Behavioural Brain Research 296 (2016) 240–248

Contents lists available at ScienceDirect

Behavioural Brain Research

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

Research report

Probabilistic reward- and punishment-based learning in opioid addiction: Experimental and computational data

Catherine E. Myers[a,b], Jony Sheynin[c,d], Tarryn Balsdon[e], Andre Luzardo[f], Kevin D. Beck[a,b], Lee Hogarth[g,h], Paul Haber[i], Ahmed A. Moustafa[e,j,*]

[a] Department of Veterans Affairs, New Jersey Health Care System, East Orange, NJ, USA
[b] Department of Pharmacology, Physiology & Neuroscience, New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, NJ, USA
[c] Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, MI, USA
[d] Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA
[e] School of Social Sciences and Psychology, University of Western Sydney, Sydney, NSW, Australia
[f] School of Mathematics, Computing Sciences & Engineering at City University London, UK
[g] School of Psychology, University of New South Wales, Sydney, NSW, Australia
[h] School of Psychology, University of Exeter, Exeter, UK
[i] Drug Health Services, Addiction Medicine, Central Clinical School, Royal Prince Alfred Hospital, The University of Sydney, Sydney, NSW, Australia
[j] Marcs Institute for Brain and Behaviour, University of Western Sydney, Sydney, NSW, Australia

HIGHLIGHTS

- Opioid-addicted individuals, and controls, performed a reward- and punishment-learning task.
- Computational (reinforcement learning) models were applied to describe individuals' performance.
- Behavioral results shows opioid-addicted individuals performed as well as controls on the task.
- Computational modeling suggested subtle differences in how the two groups made decisions.
- Specifically, the addicted group was more likely to "chase" reward when expectancies were violated.
- A bias to pursue short-term reward, rather than long-term gain, may contribute to opioid addiction.

ARTICLE INFO

Article history:
Received 26 February 2015
Received in revised form 7 September 2015
Accepted 11 September 2015
Available online 14 September 2015

Keywords:
Addiction
Reward learning
Punishment learning

ABSTRACT

Addiction is the continuation of a habit in spite of negative consequences. A vast literature gives evidence that this poor decision-making behavior in individuals addicted to drugs also generalizes to laboratory decision making tasks, suggesting that the impairment in decision-making is not limited to decisions about taking drugs. In the current experiment, opioid-addicted individuals and matched controls with no history of illicit drug use were administered a probabilistic classification task that embeds both reward-based and punishment-based learning trials, and a computational model of decision making was applied to understand the mechanisms describing individuals' performance on the task. Although behavioral results showed that opioid-addicted individuals performed as well as controls on both reward- and punishment-based learning, the modeling results suggested subtle differences in how decisions were made between the two groups. Specifically, the opioid-addicted group showed decreased tendency to repeat prior responses, meaning that they were more likely to "chase reward" when expectancies were violated, whereas controls were more likely to stick with a previously-successful response rule, despite occasional expectancy violations. This tendency to chase short-term reward, potentially at the expense of developing rules that maximize reward over the long term, may be a contributing factor to opioid addiction. Further work is indicated to better understand whether this tendency arises as a result of brain changes in the wake of continued opioid use/abuse, or might be a pre-existing factor that may contribute to risk for addiction.

1. Introduction

Addiction is a special case of impaired decision-making in which individuals continue to seek and use addictive substances despite negative consequences. In the extreme, these consequences

* Corresponding author at: Locked Bag 1797, Penrith, NSW 2751, Australia.
E-mail addresses: Catherine.Myers2@va.gov (C.E. Myers), a.moustafa@uws.edu.au (A.A. Moustafa).

http://dx.doi.org/10.1016/j.bbr.2015.09.018
0166-4328/

can include loss of income, family and friends, as well as illegal activity in acquiring and using illicit substances. Addiction can involve physical dependence, but also psychological dependence, as individuals continue to pursue and use drugs, even when they are well aware of these negative consequences and wish to stop the drug use. Thus, an important component of addiction is an abnormality in decision-making, and specifically how rewarding and punishing feedback are used to optimize behavior. One interpretation is that addicted individuals "chase" short-term reward, which may include both the acute drug effects as well as relief from withdrawal symptoms, at the expense of developing behavioral patterns that maximize reward over a longer time window.

Of particular societal concern are the highly-addictive opioid drugs including heroin, morphine, and a number of other medically-prescribed pain-killers such as oxycodone and hydrocodone. Due to an increasing use of opioids for pain management, accidental addiction leading to opioid abuse is currently a major issue. It is estimated that over 2 million people in the US alone have substance abuse disorders related to prescription opioid drugs, and abuse of these drugs may lead to abuse of heroin because it is cheaper and easier to obtain than prescription opioids [26]. Opioid addiction is notoriously difficult to overcome, even when the addicted individual strongly desires to stop using the drug; for example, one study reported 90% relapse rate for opioid-addicted individuals having undergone detoxification treatment [37]. An alternate and widely-preferred approach is maintenance therapy involving medically-supervised use of opioids such as methadone and buprenorphine; however, a recent review of outcomes following buprenorphine maintenance therapy reported that, in every study examined, 1 month following discontinuation of treatment rates of relapse to illicit opioid use exceeded 50% [4]. Given this bleak outlook, it is of great importance to better understand the mechanisms underlying impaired decision-making in opioid-addicted individuals, in order to be able to develop more effective therapies to promote and support these individuals in overcoming their dependence.

In previous work, we and others have used probabilistic reward-and-punishment learning tasks to better understand decision-making impairments in various psychiatric and neurological patient groups [5,7,13,16,25,28,29]. For example, a widely-used paradigm interleaves reward-learning trials and punishment-learning trials [5]. On the reward-learning trials, correct responses are often (but not always) rewarded with point gain, while incorrect responses trigger no feedback; on punishment-learning trials, incorrect responses are often (but not always) punished with point loss, while correct responses trigger no feedback. Thus, the task allows evaluation of the relative speed at which individuals learn to obtain reward vs. avoid punishment; it also allows investigation of how individuals deal with violation of expectancies (since the probabilistic nature of the task means that a response which is usually optimal may nevertheless be incorrect on any specific trial).

A number of prior studies have applied this task to patient populations. For example, Parkinson's disease (PD) involves progressive death of dopamine-producing neurons in the ventral striatum. Never-medicated PD patients perform similarly to matched controls on punishment-learning trials but are severely impaired at reward-learning trials, consistent with the idea that dopamine plays a key role in reward signaling; however, PD patients treated with dopaminergic drugs showed the reverse pattern: remediated reward learning but impaired punishment learning [5]. These results were recently replicated in a second study, which also considered a third group: PD patients who develop impulse control disorders (ICD) following treatment with dopaminergic medication [28]. Behaviorally, the PD–ICD group showed facilitated reward learning but impaired punishment learning. Other studies with this and similar tasks have documented facilitated reward-based learning in male veterans with severe post-traumatic stress disorder (PTSD) symptoms [25], facilitated reward and punishment learning in individuals with anxiety vulnerability [35], and correlation between reward-based (but not punishment-based) learning and negative (but not positive) symptoms in schizophrenia [14,38]. Thus, this type of task is able to dissociate qualitative patterns of behavior among different patient populations, presumably reflecting different nodes of dysfunction in the brain for the different disorders.

One way of understanding the mechanisms behind reward-and-punishment learning in these various groups is through reinforcement learning (RL) theories. RL theories suggest that actions are chosen to best maximize rewards (or reward value) in a given state [2,41,48]. RL models typically incorporate a concept of prediction error (PE), calculated by comparing actual outcomes (e.g. reward or punishment) against expected outcomes. One class of RL models, the actor-critic model [2,8], separates this prediction error from the action selection process. While one module ("critic") learns to calculate PE, and uses it to predict the value of the current environmental state, a second module ("actor") uses the PE to learn to select between competing possible actions. Specific parameters in the model can determine the rate at which an individual learns from rewarding or punishing feedback, the explore/exploit tradeoff (the degree to which an individual chooses previously-successful responses vs. occasionally trying new ones) and recency bias (the degree to which an individual simply repeats the most recent prior responses, regardless of past success).

Numerous studies suggest that the actor-critic model may provide a reasonable explanation of feedback-based learning in the brain. For example, in animals DA neurons respond especially to unexpected rewards [17,32], as well as to stimuli that signal upcoming or predicted reward [11] and reduce firing in response to omission of expected reward [17]. Similarly, in humans, healthy young males who underwent functional neuroimaging while performing the reward-and-punishment-learning task [5] showed activity in the dorsal caudate that was consistent with PE calculations [22]. Further, when several RL models were applied to data obtained from PD patients with and without ICDs, described above [28], the data were best described by an actor-critic model that allowed separate learning rates for reward and punishment trials in both the actor and the critic, a dissociation that appears to have plausible neural substrates [12,30]. Results from this model suggested that, while PD is associated with reduced reward-based learning in the actor, ICDs are associated with reduced learning in the critic, resulting in an underestimation of adverse consequences associated with stimuli that predict punishment.

Here, we ask whether the same evaluation of probabilistic reward-and-punishment learning, paired with RL modeling using an actor-critic model, can provide new insights into the mechanisms underlying the impairment in decision-making in opioid-addicted individuals, and could also suggest how this may be processed on a neural level. We applied the reward-and-punishment-learning task [5] and RL modeling to a group of individuals addicted to opioids (specifically, heroin-addicted patients on opioid maintenance treatment), and a group of individuals who had never abused illicit drugs. Observing evidence for different RL parameters in the opioid-addicted group, relative to non-drug-using controls, would promote understanding of the underlying mechanisms affecting decision-making in opioid-addicted individuals, potentially providing insight into how addiction is maintained and how it might be remediated.