Drug and Alcohol Dependence 188 (2018) 32–38

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

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

Full length article

Effects of reward and punishment on learning from errors in smokers

Leonie Duehlmeyer*, Bianca Levis, Robert Hester

School of Psychological Sciences, University of Melbourne, Redmond Barry Building, Parkville Campus, Melbourne, VIC 3010, Australia

ARTICLE INFO

Keywords:
Feedback learning
Error processing
Drug dependency
Smoking
Reward
Punishment

ABSTRACT

Background: Punishing errors facilitates adaptation in healthy individuals, while aberrant reward and punishment sensitivity in drug-dependent individuals may change this impact. Many societies have institutions that use the concept of punishing drug use behavior, making it important to understand how drug dependency mediates the effects of negative feedback for influencing adaptive behavior.
Methods: Using an associative learning task, we investigated differences in error correction rates of dependent smokers, compared with controls. Two versions of the task were administered to different participant samples: One assessed the effect of varying monetary contingencies to task performance, the other, the presence of reward as compared to avoidance of punishment for correct performance.
Results: While smokers recalled associations that were rewarded with a higher value 11% more often than lower rewarded locations, they did not correct higher punished locations more often. Controls exhibited the opposite pattern. The three-way interaction between magnitude, feedback type and group was significant, $F(1,48) = 5.288$, $p = 0.026$, $\eta^2 p = 0.099$. Neither participant group corrected locations offering reward more often than those offering avoidances of punishment. The interaction between group and feedback condition was not significant, $F(1,58) = 0.0$, $p = 0.99$, $\eta^2 p = 0.001$.
Conclusions: The present results suggest that smokers have poorer learning from errors when receiving negative feedback. Moreover, larger rewards reinforce smokers' behavior stronger than smaller rewards, whereas controls made no distinction. These findings support the hypothesis that dependent smokers may respond to positively framed and rewarded anti-smoking programs when compared to those relying on negative feedback or punishment.

1. Introduction

The capacity to learn new behaviors, rather than having them genetically imprinted, helped humans to inhabit most parts of the world. Next to observation and imitation, learning from trial and error is a crucial part of human learning (Beer, 1995). Research applying feedback learning paradigms that use trial and error learning have found punishment leading to better adaptation of future behavior compared to reward-only, with larger punishments resulting in greater future adaptation (Hester et al., 2010; Martin, 1963). Neglect of the future and dysfunctional reward circuits are common symptoms identified in a range of drug use disorders (Bechara et al., 2002; Beck et al., 2009; Bjork et al., 2008; Wrase et al., 2007). Drug-dependent individuals (DDIs) demonstrate hypersensitivity to monetary and drug-related reward and hyposensitivity to monetary and drug-related punishment (Bechara et al., 2002 Vanderschuren and Everitt, 2004; Volkow, 2004), which is argued to impede their ability to learn from negative feedback. Underlying neural mechanisms for this effect appear to precede drug consumption, potentially contributing to the propensity to consume drugs and develop dependence (Anokhin et al., 2010; van den Bree et al., 1998; Sweitzer et al., 2012; Whelan et al., 2014). Chronic drug consumption is thought to alter further this neural mechanism (Volkow, 2004), increasing sensitivity to reward-related stimuli and decreasing sensitivity to punishment-related stimuli.

In drug use disorder (DUD), neural mechanisms underlying learning from negative feedback have been shown to differ from those of healthy controls (Everitt and Robbins, 2005; Franken et al., 2010; Hester et al., 2013; Hyman et al., 2006). In healthy individuals, diverse negative emotional states, i.e., pain, anger and disgust (Naqvi and Bechara, 2009) and awareness of an outcome’s salience (Craig and Craig, 2009) have been found to activate the insula. These negative emotions could serve as feedback, promoting adjustment of behavior to avoid them. However, a range of DUD groups has exerted insula dysfunctions (Naqvi and Bechara, 2009; Paulus et al., 2005). For example, cocaine-dependent individuals had decreased insula activity during impaired adaptation of self-control after receiving monetary punishment (Hester

---

* Corresponding author at: 12th floor Redmond Barry Building, Parkville Campus, Melbourne VIC 3010, Australia.
E-mail addresses: lduhlmeyer@student.unimelb.edu.au (L. Duehlmeyer), b.levis@student.unimelb.edu.au (B. Levis), hesterr@unimelb.edu.au (R. Hester).

https://doi.org/10.1016/j.drugalcdep.2018.03.028
Received 5 September 2017; Received in revised form 6 March 2018; Accepted 6 March 2018

et al., 2013). Further brain regions with decreased activity during this task have previously been related to adjusting behavior in conflicting situations and future planning (Kerns et al., 2005; McClure et al., 2004). The DSM-V supports the importance of this topic by highlighting the failure to adapt behavior despite negative feedback as a criterion for drug dependence: Continued use despite recurring negative consequences caused/made worse by drug use. Dependent smokers present a crucial group of DDIs to investigate, with tobacco smoking remaining the primary cause of preventable mortality and morbidity around the world (Danaei et al., 2009; Ezzati and Lopez, 2003; Jha et al., 2008; Thorne et al., 2008). Previous research found smokers to show diminished adaptation of inhibitory control following negative feedback (Franken et al., 2010; Luijten et al., 2011), as well as hypersensitivity to monetary reward and hyposensitivity to monetary punishment (Luijten et al., 2013). Elucidating how positive and negative feedback influence learning from errors in smokers, may provide valuable insight into the influence these factors have on adaptive behavior. Moreover, how smokers might respond to positive and negative feedback on learning to adapt future behavior has relevance to designing effective intervention approaches.

Commonly used reward learning tasks (e.g., Iowa Gambling Task (Bechara et al., 1994), Balloon Analogue Risk Task (Lejuez et al., 2002)) require participants to risk punishment in seeking reward, often in a contingent way (larger rewards require larger risk). Learning from and avoiding future punishment in these tasks also requires self-control over risk-taking and seeking immediate reward. For example, in the Iowa Gambling Task, learning from punishment can be assumed from improved deck selection (Good > Bad) across the 100-trial task. However, failure to improve performance across the task might be due to a participant having difficulty encoding, learning from punishment (large monetary losses), or they can successfully encode the punishment but continue to choose the 'risky' decks because they are associated with larger monetary rewards. The Learning from Errors (LFE) task (Hester et al., 2007a) does not require participants to risk punishment in seeking reward, attempting to examine learning from punishment without the requirement for cognitive control over reward. Moreover, while this task requires cognitive control processes potentially confounded in addiction (e.g., spatial memory processing), it does not specifically confound one condition over another in the manner of some reward learning tasks that conflate risk and reward.

We administered two versions of the LFE task to investigate the impact of reward and punishment processing on learning from errors in dependent smokers. The first experiment examined the effect of divergent reward and punishment magnitudes (5¢ versus 50¢) on learning from errors (divergent value or divVLFE), while the second experiment increased the focus on punishment sensitivity by exchanging low (5¢) with no reward (0¢). More specifically, this version of the LFE task (Avoidance LFE or AvLFE) offered a reward of 50¢ for correctly recalled associations in contrast to only avoiding monetary punishment (0¢). Incorrect answers were punished (with a loss of 50¢) in both conditions.

We hypothesized that smokers would demonstrate a higher rate of retaining paired associates (of numbers and spatial locations) that had received larger (50¢) rewards during initial rewarding feedback, when compared to smaller (5¢) rewards. We also hypothesized that controls and dependent smokers would successfully correct their recall more often after receiving larger punishments (50¢) for initial recall errors, when compared to smaller punishments (5¢). We expected the punishment magnitude effect to be greater for controls. In the AvLFE task, we expected smokers to increasingly correct locations with a prospect for reward, whereas controls would make a lesser distinction between rewarded and non-rewarded locations.

2. Methods

2.1. DivVLFE task

2.1.1. Participants

25 dependent-cigarette smokers (10 females; mean age = 25.4 years; range = 19–36 years; years of education (YoE) = 14.4 years) and 25 controls (11 females; mean age = 24.8 years; range = 19–40 years; YoE = 14.7 participated in the experiment. Participants were recruited via advertisements at the University of Melbourne and a community website. All participants provided written informed consent, which was approved by Human Ethics Committee of The University of Melbourne and the Royal Children's Hospital. Participants were classified as dependent smokers when they smoked at least fifteen cigarettes daily, while controls had smoked less than 6 cigarettes in their lifetime. Exclusion criteria for both groups consisted of a history of neurological or psychiatric disorders, current use of psychotropic medication, and current drug abuse or dependence (other than nicotine for the smoking group). Groups did not significantly differ on variables of age ($t(48) = -0.417$, $p = 0.678$, $d = 0.112$) and YoE ($t(48) = 0.514$, $p = 0.610$, $d = 0.145$). Fagerstrom Test for Nicotine Dependence (FTND) resulted in an average score of 4.5 for smokers, describing a moderate dependence (Heatherton et al., 1991). Self-reported alcohol use was significantly higher in smokers as measured with the Alcohol Use Disorder Identification Test (AUDIT) (Saunders et al., 1993) (controls = 2.1, smokers = 9.0; $t(48) = -6.632$, $p < 0.001$, $d = 1.876$). AUDIT scores, smoker's breath carbon monoxide (CO), and craving as measured with the QSU-brief (Cox et al., 2001) were correlated with the dependent variables of interest.

2.1.2. Experimental protocols

A spatial paired-associates learning task was administered to participants (Fig. 1). All aspects of stimulus delivery and response recording were controlled by E-Prime software (version 2.0, Psychology Software Tools, Inc. Pittsburgh, PA), running on a Windows-compatible PC. The task began with an encoding phase in which eight gray squares were presented simultaneously on a black background. The locations of the squares were selected in a quasi-random fashion from an 8 × 8 matrix, with two locations randomly chosen from each of the four quadrants of the display.

First, each location, in turn, had superimposed upon it a two-digit number (1.5s), followed by an interstimulus interval (ISI) (1s). Each number's digits consisted of 1–3, or 4. Participants identified the number using a pair of response boxes (Current Designs), two-digit numbers were used to reduce the probability of guessing the correct answer to 6%. Following the encoding phase, a series of recall trials were presented. One of the eight locations were highlighted in yellow, cueing participants to respond with the associated number-location pair. Participants were required to respond within 3s, after which a variable ISI was presented (2–4s). During the ISI, the location remained highlighted by a yellow border. Feedback (2s) was then provided for the validity of the response and magnitude of reward/punishment. The location square turned blue to indicate a correct response or red to indicate an incorrect response. A photo of an Australian 5¢ or 50¢ coin was superimposed over the colored background. Feedback magnitude was randomly assigned to each location but modeled to ensure equal amounts of 5¢ or 50¢ feedback magnitudes for correct trials and error trials (separately). Once assigned, feedback magnitude of a location was fixed for round 2 recall trials, ensuring that round 1 feedback predicted future reward and punishment value of a location. Each block's gains and losses were added to an initial credit of AU$10. Following the feedback epoch, a second ISI was presented (2–4s), during which the target square remained colored (blue or red, depending on accuracy. Then, the correct two-digit number was presented on the colored location, allowing participants to re-encode the correct answer. When