. We found that systemic THC administration attenuated lose-shift responding on both tasks. However, the drug did not impair the ability to flexibly reverse choice preference in response to uncued reversals of reward probability. These results suggest that THC causes an apparent reduction in cognitive flexibility by impairing trial-by-trial loss sensitivity while sparing slower multi-trial learning from wins. This dissociation is supported by recent computational analyses of rodent and human choices in decision-making tasks, which have demonstrated that hybrid models with rapid and gradual components more accurately capture the choices of humans and rodents than do models with a unitary timeframe [34–37].

2. Methods

2.1. Animals

Subjects were 21 adult (90 days old) female Long-Evans rats (bred in-house) weighing 200–250 g. Animals were pair-housed in a climate-controlled vivarium under a 12:12 h light:dark cycle (lights on 7:30 a.m.). Animals had restricted access to water (one hour) on behavioural testing days, but otherwise had ad libitum access to food and water. All procedures were approved by the University of Lethbridge Animal Welfare Committee (AWC) in accordance with the guidelines of the Canadian Council on Animal Care. Our utilization of female rats fills knowledge gaps and is encouraged by the AWC.

2.2. Behaviour apparatus

Behavioural testing was performed in aluminum operant conditioning chambers (see Fig. 1) as described previously [33,36]. Briefly, rats were placed in the operant conditioning chamber for 45 min sessions. Trials were self-paced, and initiated by the rat performing a nose-poke into the central port. Following 150 ms of nose-poke entry, a tone (6 kHz) was presented to indicate that the animal could then locomote to one of the two adjacent sucrose delivery feeders. If the correct feeder was chosen, a reward (60 μL of 10% sucrose solution) was delivered. If the incorrect feeder was chosen, no sucrose was delivered, the house-light illuminated, and the two panel lights extinguished. The state of the lights then reverted (house-light turned off; panel light turned on). This change in lighting served to indicate that reward was not forthcoming, and was of sufficiently short duration such that it terminated by the time the rats returned to the central poke port; there was therefore no 'time-out' associated with reward omission. Once a feeder was chosen, or if no feeder was chosen in the 15 s following a nose-poke, the trial ended and the rat had to return to the central port to initiate a new trial.

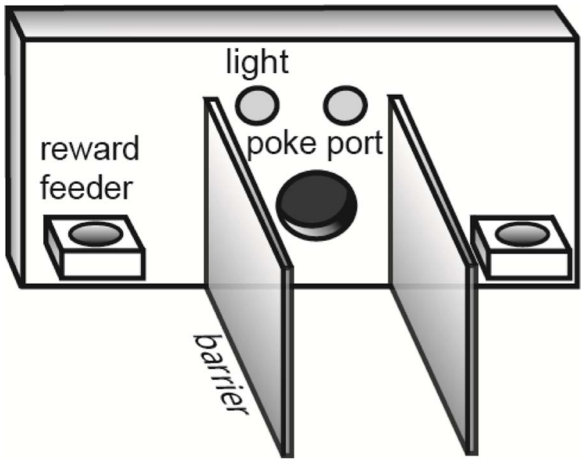

Fig. 1. Schematic diagram of the operant conditioning chamber used for behavioural tasks.

2.3. Experiment 1: acute effects of THC on the Competitive Choice Task (CCT)

The behaviour of animals in the first cohort (n = 10) was shaped during the first two training sessions. All trials were rewarded and no barriers were present in the first training session to facilitate task acquisition. In the second training session, rats were rewarded on 50% of the trials regardless of feeder choice. In all subsequent sessions, reinforcement was controlled by an algorithm that attempted to minimize the number of rewards given to the rats by predicting which feeder the rat would select. This was done by examining the choices and reinforcements from the previous four trials [20,36]. If either feeder was selected at a greater than chance rate in the context of these past trials, it would be unrewarded for the upcoming trial. In doing so, the competitive mode implements the classic 'Matching Pennies' task. Optimal performance (random responding) will result in reward on 50% of the trials. Parallel barriers positioned between the central port and feeder wells were added to introduce a choice cost and discourage feeder bias due to body position. Increasingly longer barriers (4.0, 8.5, 13.5 cm) were introduced during consecutive days of training. Rats were trained until they completed two consecutive sessions of at least 150 trials with the long barriers. All subsequent training and testing sessions were run with the long barriers.

After initial shaping (9 daily sessions), animals were randomly divided into four groups to receive acute THC (Cayman Chemicals, Ann Arbour, MI) in a counterbalanced block design. The drug was dissolved into a 1:1:1:16 solution of THC:ethanol:Cremaphor EL:sterile saline (0.9%) and delivered by intraperitoneal (IP) injection at one of three dosages (0.5 mg/kg, 1.0 mg/kg, or 2.0 mg/kg). Injections were administered 30 min prior to testing on the behavioural task over a period of 8 days using the following schedule: vehicle, injection 1, no injection, injection 2, no injection, injection 3, no injection, and injection 4. The initial vehicle injection served to habituate animals to the