Addictive Behaviors 88 (2019) 73–76

Contents lists available at ScienceDirect

Addictive Behaviors

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

Short Communication

Individual differences in learning from probabilistic reward and punishment predicts smoking status

Laura A. Rai[a], Laura O'Halloran[a], Lee Jollans[a], Nigel Vahey[a], Conall O'Brolchain[b], Robert Whelan[a,c,*]

[a] School of Psychology, Trinity College Dublin, Ireland
[b] School of Psychology, University College Dublin, Ireland
[c] Global Brain Health Institute, Trinity College Dublin, Ireland

HIGHLIGHTS

- We compared reward and punishment learning in current smokers, ex-smokers, and non-smokers using the Probabilistic Selection Task
- Probabilistic Selection Task performance predicted smoking status with moderate accuracy
- Smokers and ex-smokers showed decreased learning from reward feedback compared with non-smokers, whereas smokers showed increased learning from punishment feedback

ARTICLE INFO

Keywords:
Nicotine
Reward
Punishment
Learning
Ex-smokers

ABSTRACT

Introduction: The ability to update reward and punishment contingencies is a fundamental aspect of effective decision-making, requiring the ability to successfully adapt to the changing demands of one's environment. In the case of nicotine addiction, research has predominantly focused on reward- and punishment-based learning processes among current smokers relative to non-smokers, whereas less is known about these processes in former smokers.

Methods: In a total sample of 105 students, we used the Probabilistic Selection Task to examine differences in reinforcement learning among 41 current smokers, 29 ex-smokers, and 35 non-smokers. The PST was comprised of a training and test phase that allowed for the comparison of learning from positive versus negative feedback.

Results: The test phase of the Probabilistic Selection Task significantly predicted smoking status. Current and non-smokers were classified with moderate accuracy, whereas ex-smokers were typically misclassified as smokers. Lower rates of learning from rewards were associated with an increased likelihood of being a smoker or an ex-smoker compared with being a non-smoker. Higher rates of learning from punishment were associated with an increased likelihood of being a smoker relative to non-smoker. However, learning from punishment did not predict ex-smoker status.

Conclusions: Current smokers and ex-smokers were less likely to learn from rewards, supporting the hypothesis that deficient reward processing is a feature of chronic addiction. In addition, current smokers were more sensitive to punishment than ex-smokers, contradicting some recent findings.

1. Introduction

Addiction can be framed as a maladaptive decision-making process, in which substances are persistently sought out by the individual despite negative repercussions. The ability to update reward and punishment contingencies is a fundamental aspect of effective decision-making, requiring the ability to successfully adapt to a changing environment. Broadly, nicotine dependence is associated with increased behavioural impulsivity and higher discounting of delayed rewards. Nicotine also modulates reward-based learning in both human and rodent studies. For example, in non-smokers, a single dose of nicotine increased responsiveness to reward cues, lasting for up to one-week following administration (Barr, Pizzagalli, Culhane, Goff, & Evins, 2008), suggesting that early reinforcement of non-drug cues in the

* Corresponding author at: Lloyd Building, Trinity College Dublin, Ireland.
E-mail address: Robert.Whelan@tcd.ie (R. Whelan).

https://doi.org/10.1016/j.addbeh.2018.08.019
Received 20 March 2018; Received in revised form 10 August 2018; Accepted 14 August 2018
Available online 19 August 2018
0306-4603/

environment as a result of smoking may lead to nicotine dependence.

The probabilistic selection task (PST; Frank, Seeberger, & O'Reilly, 2004) quantifies individual differences in learning from reinforcement relative to learning from punishment (i.e., from positive relative to negative feedback). Parkinson's Disease patients on dopaminergic agonist medication (i.e., with sufficient dopamine) learned more effectively from reinforcers than from punishers, with a reverse pattern observed in off-medication patients (Frank et al., 2004). Individuals with a range of substance misuse (e.g., alcohol, cannabis, and nicotine use) were poorer at both learning from rewards and from punishers compared to non-dependent groups (Baker, Stockwell, Barnes, & Holroyd, 2011; Baker, Stockwell, & Holroyd, 2013). The findings of Baker et al. (2011, 2013) support addiction models that include desensitization of reward circuits over time (e.g., Volkow, Koob, & McLellan, 2016). While acute nicotine administration amplifies reward learning with respect to non-drug cues, chronic nicotine use may desensitize the dopamine system and consequently blunt reward sensitivity. Fehr et al. (2008) demonstrated reduced availability of striatal D_2/D_3 dopamine receptors in nicotine dependence. However, as Nestor, McCabe, Jones, Clancy, and Garavan (2018a) note, this is in contrast to the striatal hyperactivity to non-drug rewards observed in some addiction populations.

Garavan, Brennan, Hester, and Whelan (2013) proposed that successful abstinence is characterised both by the restoration of brain function once the neurotoxic effects of the drug abuse diminish, and also by the continued process of abstaining from the drug. Briggs et al. (2015) found that former- and non-smokers, versus current smokers, more effectively updated shifting reward and punishment contingencies in the Iowa Gambling Task. In a reversal task, Butler, Rusted, Gard, and Jackson (2017) showed that current smokers had more reversal errors than either ex- or non-smokers. In contrast to those studies suggesting that ex-smokers display similar decision-making processes to non-smokers rather than smokers, Nestor, McCabe, Jones, Clancy, and Garavan (2018b) reported that ex-smokers demonstrate amplified negative valence monitoring compared with smokers and non-smokers.

The current study examined reward and punishment learning in ex-smokers, current and non-smokers using the PST. Based on previous studies in substance-dependent samples (Baker et al., 2011, 2013; Nestor et al., 2018b), we hypothesised that the current smokers would show impaired reward learning compared with ex-smokers and non-smokers. We also expected that ex-smokers, relative to current and never smokers, would learn best from punishers.

2. Methods

2.1. Participants

57 current smokers had smoked over 40 lifetime cigarettes, with at least weekly smoking in the past 30 days. 40 ex-smokers smoked more than 40 cigarettes in their lifetime, with fewer than one cigarette per week (4 participants), or no cigarettes at all, in the past 30 days. 43 non-smokers smoked on fewer than 40 occasions in their entire lifetime with no cigarettes at all in the past 30 days. Exhaled carbon monoxide readings were collected from a subset of 60 participants (25 Smokers; 17 ex-smokers; 13 non-smokers). Smokers had readings of ≥ 6 ppm, and ex-smokers and non-smokers ≤ 5 ppm (Low, Ong, & Tan, 2004).

2.2. Procedure

Ethics Committees from University College Dublin and Trinity College Dublin approved the study. Current smokers were requested to smoke as normal prior to the experiment, and therefore were not in acute abstinence. Participants completed the experimental tasks alone in a sound-attenuated booth. Questionnaires were completed during the testing session, or at home via an online survey platform. The PST was part of a larger test battery that took approximately 1 h. Participants were compensated with €10 (approximately $12) plus maximum travel expenses of €10.

2.3. Measures

2.3.1. Probabilistic selection task

The PST comprised of a training and test phase. During training, three stimulus pairs (AB, CD, EF) were randomly presented. Stimulus position was random across trials. Stimulus reward probabilities were predetermined (A:80%, B:20%, C:70%, D:30%, E:60%, F:40%). Correct and incorrect were signalled by a green tick or red X, respectively. In the Test phase, novel stimulus combinations were presented without feedback. There was no intervening period between the presentation of the training and test phases of the PST. Test phase performance was quantified by selection frequency of the A stimulus versus the B stimulus in novel pairs. A should be preferable to all other stimuli following positive feedback learning, whereas B should always be avoided following negative feedback learning. Consistent with previous PST papers (Aberg, Doell, & Schwartz, 2016), only participants with over 60% accuracy for AB pairs, and over 50% accuracy for CD during training were included in the final analysis.

Statistical analyses were performed in IBM SPSS (Version 23). Non-parametric tests were used when appropriate. Alpha was 0.05 unless stated otherwise due to multiple comparison correction. Our goal was to predict group membership (see Yarkoni and Westfall, 2017, for a rationale for prediction versus explanation) and therefore percent Approach A and Avoid B selections were predictor variables in a multinomial logistic regression model. The non-smoker group was the reference category and p values were calculated using 1000 bootstrapped samples.

2.3.2. Questionnaire measures

The ESPAD questionnaire on substance use (Hibell & Bjarnason, 2008) was used to assess lifetime and past 30 days smoking, and past 30 days alcohol use. The Fagerstrom Test for Nicotine Dependence (FTND; Heatherton, Kozlowski, Frecker, & Fagerstrom, 1991) was used to measure nicotine dependence in the smoker group. The Barratt Impulsiveness Scale (BIS-11; Patton, Stanford, & Barratt, 1995) is a 30-item measure of impulsiveness. The Sensation Seeking Scale (*SSS*; Zuckerman, 1971) is a 40-item measure with four sensation seeking subscales: thrill- and adventure-seeking, disinhibition, experience-seeking and susceptibility to boredom (further details in Supplemental Materials).

3. Results

The final sample consisted of 41 smokers, 29 ex-smokers and 35 non-smokers. Participant characteristics are presented in Table 1 (see Supplemental Materials for further information on the PST training phase). Specific age was collected for 58 participants (55.2%) of the final sample; remaining participants were aged 18–21 years. There was a significant difference between groups based on the 58 participants who reported their exact age (Kruskal Wallis test; $\chi^2(2, 58) = 8.069$, $p = .018$). The ex-smoker group ($N = 17$) were older than the smoker ($N = 25$) and non-smoker ($N = 16$) groups. The mean FTND for smokers was 2.17 (SD = 2.26), indicating the 'Low' dependence that is typical for younger smokers (Li et al., 2015).

For Approach A trials, non-smokers chose A more often than smokers and non-smokers, with a median % choice (interquartile range) of 89(29), 83(32), and 77(38), respectively. For Avoid B trials, non-smokers performed similarly to smokers, with a median % choice of 75(26) and 76(21) respectively, while ex-smokers avoided the B stimulus on 72% of trials (IQR = 20). Approach A and Avoid B percentages were entered in a multinomial logistic regression. Performance on the learning from reward (i.e., Approach A) test trials successfully predicted smoker group (Approach A, $\chi^2 = 7.01$, df(2) $p = .030$), while learning

Table 1
ESPAD, CO reading, PST performance, reaction times, and personality scores (Barratt Impulsiveness Scale and Sensation-Seeking Scale) by group.

Measure	Smokers	Ex-smokers	Non-smokers	Significant difference
Gender (M/F)	26/15	18/11	13/22	–
Age (Years)[a]	21 (5)	32 (21)	22.5 (3.5)	Ex > S & NS
Lifetime smoking (ESPAD)[b]	7 (0)	7 (0)	2 (3)	
Past 30 days smoking (ESPAD)[c]	4 (1)	1 (0)	1 (0)	
Past 30 days Alcohol (ESPAD)[a]	4 (3)	4 (3)	4 (2)	–
Carbon Monoxide (ppm)[a]	11 (5.5)	2 (1)	2 (1.5)	S > Ex & NS
FTND Total[a]	2.17 (2.26)	–	–	N/A
BIS Total[a]	70 (16)	62 (11)	66.5 (10)	S > Ex
BIS Attentional[a]	19 (5.5)	16 (4.5)	17.5 (4)	–
BIS Motor[a]	24 (6)	22 (4)	23 (5.5)	–
BIS Non-planning[a]	26 (7)	26 (9)	25 (5.75)	–
SSS Total[a]	26 (11)	20 (11)	21.5 (5)	S > Ex & NS
SSS Boredom Susceptibility[a]	3 (3)	3 (2.5)	2 (1.75)	–
SSS Disinhibition[a]	7 (3)	6 (4)	6 (3)	–
SSS Experience Seeking[a]	7 (3)	6 (3)	6 (3)	S > NS
SSS Thrill & Adventure[a] Seeking	8 (5)	6 (7)	8 (3.75)	–

[a] Median(Inter Quartile Range).
[b] The ESPAD Lifetime Smoking variable was coded as follows: 1 = 0 cigarettes, 2 = 1–2 cigarettes, 3 = 3–5 cigarettes, 4 = 6–9 cigarettes, 5 = 10–19 cigarettes, 6 = 20–39 cigarettes, and 7 = 40+ cigarettes.
[c] The ESPAD Past Month Smoking variable was coded as follows: 1 = Not at all, 2 = Less than 1 cigarette per week, 3 = Less than 1 cigarette per day, 4 = 1–5 cigarettes per day, 5 = 6–10 cigarettes per day, 6 = 11–20 cigarettes per day, or 7 = 20+ cigarettes per day. S = current smoker; Ex = ex-smoker; NS = non-smoker.

Table 2
Classification table for multinomial regression with PST Approach A and Avoid B as predictor variables.

Observed	Predicted			
	Smoker	Ex-smoker	Non-smoker	Correct (%)
Smoker	29	3	9	70.7
Ex-Smoker	17	5	7	17.2
Non-Smoker	13	4	18	51.4
Overall Correct (%)	56.2	11.4	32.4	49.5

from punishment (i.e., Avoid B) test trials was just greater than the significance threshold (Avoid B, $\chi^2 = 5.96$, df(2), $p = .051$). Table 2 displays the classification accuracy of the multinomial logistic regression. As the tendency to learn from positive outcomes increased, the likelihood of being a non-smoker relative to smoker ($p = .024$; 95% Confidence Interval − 0.089 to −0.009) or ex-smoker ($p = .04$; 95% CI -0.075 to −0.001) increased. In contrast, as the tendency to learn from punishment increased, the likelihood of being a smoker relative to non-smoker increased ($p = .034$; 95% CI 0.006 to 0.095), but this was not significant for ex-smokers compared with non-smokers ($p > .05$). A separate multinomial regression was conducted to control for the inclusion of lighter smokers in our analysis, and produced similar findings (see Supplementary Materials). CO readings significantly correlated with the tendency to learn from punishment (rho = 0.31, $p = .020$), but not from reward (rho = 0.01, $p = .93$).

BIS-11 scores were compared using Kruskal Wallis tests. Groups differed in total BIS score ($\chi^2(2, 105) = 7.03$, $p = .03$) and total SSS scores ($\chi^2(2, 98) = 7.5$, $p = .02$). Questionnaire data for eight participants were missing (final sample size: 41 Smokers, 29 Ex-Smokers, 28 Non-smokers). Total BIS and SSS scores were entered as predictor variables in a separate multinomial regression model. Neither questionnaire significantly predicted smoking group status ($p > .05$). However, total SSS score predicted the likelihood of belonging to the smoker group relative to the non-smoker group ($p = .046$, 95% CI -0.008 to 0.176). Correlations between the PST (Approach A and Avoid B), and personality measures (the BIS and SSS) were not significant.

4. Discussion

Individual differences in reward learning predicted smoker status with moderate accuracy. Relative to non-smokers, smokers and ex-smokers had decreased learning from reward. Our results are concordant with those of Baker et al. (2011), in that our non-dependent (i.e., non-smoker) group showed higher reward learning in the PST compared with the dependent (i.e., smoker) group. In contrast, Potts, Bloom, Evans, and Drobes (2014), in a flanker task, reported that ex-smokers and smokers were more sensitive to rewards compared with non-smokers. Unlike Baker et al. (2011), our dependent group demonstrated increased learning from punishment relative to the non-dependent group. Butler et al. (2017), observed poor performance monitoring in smokers and found that post-punishment slowing correctly identified current smokers more so than former smokers (80% vs 60%).

Some of the contrast between the current findings and previous research may be attributable to phenotypic and methodological differences among studies. For example, Nestor et al. (2018b) included ex-smokers, abstinent for at least 12 months prior to testing. Potts et al. (2014) included only smokers who smoked over 10 cigarettes per day for the past year. The current study included smokers who were on average low in nicotine dependency, and abstinence was operationalized by the participant's self-reported smoking behaviour in the past 30 days. Carballo and López (2014) found increased length of abstinence in cocaine-dependent participants improved performance in response to negative feedback on a flanker task. Prolonged nicotine abstinence may similarly affect punishment sensitivity in the PST. Nestor et al. (2018b) used the Monetary Incentive Delay task, which focuses on gain and loss anticipation, while Potts et al. (2014) used a modified flanker task without feedback.

Many researchers (e.g., Baker, Piper, McCarthy, Majeskie, & Fiore, 2004; Blum et al., 2000; Koob, 2009) posit a negative affect addiction stage, involving avoidance of negative emotional after-effects of drug use. Lower levels of dopamine D2 receptor availability have been observed in chronic addiction. Lower dopamine levels have also been associated with increased learning from punishment (Frank et al., 2004). Martin, Cox, Brooks, and Savage (2014) showed that smokers were hyper-responsive to the anticipation of punishment. It is conceivable that our current smokers showed increased sensitivity to punishment due to decreased dopaminergic activity, and indeed smoking heaviness (indexed by CO level) correlated with learning from punishment. This may also explain why punishment learning did not predict ex-smoker group status, as this group was no longer experiencing the negative affect stage of their former addiction.

In conclusion, these findings provide an insight into the effects of smoking status on reward and punishment learning using the PST. The results suggest that the PST has some utility in discriminating between smokers, ex-smokers, and non-smokers. These behavioural findings may be useful in understanding which smoking-cessation techniques are most effective, based on their use of positive and negative reinforcement.

Conflict of interest

The authors report no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.addbeh.2018.08.019.

References

Aberg, K. C., Doell, K. C., & Schwartz, S. (2016). Linking Individual Learning Styles to Approach-Avoidance Motivational Traits and Computational Aspects of Reinforcement Learning. *PLoS One, 11*(11), e0166675. https://doi.org/10.1371/journal.pone.0166675.

Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R., & Fiore, M. C. (2004). Adiction motivation reformulated: An affective processing model of negtive reinforcement. *Psychological Review, 111*(1), 33–51. https://doi.org/10.1037/0033-295X.111.1.33.

Baker, T. E., Stockwell, T., Barnes, G., & Holroyd, C. B. (2011). Individual differences in substance dependence: At the intersection of brain, behaviour and cognition: Substance dependence. *Addiction Biology, 16*(3), 458–466. https://doi.org/10.1111/j.1369-1600.2010.00243.x.

Baker, T. E., Stockwell, T., & Holroyd, C. B. (2013). Constraints on decision making: Implications from genetics, personality, and addiction. *Cognitive, Affective, & Behavioral Neuroscience, 13*(3), 417–436. https://doi.org/10.3758/s13415-013-0164-8.

Barr, R. S., Pizzagalli, D. A., Culhane, M. A., Goff, D. C., & Evins, A. E. (2008). A single dose of nicotine enhances reward responsiveness in nonsmokers: implications for development of dependence. *Biological Psychiatry, 63*(11), 1061–1065. https://doi.org/10.1016/j.biopsych.2007.09.015.

Blum, K., Braverman, E. R., Holder, J. M., Lubar, J. F., Monastra, V. J., Miller, D., ... Comings, D. E. (2000). Reward deficiency syndrome: A biogenetic model for the diagnosis and treatment of impulsive, addictive, and compulsive behaviors. *Journal of Psychoactive Drugs, 32*, 1–112.

Briggs, Z., O'Connor, M., Jollans, E. K., O'Halloran, L., Dymond, S., & Whelan, R. (2015). Flexible emotion-based decision-making behavior varies in current and former smokers. *Addictive Behaviors, 45*, 269–275. https://doi.org/10.1016/j.addbeh.2015.02.011.

Butler, K., Rusted, J., Gard, P., & Jackson, A. (2017). Performance monitoring in nicotine dependence: Considering integration of recent reinforcement history. *Pharmacology, Biochemistry, and Behavior, 156*, 63–70. https://doi.org/10.1016/j.pbb.2017.04.004.

Carballo, M., & López, V. (2014). Error monitoring recovery after abstinence and treatment in cocaine addicts / Recuperación del procesamiento de errores tras abstinencia y tratamiento en adictos a cocaína. *Estudios de Psicología, 35*(1), 193–199. https://doi.org/10.1080/02109395.2014.893656.

Fehr, C., Yakushev, I., Hohmann, N., Buchholz, H.-G., Landvogt, C., Deckers, H., ... Schreckenberger, M. (2008). Association of low striatal dopamine d2 receptor availability with nicotine dependence similar to that seen with other drugs of abuse. *The American Journal of Psychiatry, 165*(4), 507–514. https://doi.org/10.1176/appi.ajp.2007.07020352.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.), 306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941.

Garavan, H., Brennan, K. L., Hester, R., & Whelan, R. (2013). The Neurobiology of successful Abstinence. *Current Opinion in Neurobiology, 23*(4), 668–674. https://doi.org/10.1016/j.conb.2013.01.029.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The Fagerstrom Test for Nicotine Dependence: A revision of the Fagerstrom Tolerance Questionnaire. *British Journal of Addiction, 86*, 1119–1127.

Hibell, B., & Bjarnason, T. (2008). *Report from the ESPAD 07 questionnaire test (manuscript)*. Stockholm: Swedish Council for Information on Alcohol and Other Drugs.

Koob, G. F. (2009). Neurobiological substrates for the dark side of compulsivity in addiction. *Neuropharmacology, 56*(Suppl. 1), 18–31. https://doi.org/10.1016/j.neuropharm.2008.07.043.

Li, H., Zhou, Y., Li, S., Wang, Q., Pan, L., Yang, X., ... Jia, C. (2015). The relationship between nicotine dependence and age among current smokers. *Iranian Journal of Public Health, 44*(4), 495–500.

Low, E. C. T., Ong, M. C. C., & Tan, M. (2004). Breath carbon monoxide as an indication of smoking habit in the military setting. *Singapore Medical Journal, 45*, 578–582.

Martin, L. E., Cox, L. S., Brooks, W. M., & Savage, C. R. (2014). Winning and losing: Differences in reward and punishment sensitivity between smokers and nonsmokers. *Brain and Behavior: A Cognitive Neuroscience Perspective, 4*(6), 915–924. https://doi.org/10.1002/brb3.285.

Nestor, L. J., McCabe, E., Jones, J., Clancy, L., & Garavan, H. (2018a). Shared and divergent neural reactivity to non-drug operant response outcomes in current smokers and ex-smokers. *Brain Research, 1680*, 54–61. https://doi.org/10.1016/j.brainres.2017.12.003.

Nestor, L. J., McCabe, E., Jones, J., Clancy, L., & Garavan, H. (2018b). Smokers and ex-smokers have shared differences in the neural substrates for potential monetary gains and losses: Reward processing in smokers. *Addiction Biology, 23*(1), 369–378. https://doi.org/10.1111/adb.12484.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology, 51*, 768–774.

Potts, G. F., Bloom, E., Evans, D. E., & Drobes, D. J. (2014). Neural reward and punishment sensitivity in cigarette smokers. *Drug and Alcohol Dependence, 144*, 245–253. https://doi.org/10.1016/j.drugalcdep.2014.09.773.

Volkow, N. D., Koob, G. F., & McLellan, A. T. (2016). Neurobiologic advances from the brain disease model of addiction. *New England Journal of Medicine, 374*(4), 363–371. https://doi.org/10.1056/NEJMra1511480.

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393.

Zuckerman, M. (1971). Dimensions of sensation seeking. *Journal of Consulting and Clinical Psychology, 36*, 45–52.