Cite as: Hogarth, L. (2012). Goal-directed and transfer-cue-elicited drug-seeking are dissociated by pharmacotherapy: Evidence for independent additive controllers. Journal of Experimental Psychology: Animal Behavior Processes, 38, 266-278.

Goal-directed and transfer-cue-elicited drug-seeking are dissociated by drug preload: Evidence for independent additive controllers

Lee Hogarth

School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia.

Now at: School of Psychology, University of Exeter, Washington Singer Building, Perry Road, Exeter EX4 4QG, UK.

Acknowledgements

The research reported in this manuscript was supported by a UK MRC Grant (G0701456) held by Lee Hogarth whilst at the University of Nottingham UK. The Author would like to thank Amy Inman, Emily Hassall & Kristy Martin for data collection.

Address for Correspondence:

Dr. Lee Hogarth
A/Prof
School of Psychology,
University of Exeter,
Washington Singer Building,
Perry Road, Exeter
EX4 4QG, UK.

Abstract

According to contemporary learning theory, drug-seeking behaviour reflects the summation of two dissociable controllers. Whereas goal-directed drug-seeking is determined by the expected current incentive value of the drug, stimulus-elicited drug-seeking is determined by the expected probability of the drug independently of its current incentive value, and these two controllers contribute additively to observed drug-seeking. One applied prediction of this model is that smoking cessation pharmacotherapies selectively attenuate tonic but not cue-elicited craving because they downgrade the expected incentive value of the drug but leave expected probability intact. To test this, the current study examined whether nicotine replacement therapy (NRT) nasal spray would modify goal-directed tobacco choice in a human outcome devaluation procedure, but leave cue-elicited tobacco choice in a Pavlovian to instrumental transfer (PIT) procedure intact. Smokers (n=96) first underwent concurrent choice training in which two responses earned tobacco or chocolate points respectively. Participants then ingested either NRT nasal spray (1mg) or chocolate (147g) to devalue one outcome. Concurrent choice was then tested again in extinction to measure goal-directed control of choice, and in a PIT test to measure the extent to which tobacco and chocolate stimuli enhanced choice of the same outcome. It was found that NRT modified tobacco choice in the extinction test but not the extent to which the tobacco stimulus enhanced choice of the tobacco outcome in the PIT test. This dissociation suggests that the propensity to engage in drug-seeking is determined independently by the expected value and probability of the drug, and that the partial efficacy of pharmacotherapy is due to its selective effect on expected drug value.

Key words: Addiction; outcome revaluation; goal-directed action; Pavlovian transfer

Introduction

Animal learning theorists have recently articulated a dual-controller theory to describe the unique psychological processes mediating intentional versus cue-provoked action selection (Balleine & O'Doherty, 2010; Belin, Jonkman, Dickinson, Robbins, & Everitt, 2009; de Wit & Dickinson, 2009; Hogarth & Chase, 2011b; Ostlund & Balleine, 2008). The core empirical source for this dual-controller account is the dissociation between the outcome devaluation and Pavlovian to instrumental transfer (PIT) procedures (Colwill & Rescorla, 1990; Corbit, Janak, & Balleine, 2007; Holland, 2004; Rescorla, 1994). In these designs, Pavlovian training is used to establish two distinct conditioned stimuli as differential predictors of two outcomes (e.g. pellets and sucrose). Instrumental training separately establishes distinct instrumental responses (e.g. lever press and chain pull) as differentially causal in producing these same two outcomes. One of the outcomes is then devalued by specific satiety or pairing it with lithium chloride induced sickness, before the performance of the two instrumental responses is tested in extinction with no stimuli, and in a Pavlovian to instrumental transfer (PIT) test where the two stimuli are presented.

These designs have found that selection between the two actions in the extinction test is sensitive to outcome devaluation, that is, there is a selective reduction in choice of the action that produced the devalued outcome. As this choice is made in extinction (i.e. no outcomes are presented), choice must be mediated by retrieval of a representation of the current incentive value of the outcomes earned by the two responses which determines the propensity to perform each response, that is, choice is goal-directed or intentional. By contrast, in the PIT test, each stimulus selectively enhances choice of the response that was associated with the same outcome, indicating that each stimulus must have retrieved a

representation of its associated outcome, which in turn elicited the response that was associated with that outcome.

The extraordinary and paradoxical finding, however, is that the magnitude of this outcome-specific transfer effect is not modulated by the devaluation treatment, indicating that each stimulus must have retrieved a representation of the perceptual identity of its paired outcome, which primed the associated response, without retrieving the current incentive value of the outcome (Colwill & Rescorla, 1990; Corbit, et al., 2007; Holland, 2004; Rescorla, 1994). By contrast, the transfer effect can be modified by manipulating the strength of the cue-outcome contingency in training, suggesting that stimulus-evoked expectations regarding the probability of the outcome drive choice of that outcome in the PIT test (Balleine, Leung, & Ostlund, 2011; Delamater, 1995; Gámez & Rosas, 2005; Trick, Hogarth, & Duka, 2011). The important and unique point made dual-controller theory is that this probability estimate evoked by the stimulus is not weighted by the current value of the outcome, given that the transfer effect is not sensitive to devaluation. Thus, the propensity to select an action appears to reflect the summation of two independent controllers: Whereas the overall propensity to engage in a goal-directed choice is determined by the expected incentive value of the outcome, the capacity of cues to elicit this choice is determined by the expected probability of the outcome. These value and probability estimates contribute independently and additively to choice.

A number of addiction theories have proposed that drug-seeking transitions between an intentional and more automatic cue-locked form, broadly consistent with the dual-controller framework (Robinson & Berridge, 1993; Tiffany, 1990). Research on nicotine dependence has amassed a substantial body of evidence in support of this position. The typical behavioural model employed in this research involves recording smokers' level of cigarette

craving in the absence of any smoking cues to measure tonic or background craving. In addition, craving is recorded in the presence of smoking and neutral cues, and the difference reflects the degree of cue-elicited craving. A remarkably consistent finding across these studies is that manipulations of smoking satiety versus deprivation modulate the magnitude of tonic craving measured both in the absence of cues and across the smoking and neutral cue conditions collectively. By contrast, these manipulations have no effect on the cue-elicited craving, that is, no effect on the increase in craving produced by the smoking cue over the neutral cue, or relative to a pre-cue baseline (Drobes & Tiffany, 1997; Maude-Griffin & Tiffany, 1996). More evidence for this dissociation comes from a study by Hogarth, Dickinson, & Duka (2010). In this design, smokers first learned that two arbitrary stimuli, S+ and S-, predicted that a button press response would earn and lose ¼ of a cigarette, respectively. These stimuli were then tested (in a transfer test) for their capacity to augment craving and the number of cigarette puffs consumed on a trial by trial basis in an ad libitum smoking session. The results showed that craving and puff number declined across the two blocks of this ad libitum smoking session, reflecting burgeoning satiety, but the capacity of the S+ to enhance craving and puff number above the S- was statistically equivalent across blocks. Thus, whereas the overall propensity for craving and puffing were determined by the incentive value of smoking, cue-elicited craving and puffing were autonomous of incentive value and were instead dependent on participants explicit knowledge of the cue-outcome contingencies in training. These dissociations accord with the additive dual-controller model in suggesting that value and probability make independent additive contributions to craving and drug-seeking.

Direct evidence for the dual-controller theory of addiction, however, has only recently been demonstrated by the dissociable effect of outcome devaluation on goal-directed versus transfer-cue-elicited drug-seeking (Hogarth & Chase, 2011a). In this study, smokers first

learned distinct instrumental responses for tobacco and chocolate points before one outcome was devalued by health warnings against consumption of that outcome. Choice was then tested in extinction, i.e. without feedback from the outcomes, and as expected, participants reduced choice of the devalued outcome indicating this behaviour was goal-directed in being mediated by the expected value of the outcome. Moreover, a cigarette and chocolate picture presented at the choice point in a nominal PIT test enhanced choice of the same outcome (also replicated by Hogarth & Chase, 2012). That is, the cigarette picture enhanced choice of tobacco points and the chocolate picture enhanced choice of chocolate points. The implication is that the pictures, through a process of generalisation (McLaren & Mackintosh, 2002) and inference (Mitchell, De Houwer, & Lovibond, 2009), increased the expected probability of the outcome with which it shared category membership (Wills & Pothos, 2012), which in turn, enhanced choice of that outcome. This nominal PIT test is comparable to standard PIT procedures insofar as the pictures should have been associated with their consummatory rewards in participants' extra-experimental history, but have never been reinforced contiguously with the instrumental responses, precluding interpretation of the cue effect on choice by the formation of a direction association between the stimulus and the response, which was the original objective of the PIT design (Balleine & Ostlund, 2007).

The crucial finding from the PIT test, however, was that the capacity of cues to enhance choice of the same outcome was not modulated by health warnings. Specifically, although health warnings against smoking decreased the overall propensity to choose the tobacco response in the PIT test, the capacity of the tobacco picture to enhance tobacco choice above a no-stimulus baseline was entirely unaffected by health warnings. Thus, whereas the overall propensity to choose the tobacco response was goal-directed in being determined by the expected value of the tobacco outcome, cue-elicited tobacco choice was apparently determined by the expected probability of the tobacco outcome, which itself was not

weighted by expected value. Thus, expected drug value and probability appear to converge separately on the propensity to engage in drug-seeking.

The dual-controller model may explain the partial therapeutic efficacy of addiction pharmacotherapy. As with manipulations of smoking deprivation and satiety described above, smoking cessation medications have been shown to attenuate tonic cigarette craving but not cue-elicited craving (for a review see Ferguson & Shiffman, 2009). This dissociation has been found with nicotine replacement therapy (Havermans, Debaere, Smulders, Wiers, & Jansen, 2003; Morissette, Palfai, Gulliver, Spiegel, & Barlow, 2005; Niaura et al., 2005; Rohsenow et al., 2007; Shiffman et al., 2003; Tiffany, Cox, & Elash, 2000; Waters et al., 2004), bupropion (Hussain et al., 2010) and varenicline (Brandon et al., 2011; Franklin et al., 2011; Hitsman et al., under review; Hitsman, Niaura, Shadel, Britt, & Price, 2006; for related animal data see O'Connor, Parker, Rollema, & Mead, 2010). Dual-controller theory explains these dissociations by suggesting that although pharmacotherapies modulate the expected value smoking and thus impact on goal-directed tobacco-seeking, they do not modify expected drug probability and so leave cue-elicited tobacco-seeking intact.

The aim of the current experiment was to test this dual-controller explanation for the partial efficacy of the front-line smoking cessation agent, nicotine replacement therapy (NRT). The design involved administering NRT as the devaluation treatment (versus chocolate satiety) in a procedure otherwise comparable to Hogarth & Chase (2011a). The key predictions were that NRT would modulate goal-directed tobacco choice in the extinction test, indicating that this agent can change the expected value of tobacco. By contrast, NRT should not modify the extent to which the tobacco picture enhances tobacco choice in the PIT test, indicating that this agent cannot modify expected drug probability. This dissociation would support dual-

controller theory of drug-seeking and explain the partial clinical efficacy of addiction pharmacotherapy.

Method

Participants

The study recruited 96 smokers, half of whom reported smoking on a daily basis and half of whom reported smoking less than every day of the week. The daily/non-daily smoker groups were balanced with respect to gender and assigned in equal proportions to the two devaluation treatments (NRT, chocolate). This recruitment strategy was undertaken to ensure broad variance in tobacco use, dependence and craving, such that these individual differences could be evaluated against behavioural effects.

All participants were informed that they might receive nicotine nasal spray, and reviewed a list (provided with the Nicorette® nasal spray 10ml pump) of conditions under which this medication should not be taken (specifically, ailments e.g. peptic ulcer, or concurrent medications e.g. clozapine, bupropion or other NRTs) and the potential side effects (e.g. headache) to report if experienced. Participants reviewed and signed this sheet at the outset of the laboratory session, and reviewed this information prior to recruitment during initial email communications. In addition, participants provided written informed consent, and the study was approved by the school of psychology ethics committee. No participant excluded themselves during the laboratory session or reported adverse reactions to the NRT administration either during or after the laboratory session.

Apparatus

Participants completed demographic then the following drug relevant questionnaires: the Diagnostic and Statistical Manual of Mental disorders tobacco dependence questionnaire

(Grant et al., 2003); the cigarette dependence scale (CDS-5 - Etter, Le Houezec, & Perneger, 2003); the brief questionnaire of smoking urges (QSU - Cox, Tiffany, & Christen, 2001) which yielded a factor 1 score reflecting desire for positive tobacco reward, and factor 2 score reflecting desire to avoid negative abstinence states, using the updated scoring system (Cappelleri et al., 2007); the alcohol use questionnaire which estimates units of alcohol consumed per week and binge drinking score (Townshend & Duka, 2005); and the assessment of substance misuse in adolescence (ASMA) scale which estimates illicit drug use and dependence (Willner, 2000 – using a 0-9 scale).

The computer task was programmed with E-prime running on a PC with 15-inch monitor, and responses were recorded using a five-key serial response box (Psychology Software Tools, Pittsburgh, USA). The rewards that participants believed they could earn during the computer task were held in two containers, one holding 20 Marlboro Lights cigarettes (Tar 6mg, Nicotine 0.5mg) and the other holding 20 Cadbury Dairy Milk Treatsize chocolate bars (15g with 4 chunks per bar). Two further containers were labelled “Your Cigarette Box” and “Your Chocolate Box”, in which participants cached the rewards they earned in the task.

The NRT devaluation treatment constituted two sprays from a Nicorette® nasal spray 10ml pump, with each spray delivering 0.5 mg of nicotine. The chocolate devaluation treatment constituted consumption of three 49g Cadbury Dairy Milk Chocolate bars, totalling 147g or 530Kcal. During payment, participants were told that ethics did not permit payment by cigarettes and chocolate (in contrast to what they had been told at the study outset), and asked whether the monetary equivalent of UK £5 would be acceptable. All participants accepted this arrangement, which when added to the base payment of UK £10, amounted to UK £15 in total.

Concurrent choice training

Figure 1 illustrates the procedure. Prior to concurrent training, participants first answered the questions: "I would like to smoke a cigarette right now" and "I would like to eat a chocolate bar right now" on 7 point Likert scales. Then, participants were presented with the following on-screen instructions: "*This is a game in which you can win cigarettes and chocolate. In each trial, press the left or right key to try and win these rewards. You will only win on some trials. Press any key to begin*". Each trial began with the centrally presented text, "Choose a key", which remained until either the leftmost or rightmost key of the five-key serial response box was pressed. A response on one key replaced this text with the outcome, "You win ¼ of a cigarette", whereas a response on the other key produced the outcome, "You win ¼ of chocolate bar". The key-outcome assignment was counterbalanced between-participants with respect to devaluation group, gender and smoker type (daily/non-daily). Only 1 outcome was scheduled to be available in each trial (at random), such that each key had only a 50% chance of yielding its respective outcome. On non-rewarded trials (in which the incorrect key was selected), the text "You win nothing" was presented. These three potential outcomes texts were presented for 1500msec, followed by a random inter-trial interval (ITI) between 1000 and 2000 msecs prior to the next trial.

Concurrent training comprised four 16 trial blocks. Earned outcomes were summed across trials and at the end of each block a "totalizer" screen reported the quantity of each reward type earned. Where whole cigarettes or chocolate bars had been earned, participants were instructed to move that many units from the loaded containers into their boxes present on the desk (see apparatus), such that the rewards were actually contacted. Any remainder of each reward type was added to the sum of the next block. The percent choice of the tobacco versus the chocolate key was recorded as the dependent measure.

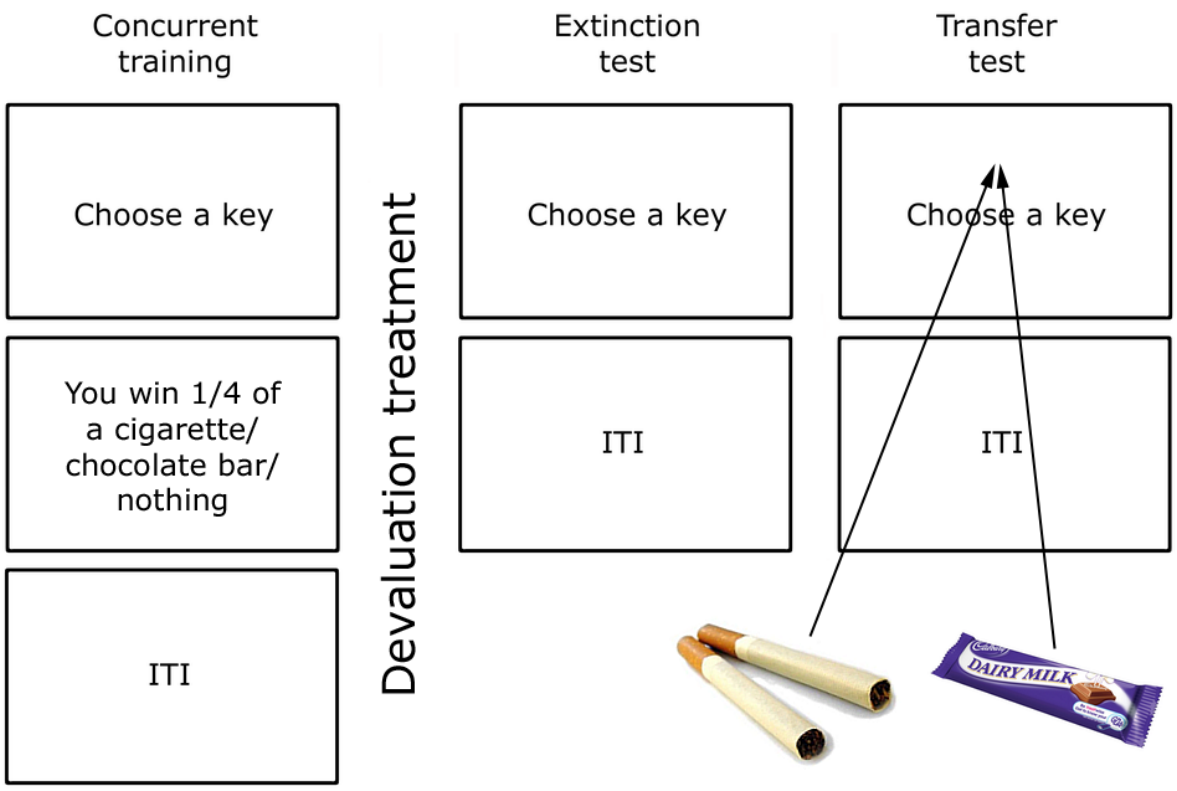

Figure 1: An illustration of the experiments procedures. See text for details. Arrows indicate the presentation of stimuli during the PIT test.

Contingency knowledge test 1

Awareness of the response-outcome contingencies was assessed following concurrent training by presentation of the on-screen instructions: "*We would now like to test whether you know which key earned which reward. Press any key to begin*", followed by the question: "*Which key earned cigarettes/chocolate, the left or the right key? Please choose carefully*". The order of presentation of the question about the two rewards was random and spaced by a 1000-2000msec ITI.

Devaluation treatments

The devaluation treatments were then conducted. Participants in the chocolate satiety condition were presented with three 49g Cadbury Dairy Milk Chocolate bars broken into 18 chunks on a plate and told: "*You will now be required to eat 3 bars of chocolate. You should*

try to eat all the chocolate but you can stop if you feel sick. Please complete the questionnaires while you are consuming the chocolate". During consumption, participants completed a questionnaire containing 18 visual analogue scales which rated the pleasantness of each successive chunk rated ranging from not very enjoyable to very enjoyable. Participants in the NRT group were handed a Nicorette® nasal spray 10ml pump and given the following instructions: (a) Tip head back slightly; (b) insert spray tip into one nostril; (c) point towards back of nose; (d) press firmly and quickly; (e) spray into other nostril. Each spray delivered approximately 0.5mg nicotine, totalling 1mg per participant. An eight minute wait period was then instituted to allow plasma nicotine concentrations to peak (Gourlay & Benowitz, 1997; Guthrie et al., 1999; Sutherland, Russell, Stapleton, Feyerabend, & Ferno, 1992), before a drug effects questionnaire was administered which contained four visual analogue scales assessing good effects, bad effects, head rush and high.

Contingency knowledge test 2

Participants' knowledge of the response-outcome contingencies was then assessed again to identify those who had forgotten the contingencies during the devaluation treatment.

Extinction test

The extinction test following devaluation comprised a single block of 24 trials identical to concurrent training apart from the omission of outcomes (see Figure 1). The on-screen instructions stated: "*In this part of the game, you will continue to earn cigarettes and chocolate bars in the same way as before. However, you will only be told how many of each reward you have earned at the end. Press any key to begin*". The trials were identical to concurrent training, but the outcomes were omitted. Instead, a left or right choice at the prompt, "Choose a key", simply launched a random ITI between 1000-2000 msec prior to the next trial. There were also no totalizer screens to provide feedback about outcomes

earned. We regard this test procedure as nominal extinction because despite the omission of outcomes, participants were led to believe that outcomes were nevertheless being earned. The purpose was to avoid experience of outcomes affecting choice, and the rapid extinction of responding that is typically found in devaluation procedures where the responses simply cease to earn rewards without prior instructions (e.g. Tricomi, Balleine, & O'Doherty, 2009). The dependent measure in the extinction test was again percent choice of the tobacco versus chocolate key, and the question was whether the devaluation treatments would modify choice relative to concurrent training.

Transfer test

The transfer test which followed extinction was headed by the instructions: "*In this part of the task, you can earn cigarettes and chocolate by pressing the left or right key in the same way as before. However, you will sometime be shown pictures before you choose which key to press. Press the space bar to begin*". The trial procedure was identical to extinction, i.e. there were no outcomes. However, the prompt "Choose a key" was compounded with either a cigarette or chocolate stimulus (see Figure 1) presented directly above the prompt, intermixed with no stimulus trials which were wholly identical to extinction trials. The transfer test totalled 24 trials, comprising 2 cycles of 12 where each cycle presented the cigarette, chocolate and no stimulus 4 times each in random order. Again, percent choice of the tobacco versus chocolate key was the dependent measure, and the question was whether cues would drive choice of the same outcome, and whether this effect would be modulated by devaluation treatment.

Results

Participants

Of 96 participants, five were excluded because they reported inaccurate knowledge of the response-outcome contingencies in either of the two knowledge tests. The characteristics of the remaining 91 participants, split by devaluation group, are shown in Table 1. There were no significant group differences in these characteristics.

Devaluation treatments

The chocolate devaluation group consumed an average of 62.7% (standard deviation, std = 3.7, range = 16.7-100) of the total 147g of chocolate to be consumed, and showed a decline in chocolate liking from 83.3% of the visual analogue scale (3.1, 3.6-100) to 21.5% (3.7, .0-92) from the first to last chocolate chunk consumed, $F(1,44) = 201.30$, $p < .001$. The NRT devaluation group reported the following subjective drug effects as percent of visual analogue scale: Good effects = 19.1% (3.1, 0-75); bad effects = 65.6% (3.8, 0-98.2); drug rush = 35.2% (4.6, 0-95.5); drug high = 26.5% (3.8, 0-87.5).

	NRT (n=46)	**Chocolate (n=45)**	***p*≤**
Gender ratio M:F	23:23	23:22	1
Age	20.2 (1.2, 18-24)	20.4 (1.7, 18-29)	.88
Smoking days per week	5.6 (1.9, 1-7)	4.9 (2.4, 1-7)	.34
Cigarettes on smoking days	5.7 (3.3, 1-15)	5.1 (3.8, 1-22.5)	.17
Smoking years	4.2 (1.9, 1-8)	3.6 (1.6, .6-7)	.15
Hours since a cigarette	17.3 (26.3, .1-160)	33.9(91.8, .1-552)	.30
Age of smoking onset	16.8 (1.8, 13-21)	17.4 (1.7, 14-21)	.15
Cigarette dependence scale	11.3 (3.8, 6-18)	10.1 (3.8, 5-20)	.14
DSM tobacco dependence	4.2 (1.8, .0-7)	4.1 (1.7, .0-7)	.70
Fagerstrom nicotine dependence	1.3 (1.8, .0-7)	.9 (1.5, .0-6)	.20
Smoking urges factor 1	3.7 (1.7, 1-7)	3.3 (1.6, 1-7)	.21
Smoking urges factor 2	1.7 (.7, 1-3.3)	1.9 (1.3, 1-7)	.64
1-item desire to smoke	3.6 (1.9, 1-7)	3.6 (1.9, 1-7)	.85
1-item desire for chocolate	3.1 (1.8, 1-7)	3.7 (1.9, 1-7)	.1
Units of alcohol per week	36.6 (26.6, 2.3-145)	36.2 (23.9, 3.8-108)	.91
Alcohol binge score	46.3 (29.8, 8-118)	41.6 (30.5, 2.6-158)	.46
% illicit substance use	71.7	68.0	.82
ASMA substance misuse score	1.4 (1.7, .0-6)	1.1 (1.5, .0-6)	.50

Table 1: Characteristics of the NRT and chocolate satiety groups.

Choice behaviour

Figure 2 shows the percent choice of tobacco versus chocolate key during the five conditions of the design. To examine goal-directed control over choice, the data from concurrent training and the extinction test (block variable) were entered into ANOVA with the variables devaluation group (NRT, chocolate). This analysis yielded a significant interaction between block and group, F (1,89) = 11.63, p = .001, with the block effect being significant in the chocolate devaluation group, F (1,44) = 18.00, p < .001, but not in the NRT devaluation group, F < 1. Thus, relative to concurrent training, chocolate devaluation reduced choice of the chocolate key in the extinction test that followed, demonstrating goal-directed control of action selection. However, NRT appeared to produce no such devaluation effect.

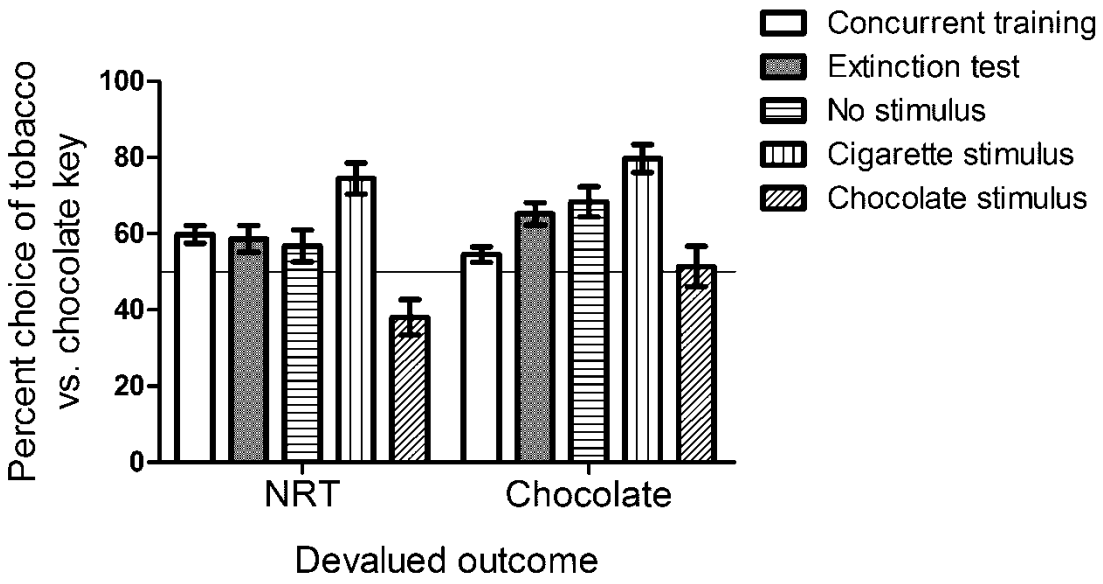

Figure 2: Percent choice of the tobacco versus chocolate key in concurrent choice training and extinction test, as well as the no stimulus, cigarette stimulus and chocolate stimulus condition of the transfer test.

To examine the PIT effect, data from the no stimulus, cigarette stimulus and chocolate stimulus condition were entered into ANOVA with the variables stimulus (3) and devaluation group (NRT, chocolate). There was a main effect of devaluation group, F (1,89) = 4.41, p = .04, indicating that the chocolate devaluation resulted in less choice of the chocolate key during the transfer test as a whole. There was also a main effect of stimulus, F (2,178) = 47.56, p < .001. However, crucially, there was no interaction between stimulus and group, F < 1, indicating that the ability of stimuli to selectively enhance choice of the same outcome was unaffected by whether NRT or chocolate had been devalued. The differential sensitivity of extinction compared to the transfer test to devaluation supports dual-controller theory of action selection.

Individual differences

Correlational analyses

As NRT produced no devaluation effect overall, individual differences in this effect were examined. An NRT devaluation score was calculated by subtracting the percent tobacco choice in the extinction test from that in concurrent training, such that positive scores reflected a greater reduction in tobacco choice following NRT. These scores from the NRT group (n=46) correlated significantly with number of cigarettes smoked on smoking days, r = -.41, p = .005, smoking urges factor 1, r = -.31, p = .03, the single item cigarette desire score, r = -.35, p = .02, and chocolate desire score, r = .29, p < .05, all obtained prior to concurrent training. These four proxies were translated into a compound smoking severity/desire score by $\log^{10}$ transforming them, to normalise their distribution, before summing them thus: cigarettes on smoking days + smoking urges factor 1 + (1-item cigarette desire – 1-item chocolate desire). The correlation between this compound smoking severity/desire and the NRT devaluation effect is shown in Figure 3A, r = -.57, p < .001, and

suggests that smokers of lower severity/desire showed an NRT devaluation effect, whereas smokers of higher severity/desire showed an NRT priming effect on their goal-directed tobacco choice in the extinction test following NRT. By contrast, the chocolate devaluation effect showed no such association with compound smoking severity/desire scores, r = -.02, p = .91, demonstrating that individual differences in severity/desire conferred sensitivity to NRT devaluation specifically, rather than sensitivity to devaluation generally.

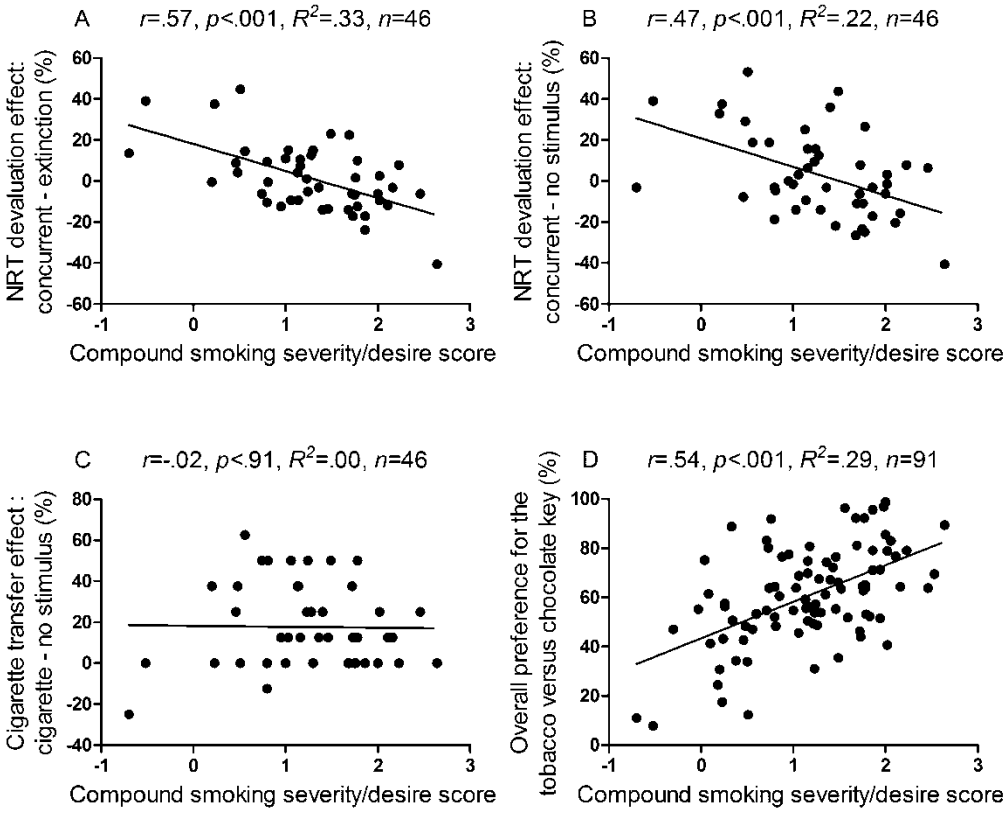

Figure 3: Correlations between compound smoking severity/desire scores and the NRT devaluation effect in the extinction test (A) and no stimulus condition of the transfer test (B) relative to concurrent training. (C) No relationship between compound smoking severity/desire scores in the NRT devaluation group and the capacity of the cigarette stimulus to increase tobacco choice above the no stimulus condition in the transfer test. (D) Correlation between compound smoking severity/desire scores and overall preference for the tobacco versus chocolate key (collapsed across all training and testing phases).

To examine whether the NRT devaluation effect persisted into the transfer test, the percent tobacco choice in the no stimulus condition was subtracted from that in concurrent training, such that positive scores reflected a greater reduction in tobacco choice following NRT in the no stimulus condition. As can be seen, in Figure 3B, the NRT devaluation effect in the no stimulus condition correlated with smoking severity/desire scores, r = -.47, p = .001. Moreover, the devaluation effects in the extinction test and no stimulus condition of the transfer test were themselves correlated, r = .61, p < .001, demonstrating individual consistency in the persistence of the devaluation effect across these two test periods. Finally, the chocolate devaluation effect in the no stimulus condition did not correlate with smoking severity/desire scores, r = .22, p = .14, again confirming that individual differences in severity/desire conferred specific sensitivity to NRT rather than devaluation generally.

By contrast, Figure 3C illustrates the insensitivity of the cigarette transfer effect to NRT devaluation. Here, a cigarette transfer score was calculated by subtracting percent tobacco choice in the cigarette stimulus condition from that in the no stimulus condition such that positive scores reflect an increased choice of the tobacco key driven by the cigarette stimulus. As can be seen in Figure 3C, the magnitude of this cueing effect on tobacco choice was not sensitive to individual differences in smoking severity/desire, r = -.02, p = .91. The three correlations therefore illustrate that whereas free choice of the tobacco key in the extinction and no stimulus conditions were sensitive to individual differences in NRT devaluation (Figure 3A/B), cued-choice of the tobacco key in the PIT test was autonomous of individual differences in NRT devaluation (Figure 3C).

Parametric analyses

ANCOVA was used to evaluate the individual differences in these devaluation effects more comprehensively (see Figure 4). Concurrent training and extinction test data (the block

variable) were entered with the compound smoking severity/desire score as a covariate and devaluation group (NRT, chocolate). There was a main effect of smoking severity/desire, F (1,87) = 37.87, $p < .001$ (illustrated in Figure 3D), demonstrating greater tobacco preference across both the concurrent and extinction phases as severity/desire increased, confirming the concurrent choice task as an assay of drug value (Hogarth & Chase, 2011a). More importantly, there was a significant interaction between severity/desire, group and block, F (1,87) = 8.63, $p = .004$, with the severity/desire by block interaction being reliable in the NRT devaluation group, F (1,44) = 21.56, $p < .001$ (illustrated in Figure 4), but not the chocolate devaluation group, $F < 1$. Thus, smoking severity/desire conferred sensitivity to NRT devaluation specifically, rather than devaluation generally. Participants as a whole were split into 3 ntiles by their smoking severity/desire score into a low, middle and high group. Table 2 and Figure 4 show data from these three ntiles within the NRT group. ANOVA contrasting the concurrent and extinction phase of the NRT group showed a significant NRT devaluation effect in the low group, F (1,10) = 6.23, $p = .03$, no change in the middle group, $F < 1$, and a significant NRT priming effect in the high group, F (1,17) = 4.95, $p = .04$.

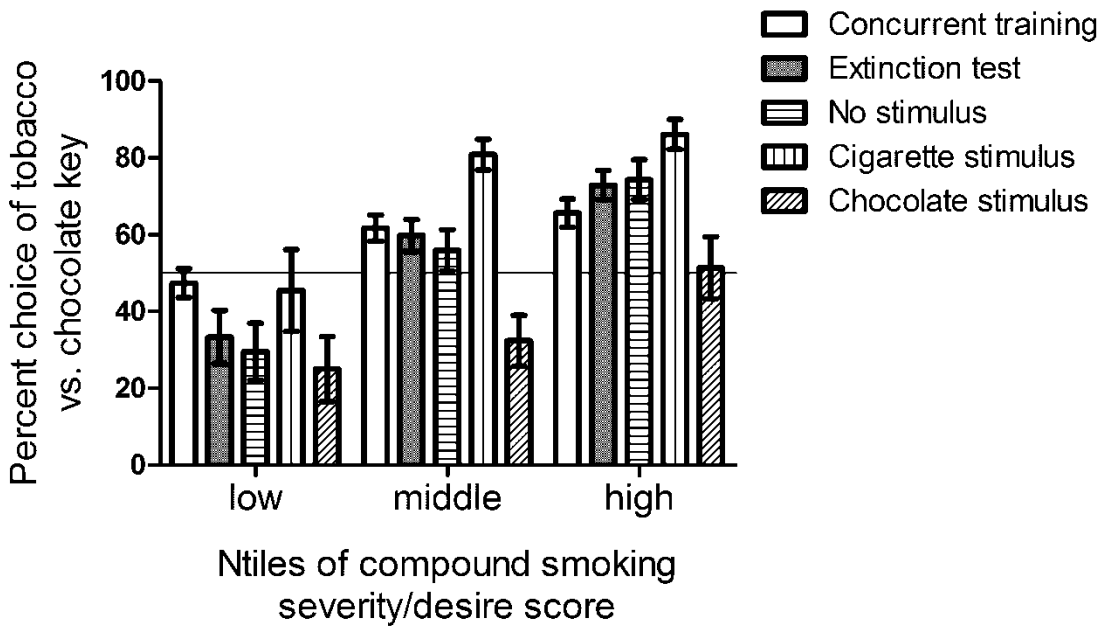

Figure 4: Percent choice of the tobacco versus chocolate key across all training and test phases, for the NRT devaluation group split into 3 ntiles on the basis of compound smoking severity/desire scores.

These NRT effects persisted into no stimulus trials of the transfer test. Concurrent and no stimulus data were entered into ANCOVA with smoking severity/desire as a covariate and devaluation group (NRT, chocolate). Although the main effect of severity/desire was again significant, F (1,87) = 37.53, $p < .001$, the interaction between severity/desire, group and block did not reach significance, F (1,87) = 1.25, $p = .27$. Nevertheless, the severity/desire by block interaction was reliable in the NRT group, F (1,44) = 12.73, $p = .001$, but not in the chocolate group, F (1,43) = 2.20, $p = .15$. Finally, specific contrasts on the 3 ntiles of the NRT group found a significant NRT devaluation effect in the low group, F (1,10) = 6.59, $p < .03$, no change in the middle group, F (1,16) = 1.81, $p = .20$, and a significant NRT priming effect in the high group, F (1,17) = 5.32, $p = .03$. Thus, baseline choice in the transfer test was mediated by individual differences in sensitivity to NRT devaluation.

	Low (n=11)	Middle (n=17)	High (n=18)	*p*≤
Compound smoking severity/desire score (lg^{10})	.3 (.5, -.7-.8)	1.2 (.2, .8-1.5)	2.0 (.3, 1.7-2.6)	.001
Gender ratio M:F	5:6	10:7	8:10	.66
Age	20.3 (1.3, 19-24)	20.2 (1.4, 18-22)	20.3 (1.1, 19-22)	.97
Smoking days per week	4.1 (2.4, 1-7)	5.6 (1.7, 3-7)	6.5 (1.1, 3-7)	.006
Cigarettes on smoking days	2.7 (1.0, 1-4)	6.4 (3.5, 2.5-15)	6.8 (2.9, 4-15)	.001
Smoking years	3.5 (2.1, 1.5-8)	4.1 (1.6, 2-7)	4.8 (2.0, 1-8)	.16
Hours since a cigarette	33.4 (45.0, 1-160)	16.4 (18.1, .1-60)	8.3 (8.6, .3-34)	.04
Age of smoking onset	17.4 (2.0, 13-20)	16.9 (1.5, 14-19)	16.3 (2.0, 13-21)	.17
Cigarette dependence scale	9.3 (2.8, 6-15)	11.0 (4.3, 6-18)	12.7 (3.3, 6-18)	.05
DSM tobacco dependence	3.4 (2.0, .0-6)	4.5 (1.4, 1-7)	4.4 (1.8, 1-7)	.22
Fagerstrom nicotine dependence	0(0, 0-0)	1.5 (1.8, .0-5)	2.1 (2.0, 0-7)	.002
Smoking urges factor 1	1.7 (.6, 1-3)	3.7 (1.3, 1.8-6.8)	4.9 (1.2, 2.4-7)	.001
Smoking urges factor 2	1.4 (.6, 1-2.7)	1.7 (.7, 1-3)	1.8 (.8, 1-3.3)	.31
1-item desire to smoke	1.7 (.5, 1-2)	2.8 (1.2, 1-6)	5.4 (1.2, 3-7)	.001
1-item desire for chocolate	3.5 (1.4, 2-5)	3.9 (2.1, 1-7)	2.0 (1.1, 1-5)	.006
Units of alcohol per week	25.2 (15.0, 5.6-58.8)	33.1 (19.6, 2.3-63.6)	46.8 (34.2, 5-145)	.20
Alcohol binge score	27 (18.4, 11-70)	41.6 (22.7, 8-77)	62.5 (33.5, 9-118)	.007
% illicit substance use	55	59	94	.02
ASMA substance misuse score	1.0(1.9, 0-6)	.8 (1.1, 0-3)	2.2 (1.9, 0-6)	.02
Good effect (%VAS)	15.3 (6.6, .0-70.5)	23.6 (5.6, .0-75.0)	17.3 (4.2, .0-71.4)	.53
Bad effect (%VAS)	60.2 (7.7, 24.1-96.4)	70.1 (5.9, 19.6-98.2)	64.7 (6.5, .0-98.2)	.61
Rush (%VAS)	43.4 (10.6, .0-95.5)	36.3 (7.9, .0-84.8)	29.0 (6.6, .0-76.8)	.48
High (%VAS)	27.2 (7.2, .0-65.2)	28.2 (6.6, .0-87.5)	24.5 (6.3, .0-75.9)	.91

Table 2: Characteristics of the NRT devaluation group split by 3 ntiles of the compound smoking severity/desire score into low, middle and high. VAS = visual analogue scale.

By contrast, the capacity of cues to augment choice of the same outcome showed no sensitivity to devaluation. ANCOVA was conducted on the no stimulus, cigarette stimulus and chocolate stimulus conditions (3) with smoking severity/desire as a covariate and devaluation group (NRT, chocolate). Again there was a main effect of severity/desire, $F(1,87) = 42.25$, $p < .001$, reflecting greater tobacco preference, and a main effect of stimulus, $F(2,174) = 13.48$, $p < .001$, reflecting cue-elicited choice of the same outcome. However, the stimulus variable showed no 2-way interaction with either severity/desire or devaluation group, and there was no 3-way interaction between all these variables, Fs < 1. Thus, cue-elicited choice of the same outcome was not modulated by the devaluation treatments and/or individual differences in smoking severity/desire. This autonomy of the PIT effect is most striking when one considers the specific contrast of low and high severity/desire participants in the NRT group shown in Figure 4. When the no stimulus and cigarette stimulus data from

these two groups were analysed, this stimulus (2) variable showed no interaction with severity/desire as either a covariate or a group variable, *F*s < 1. Thus, despite these two severity/desire groups showing a diametrically opposite change in the value of tobacco following NRT in the extinction and no stimulus conditions, by contrast, the capacity of the tobacco stimulus to control tobacco choice was entirely unaffected by this change in tobacco value. Such autonomy of the drug PIT effect to drug value is paradoxical because it must involve the cue retrieving of a representation of the drug outcome.

Remarkably, as shown in Table 2, the three severity/desire groups within the NRT devaluation sample showed no significant difference in their subjective response to NRT (good, bad, rush, high), and these subjective reactions did not correlate with either severity/desire scores or the NRT devaluation effect in either the extinction or no stimulus conditions relative to concurrent training, *r*s < .24, *p*s > .10. Thus, the differential NRT devaluation effects across these three severity/desire groups did not appear to be mediated by differential subjective reaction to NRT.

Discussion

The key findings of the study were that choice between tobacco and chocolate points in the extinction test was modified by outcome devaluation, indicating that this behaviour was goal-directed in being determined by the current value of the outcomes. By contrast, the capacity of tobacco and chocolate pictures to enhance choice of the same outcome in the PIT test was not modified by outcome devaluation indicating that this behaviour was determined by the expected probability of the outcome independently of its value, corroborating other studies showing this dissociation (Colwill & Rescorla, 1990; Corbit, et al., 2007; Hogarth & Chase, 2011a; Holland, 2004; Rescorla, 1994). The study therefore provided evidence for dual-controller theory wherein action selection is determined independently by the expected

value and probability of the outcome. Furthermore, the study demonstrated the importance of dual-controller theory for understanding the psychological basis of drug-seeking per se, and the partial efficacy of addiction pharmacotherapy. Specifically, NRT modified the expected value of tobacco underpinning goal-directed tobacco choice in the extinction test, but did not modify the expected probability of tobacco underpinning cue-elicited tobacco choice in the PIT test. This dissociation suggests that the propensity to engage in drug-seeking is determined independently by the expected value and probability of the drug, and that the partial efficacy of pharmacotherapy is due to its selective effect on expected drug value.

To summarise the results more systematically; first, chocolate satiety produced a reduction of chocolate choice in the extinction test compared to concurrent training. This effect demonstrates goal-directed control of action selection, because, in order to modify choice in the extinction test participants must have integrated knowledge of the response-outcome (R-O) contingencies acquired in concurrent training with knowledge of the current low value of the chocolate outcome acquired during the devaluation treatment to determine the propensity to make the chocolate choice. This effect replicates previous demonstrations of goal-directed control of naturally rewarded action selection in humans (de Wit, Barker, Dickinson, & Cools, 2011; de Wit, Corlett, Aitken, Dickinson, & Fletcher, 2009; de Wit, Niry, Wariyar, Aitken, & Dickinson, 2007; Hogarth, Attwood, Bate, & Munafò, 2012; Hogarth, Chase, & Baess, 2012; Kenward, Folke, Holmberg, Johansson, & Gredeback, 2009; Klossek, Russell, & Dickinson, 2008; L. Schwabe & Wolf, 2009; Lars Schwabe & Wolf, 2011; Tricomi, et al., 2009; Valentin, Dickinson, & O'Doherty, 2007).

Acute NRT administration, by contrast, produced no change in tobacco choice between concurrent training and the extinction test in participants as a whole. Instead, the effect of NRT on tobacco choice depended on participants' level of smoking severity and desire. That

is, participants who reported low smoking severity/desire reduced their choice of the tobacco key in the extinction test following NRT compared to concurrent training (a devaluation effect), whereas participants who reported high smoking severity/desire increased their choice of the tobacco key (a priming effect). Given that these changes in choice must be mediated by retrieval of the current value of tobacco, we can conclude that a 1mg dose of nicotine replacement nasal spray differentially modified the value of tobacco according to smokers' level of smoking severity/desire.

There are several explanations of how NRT modified tobacco choice in the extinction test. According to the incentive learning account, smokers have learned in their extra-experimental history that the current value of tobacco is dependent on internal states for smoking deprivation and satiety, which enables these states to raise and lower the expected value of tobacco, respectively (Dickinson & Balleine, 2002; Hutcheson, Everitt, Robbins, & Dickinson, 2001). Arguably, NRT modified the expected value of tobacco by mimicking such internal states which have themselves undergone incentive learning. In order to explain the individual differences, it must be further proposed that the optimally rewarding dose of nicotine (Corrigall & Coen, 1989) varies with level of nicotine dependence, wherein more dependent individual achieve optimal reinforcement with a higher dose (Pomerleau, 1995). On this basis, one might argue that the 1mg dose of NRT exceeded the optimal dose for low smokers and thus mimicked an internal satiety state which in the past has predicted that further nicotine ingestion would be unrewarding or aversive. Accordingly, this NRT induced satiety state reduced the expected value of tobacco and thus reduced tobacco choice in the extinction test (produced a devaluation effect) in low smokers.

It is less obvious why the 1mg NRT dose produced a priming effect in high smokers. On the incentive learning account, the 1 mg NRT dose fell below the optimally rewarding dose for

this group and thus mimicked the internal state for initial smoking. Arguably, this state has in the past provided a salient and contiguous signal that further smoking will be rewarding, and so NRT mimicry of this state raised the expected value of tobacco and thus increased tobacco choice in the extinction test. A more troubling explanation of this priming effect, however, is that NRT acted as a discriminative stimulus to retrieve the O-R association acquired in concurrent training, thus priming tobacco choice in much the same way as the smoking pictures in the PIT test (Ostlund & Balleine, 2007b). Stated less formally, NRT reminded high smokers of tobacco points through generalisation, which in turn primed choice of that outcome. This view, however, would predict that high smokers are more prone to O-R based control of choice than low smokers, and thus should show a bigger tobacco PIT effect, which Figure 3C shows was not the case. One might counter argue by suggesting an interaction between systems, wherein NRT enhanced tobacco choice via the O-R retrieval mechanism equivalently across the sample as a whole, but this effect was counteracted by NRT inducing satiety in the low smokers thus reducing tobacco choice in this group. This interactive account, however, finds disfavour with the observation that outcome induced priming of choice is not sensitive to devaluation of the outcome (Ostlund & Balleine, 2007b). Thus, incentive learning provides the most parsimonious explanation for the NRT priming effect on goal-directed tobacco choice.

Another consideration in explaining the NRT effect on choice is the possibility that health related cognitions evoked by the NRT administration, rather than internal drug states, played a role. The current design is exposed to this interpretation because no placebo group was included where cognitions arising from NRT administration were matched without active nicotine being ingested. Moreover, the effect of health warnings on goal-directed tobacco choice has been demonstrated in a comparable procedure (Hogarth & Chase, 2011a). The key weakness of this explanation, however, is that smoking health warnings produced an

equal devaluation effect across levels of smoking severity in this earlier study, whereas NRT in the current study produced opposing devaluation/priming effects across levels of smoking severity. Thus, the effect of NRT on goal-directed tobacco choice cannot be readily explained by health related cognitions accompanying NRT administration. Overall, therefore, it is concluded, that incentive learning provides the most compelling account of NRT effects on tobacco choice: Arguably, NRT mimicked internal states associated with pre- and post-optimal nicotine ingestion in high and low smokers respectively, thus producing corresponding changes in the expected value of tobacco and hence selection of goal-directed tobacco choice.

In contrast to the extinction test, cue-elicited choice in the PIT test was insensitive to the devaluation treatments. To be more specific, chocolate devaluation reduced chocolate choice in the extinction test and in the PIT test (overall) compared to concurrent training, but NRT administration produced no change in these measures. Thus, the devaluation treatments produced a divergence in goal-directed choice. By contrast, neither chocolate devaluation nor NRT administration modified the extent to which the chocolate or tobacco stimuli enhanced choice of the same outcome, over the no-stimulus condition, in the PIT test. Thus, the capacity of cues to bias choice towards the signalled outcome was autonomous of the devaluation treatments. This dissociation between free- versus cued-choice to outcome devaluation confirms related animal studies (Colwill & Rescorla, 1990; Corbit, et al., 2007; Holland, 2004; Rescorla, 1994) and one human study in which free-choice but not cued-choice was sensitive to outcome devaluation produced by health warnings (Hogarth & Chase, 2011a). This behavioural dissociation also accords studies which have shown goal-directed and transfer-cue-elicited action to be mediated by dissociable neural substrates (Corbit & Balleine, 2003; Ostlund & Balleine, 2007a). Together, these studies support dual controller theory in suggesting that outcome evaluation and cue-elicited outcome prediction

converge independently on action selection (Balleine, et al., 2011; de Wit & Dickinson, 2009).

The prevailing interpretation of the PIT effect, when applied to the current data, is that Pavlovian or predictive learning outside the experiment endowed tobacco and chocolate stimuli with the capacity to elicit an expectation of a range of associated outcomes, including for example, smoking and eating chocolate. In addition, concurrent training endowed participants with instrumental knowledge of the response-points contingencies, which were bidirectional such that retrieval of a representation of the points could elicit the associated response; a process known as ideomotor or O-R control (Dutzi & Hommel, 2009). Consequently, during the PIT test, the tobacco and chocolate stimuli retrieved an expectation of a range of their associated outcomes (S-O), which through a process of generalisation (McLaren & Mackintosh, 2002) and inference (Mitchell, De Houwer, & Lovibond, 2009), increased the expected probability of the instrumental outcomes (points) with shared category membership (Wills & Pothos, 2012), which in turn elicited the associated response through the O-R or ideomotor link. In short, PIT was mediated by an S-O-R inference between predictive and instrumental knowledge (Balleine, et al., 2011; de Wit & Dickinson, 2009).

The paradoxical implication of this S-O-R account is that although the tobacco and chocolate stimuli must have retrieved a representation of tobacco or chocolate points respectively, in order to bias choice towards the associated response, these outcome representations did not make contact with their current incentive value, whereas the outcome representation retrieved during goal-directed choice did. Rather, the PIT effect appears to depend upon the signalled probability of the outcome. In support of this claim, the magnitude of the PIT effect can be reduced by degradation of the S-O predictive contingency prior to the PIT test

(Delamater, 1995; see also Gámez & Rosas, 2005) and shows an orderly decline as the strength of the predictive S-O contingency is reduced (Trick, et al., 2011). On this basis, it is argued that whilst the expected value of the outcome determines the overall propensity to engage in goal-directed choice of that outcome, the expected probability of the outcome adds an order of magnitude to this propensity, but this magnitude is itself not weighted by value. Hence action selection reflects the summation of two independent controllers.

Expected utility theory provides the key theoretical construction against which to juxtapose this dual-controller account. According to utility theory, the propensity to select an action is determined by the expected utility of the outcome, that is, the expected value of the outcome weighted by its probability (Rangel, Camerer, & Montague, 2008; Vlaev, Chater, Stewart, & Brown, 2011). Although expected value and probability may be encoded within separate neural channels, ultimately they are amalgamated prior to action selection (Knutson, Taylor, Kaufman, Peterson, & Glover, 2005). Thus, the cue effect on choice should be modulated by devaluation if mediated by expected utility, and insensitive to devaluation if mediated by expected probability in isolation of value. Clearly, the current data favour this latter claim of the dual-controller hypothesis.

The strongest challenge to dual-controller model is the finding that the PIT effect can be modulated by outcome devaluation (Allman, DeLeon, Cataldo, Holland, & Johnson, 2010). In this study, participants played the role of a stockbroker trading on foreign currency. In Pavlovian training they learned whether companies, signalled by their logos, traded in Hong Kong or American Dollars. Then, in the instrumental phase, participants sought to earn Hong Kong or American Dollars by pressing different keys. Devaluation was achieved by informing participants that one or other currency had crashed rendering it worthless. Finally, choice between earning the two currencies was tested in the presence of the company logos

established in Pavlovian training. The results showed that logos selectively enhanced choice of their associated currency only if that currency was still valued. Cues that predicted the devalued currency lost their capacity to enhance choice of that outcome.

Resolving the discrepancy between Allman, et al. (2010) and studies which have found no such effect of devaluation on PIT (Colwill & Rescorla, 1990; Corbit, et al., 2007; Hogarth & Chase, 2011a; Holland, 2004; Rescorla, 1994) relies on identifying the procedural variables that are unique to Allman, et al. (2010). In this regard, one cannot plausibly appeal to species (rat, human), instrumental response measure (rate, forced-choice), devaluation method (instructions, satiety, taste aversion), encoding of contingencies (propositional, associative), length of training (short, long) or type of stimuli (conditioned, pictorial, discriminative). Perhaps the most plausible explanation is that Allman, et al.'s (2010) devaluation instructions which indicated that one currency was now worthless inadvertently implied to participants that the companies that previously traded in this currency had ceased to do so, i.e. degraded beliefs about the cue-outcome probability. This claim is bolstered by the finding that instructions that downgrade beliefs about cue-outcome relationships immediately attenuate conditioned responding to cues (Field & Duka, 2001; Lovibond, 2004). Moreover, beliefs about the status of a cue-outcome contingency can change through an inference process following modification of other related contingencies (Dickinson & Burke, 1996; Lovibond, 2003). Such an inference process may have created ambivalence about the predictive status of logos, such that they did not evoke an expectation of the devalued currency, and consequently, did not prime choice of that currency, rather than because these logos retrieved an expectation of the low value of the devalued currency. This explanation resolves the discrepancy posed by Allman, et al. (2010) by suggesting that their devaluation treatment was more akin to extinction manipulations which have been shown to attenuate the PIT effect (Delamater, 1995; Gámez & Rosas, 2005; Trick, et al., 2011) than

with ‘standard’ devaluation protocols which do not impact on PIT (Colwill & Rescorla, 1990; Corbit, et al., 2007; Hogarth & Chase, 2011a; Holland, 2004; Rescorla, 1994).

Is should be pointed out the S-O-R theory espoused to explain the autonomy of transfer effects has prescribed limits as a general theory of stimulus control. Specifically, in Pavlovian devaluation procedures, stimuli are trained to signal that a response will be rewarded, and consequently these stimuli acquire the capacity to motivate the performance of the response. Crucially, such stimulus control of responding is modulated by outcome devaluation, indicating that the stimulus must retrieve a representation of the current value of the outcome to determine responding, i.e. an S-O-R chain (Colwill & Motzkin, 1994; Holland & Straub, 1979; Pickens et al., 2003; Rescorla, 1991). These findings are clearly at odds with the autonomy of the transfer effect, so one must consider how these forms of stimulus control differ. Resolution of this discrepancy may lie in the additional computational step required in transfer, that is, the generalisation between the outcome predicted by the stimulus and the outcome delivered by the response. Presumably, encoding of outcome value is sacrificed during such generalisation resulting in the autonomy of PIT from devaluation. On this view, the Pavlovian devaluation effect relies on the outcome predicted by the stimulus and earned by the response being identical, precluding generalisation thus enabling outcome value to modulate the propensity to response.

The dual-controller account elaborated above provides a framework for explaining why tonic cigarette craving but not cue-elicited craving is modulated by abstinence/satiety (Drobes & Tiffany, 1997; Hogarth, et al., 2010; Maude-Griffin & Tiffany, 1996) nicotine replacement therapy (Havermans, et al., 2003; Morissette, et al., 2005; Niaura, et al., 2005; Rohsenow, et al., 2007; Shiffman, et al., 2003; Tiffany, et al., 2000; Waters, et al., 2004), bupropion (Hussain, et al., 2010) and varenicline (Brandon, et al., 2011; Franklin, et al., 2011; Hitsman,

et al., under review; Hitsman, et al., 2006). The current analysis suggests that these treatments induce or mimic internal states which have undergone incentive learning and so modulate expectations regarding the current incentive value of tobacco, which determines overall propensity to engage in goal-directed tobacco-seeking (tonic craving). By contrast, these manipulations do not affect the capacity of smoking cues to enhance the expected probability of tobacco, which underpins cue-elicited craving and tobacco-seeking (Carter & Tiffany, 2001; Dols, Hout, Kindt, & Willems, 2002), and so these treatments are ineffective in modulating such cue-reactivity.

Dual-controller theory can also explain two puzzling effects of pharmacotherapy on smoking cessation outcomes in clinical trials. First, (Waters, et al., 2004) found that cue-elicited craving measured prior to quitting only predicted smoking cessation in a group that had been treated with NRT during the cue-test and cessation phase, but not in the group so treated with placebo. Dual-controller theory can explain this finding by suggesting that NRT selectively attenuated goal-directed tobacco-seeking, leaving cue-elicited tobacco-seeking as the principle determinant of craving and relapse, thus strengthening the correlation between these measures. By contrast, in the placebo group, both controllers contributed to relapse, thus degrading the correlation between this relapse and cue-elicited craving. The second finding by (Brandon, et al., 2011) was that although cue-elicited craving was not attenuated by acute varenicline, it was attenuated by chronic varenicline that was administered alongside continued smoking (so called preloading; Lindson & Aveyard, 2011). Dual-controller theory explains this by suggesting that chronic varenicline blocked experience of smoking reward (West, Baker, Cappelleri, & Bushmakin, 2008) thus degrading the cue-outcome contingency by reducing smoking whilst cue exposure remained constant. Thus, chronic varenicline arguably attenuated cue-induced craving by promoting extinction rather than by reducing expectations of drug value retrieved by the cue.

To summarise, the study found that the pharmacotherapy for smoking cessation, NRT, attenuated goal-directed but not transfer cue-elicited tobacco choice. This dissociation substantiates the view that expected drug value and probability independently determine drug-seeking, and that pharmacotherapy selectively impacts on expected drug value. The study therefore offers a principled explanation for the partial clinical efficacy of addiction pharmacotherapy, and a behavioural model for screening combined therapies which seek to achieve broader protection by impacting on both controllers. Thus, identifying the neuropharmacological substrates of the dual-controllers underpinning drug-seeking should be a priority for improving the efficacy of addiction pharmacotherapies.

Reference

Allman, M. J., DeLeon, I. G., Cataldo, M. F., Holland, P. C., & Johnson, A. W. (2010). Learning processes affecting human decision making: An assessment of reinforcer-selective Pavlovian-to-instrumental transfer following reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes, 36*(3), 402-408. doi: 10.1037/a0017876

Balleine, B. W., Leung, B. K., & Ostlund, S. B. (2011). The orbitofrontal cortex, predicted value, and choice. *Annals of the New York Academy of Sciences, 1239*(1), 43-50. doi: 10.1111/j.1749-6632.2011.06270.x

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology, 35*(1), 48-69. doi: 10.1038/npp.2009.131

Balleine, B. W., & Ostlund, S. B. (2007). Still at the choice-point - Action selection and initiation in instrumental conditioning *Reward and Decision Making in Corticobasal Ganglia Networks* (Vol. 1104, pp. 147-171).

Belin, D., Jonkman, S., Dickinson, A., Robbins, T. W., & Everitt, B. J. (2009). Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction. *Behavioural Brain Research, 199*(1), 89-102. doi: 10.1016/j.bbr.2008.09.027

Brandon, T., Drobes, D., Unrod, M., Heckman, B., Oliver, J., Roetzheim, R., . . . Small, B. (2011). Varenicline effects on craving, cue reactivity, and smoking reward. *Psychopharmacology, 218*(2), 391-403. doi: 10.1007/s00213-011-2327-z

Cappelleri, J. C., Bushmakin, A. G., Baker, C. L., Merikle, E., Olufade, A. O., & Gilbert, D. G. (2007). Multivariate framework of the brief questionnaire of smoking urges. [Article]. *Drug and Alcohol Dependence, 90*(2-3), 234-242. doi: 10.1016/j.drugalcdep.2007.04.002

Carter, B. L., & Tiffany, S. T. (2001). The cue-availability paradigm: The effects of cigarette availability on cue reactivity in smokers. *Experimental and Clinical Psychopharmacology, 9*(2), 183-190.

Colwill, R. M., & Motzkin, D. K. (1994). Encoding of the unconditioned stimulus in Pavlovian conditioning. *Animal Learning & Behavior, 22*(4), 384-394. doi: 10.3758/BF03209158

Colwill, R. M., & Rescorla, R. A. (1990). Effects of reinforcer devaluation on discriminative control of instrumental behavior. *Journal of Experimental Psychology-Animal Behavior Processes, 16*(1), 40-47.

Corbit, L. H., & Balleine, B. W. (2003). The role of prelimbic cortex in instrumental conditioning. *Behavioural Brain Research, 146*(1-2), 145-157. doi: 10.1016/j.bbr.2003.09.023

Corbit, L. H., Janak, P. H., & Balleine, B. W. (2007). General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *European Journal of Neuroscience, 26*(11), 3141-3149. doi: 10.1111/j.1460-9568.2007.05934.x

Corrigall, W. A., & Coen, K. M. (1989). Nicotine maintains robust self-administration in rats on a limited-access schedule. *Psychopharmacology, 99*(4), 473-478. doi: 10.1007/BF00589894

Cox, L. S., Tiffany, S. T., & Christen, A. G. (2001). Evaluation of the brief questionnaire of smoking urges (QSU-brief) in the laboratory and clinical settings. *Nicotine and Tobacco Research, 3*, 7-16. doi: 10.1080/14622200124218

de Wit, S., Barker, R. A., Dickinson, A. D., & Cools, R. (2011). Habitual versus goal-directed action control in Parkinson disease. *Journal of Cognitive Neuroscience, 23*(5), 1218-1229. doi: 10.1162/jocn.2010.21514

de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., & Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *Journal of Neuroscience, 29*(36), 11330-11338. doi: 10.1523/jneurosci.1639-09.2009

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychological Research, 73*, 463-476. doi: 10.1007/s00426-009-0230-6

de Wit, S., Niry, D., Wariyar, R., Aitken, M. R. F., & Dickinson, A. (2007). Stimulus-outcome interactions during instrumental discrimination learning by rats and humans. *Journal of Experimental Psychology-Animal Behavior Processes, 33*(1), 1-11. doi: 10.1037/0097-7403.33.1.1

Delamater, A. R. (1995). Outcome-selective effects of intertrial reinforcement in Pavlovian appetitive conditioning with rats. *Animal Learning & Behavior, 23*(1), 31-39. doi: 10.3758/BF03198013

Dickinson, A., & Balleine, B. W. (2002). The role of learning in the operation of motivational systems. In C. R. Gallistel (Ed.), *Stevens' Handbook of Experimental Psychology, Vol 3, Learning, Motivation and Emotion, Ed 3* (pp. pp 497-533). New York: Wiley.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 49*(1), 60-80.

Dols, M., Hout, M. v. d., Kindt, M., & Willems, B. (2002). The urge to smoke depends on the expectation of smoking. *Addiction, 97*(1), 87-93. doi: 10.1046/j.1360-0443.2002.00010.x

Drobes, D. J., & Tiffany, S. T. (1997). Induction of smoking urge through imaginal and in vivo procedures: Physiological and self-report manifestations. *Journal of Abnormal Psychology, 106*(1), 15-25.

Dutzi, I. B., & Hommel, B. (2009). The microgenesis of action-effect binding. *Psychological Research 73*, 425-435.

Etter, J. F., Le Houezec, J., & Perneger, T. V. (2003). A self-administered questionnaire to measure dependence on cigarettes: The Cigarette Dependence Scale. *Neuropsychopharmacology, 28*(2), 359-370. doi: 10.1038/sj.npp.1300030

Ferguson, S. G., & Shiffman, S. (2009). The relevance and treatment of cue-induced cravings in tobacco dependence. *Journal of Substance Abuse Treatment, 36*(3), 235-243. doi: 10.1016/j.jsat.2008.06.005

Field, M., & Duka, T. (2001). Smoking expectancy mediates the conditioned responses to arbitrary smoking cues. *Behavioural Pharmacology, 12*(3), 183-194.

Franklin, T., Wang, Z., Suh, J. J., Hazan, R., Cruz, J., Li, Y., . . . Childress, A. R. (2011). Effects of varenicline on smoking cue-triggered neural and craving responses. *Archives of General Psychiatry*, archgenpsychiatry.2010.2190. doi: 10.1001/archgenpsychiatry.2010.190

Gámez, A. M., & Rosas, J. M. (2005). Transfer of stimulus control across instrumental responses is attenuated by extinction in human instrumental conditioning. *International Journal of Psychology & Psychological Therapy, 5*(3), 207-222.

Gourlay, S. G., & Benowitz, N. L. (1997). Arteriovenous differences in plasma concentration of nicotine and catecholamines and related cardiovascular effects after smoking, nicotine nasal spray, and intravenous nicotine. *Clinical Pharmacology & Therapeutics, 62*(4), 453-463.

Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, P. S., Kay, W., & Pickering, R. (2003). The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence, 71*(1), 7-16.

Guthrie, S. K., Zubieta, J. K., Ohl, L., Ni, L., Koeppe, R. A., Minoshima, S., & Domino, E. F. (1999). Arterial/venous plasma nicotine concentrations following nicotine nasal spray. *European Journal of Clinical Pharmacology, 55*(9), 639-643.

Havermans, R. C., Debaere, S., Smulders, F. T. Y., Wiers, R. W., & Jansen, A. T. M. (2003). Effect of cue exposure, urge to smoke, and nicotine deprivation on cognitive performance in smokers. *Psychology of Addictive Behaviors, 17*(4), 336-339. doi: 10.1037/0893-164x.17.4.336

Hitsman, B., Hogarth, L., Teige, Shadel, DiBendetti, Price, & Niaura, R. (under review). Dissociable effect of acute varenicline on tonic versus cue-provoked craving in non-treatment motivated smokers.

Hitsman, B., Niaura, R., Shadel, W. G., Britt, D. M., & Price, L. H. (2006). Effect of varenicline on cue-provoked cigarette craving and acute nicotine withdrawal. *Neuropsychopharmacology, 31*, S75-S76.

Hogarth, L., Attwood, A. S., Bate, H. A., & Munafò, M. R. (2012). Acute alcohol impairs human goal-directed action. *Biological Psychology, in press*. doi: 10.1016/j.biopsycho.2012.02.016

Hogarth, L., & Chase, H. W. (2011a). Parallel goal-directed and habitual control of human drug-seeking: Implications for dependence vulnerability. *Journal of Experimental Psychology: Animal Behavior Processes, 37*(3), 261-276. doi: 10.1037/a0022913

Hogarth, L., & Chase, H. W. (2011b). Vulnerabilities underlying human drug dependence: goal-valuation versus habit learning. In M. Haselgrove & L. Hogarth (Eds.), *Clinical Applications of Learning Theory* (pp. 75-101). Hove: Psychology Press.

Hogarth, L., & Chase, H. W. (2012). Evaluating psychological markers for human nicotine dependence: Tobacco choice, extinction and Pavlovian to instrumental transfer. *Experimental and Clinical Psychopharmacology, in press*.

Hogarth, L., Chase, H. W., & Baess, K. (2012). Impaired goal-directed behavioural control in human impulsivity. *Quarterly Journal of Experimental Psychology, 65*(2), 305-316. doi: 10.1080/17470218.2010.518242

Hogarth, L., Dickinson, A., & Duka, T. (2010). The associative basis of cue elicited drug taking in humans. *Psychopharmacology, 208*(3), 337-351. doi: DOI 10.1007/s00213-009-1735-9

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology-Animal Behavior Processes, 30*(4), 258-258.

Holland, P. C., & Straub, J. J. (1979). Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 5*(1), 65-78. doi: 10.1037/0097-7403.5.1.65

Hussain, S., Zawertailo, L., Busto, U., Zack, M., Farvolden, P., & Selby, P. (2010). The impact of chronic bupropion on plasma cotinine and on the subjective effects of ad lib smoking: A randomized controlled trial in unmotivated smokers. *Addictive Behaviors, 35*(2), 164-167. doi: 10.1016/j.addbeh.2009.09.004

Hutcheson, D. M., Everitt, B. J., Robbins, T. W., & Dickinson, A. (2001). The role of withdrawal in heroin addiction: enhances reward or promotes avoidance? *Nature Neuroscience, 4*(9), 943-947. doi: 10.1038/nn0901-943

Kenward, B., Folke, S., Holmberg, J., Johansson, A., & Gredeback, G. (2009). Goal Directedness and Decision Making in Infants. *Developmental Psychology, 45*(3), 809-819. doi: 10.1037/a0014076

Klossek, U. M. H., Russell, J., & Dickinson, A. (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *Journal of Experimental Psychology-General, 137*(1), 39-51. doi: 10.1037/0096-3445.137.1.39

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *The Journal of Neuroscience, 25*(19), 4806-4812. doi: 10.1523/jneurosci.0642-05.2005

Lindson, N., & Aveyard, P. (2011). An updated meta-analysis of nicotine preloading for smoking cessation: investigating mediators of the effect. *Psychopharmacology, 214*(3), 579-592. doi: 10.1007/s00213-010-2069-3

Lovibond, P. F. (2003). Causal beliefs and conditioned responses: Retrospective revaluation induced by experience and by instruction. *Journal of Experimental Psychology-Learning Memory and Cognition, 29*(1), 97-106.

Lovibond, P. F. (2004). Cognitive processes in extinction. *Learning & Memory, 11*(5), 495-500.

Maude-Griffin, P. M., & Tiffany, S. T. (1996). Production of smoking urges through imagery: The impact of affect and smoking abstinence. *Experimental and Clinical Psychopharmacology, 4*(2), 198-208.

McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior, 30*(3), 177-200.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*(2), 183-+. doi: 10.1017/s0140525x09000855

Morissette, S. B., Palfai, T. P., Gulliver, S. B., Spiegel, D. A., & Barlow, D. H. (2005). Effects of transdermal nicotine during imaginal exposure to anxiety and smoking cues in college smokers. *Psychology of Addictive Behaviors, 19*(2), 192-198. doi: 10.1037/0893-164x.19.2.192

Niaura, R., Sayette, M., Shiffman, S., Glover, E. D., Nides, M., Shelanski, M., . . . Sorrentino, J. (2005). Comparative efficacy of rapid-release nicotine gum versus nicotine polacrilex gum in relieving smoking cue-provoked craving. *Addiction, 100*(11), 1720-1730. doi: 10.1111/j.1360-0443.2005.01218.x

O'Connor, E. C., Parker, D., Rollema, H., & Mead, A. N. (2010). The alpha 4 beta 2 nicotinic acetylcholine-receptor partial agonist varenicline inhibits both nicotine self-administration following repeated dosing and reinstatement of nicotine seeking in rats. *Psychopharmacology, 208*(3), 365-376. doi: 10.1007/s00213-009-1739-5

Ostlund, S. B., & Balleine, B. W. (2007a). Orbitofrontal cortex mediates outcome encoding in pavlovian but not instrumental conditioning. *Journal of Neuroscience, 27*(18), 4819-4825.

Ostlund, S. B., & Balleine, B. W. (2007b). Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome. *Learning & Behavior, 35*(1), 43-52.

Ostlund, S. B., & Balleine, B. W. (2008). The disunity of Pavlovian and instrumental values. *Behavioral and Brain Sciences, 31*(4), 456. doi: 10.1017/s0140525x08004925

Pickens, C. L., Saddoris, M. P., Setlow, B., Gallagher, M., Holland, P. C., & Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *Journal of Neuroscience, 23*(35), 11078-11084.

Pomerleau, O. (1995). Individual differences in sensitivity to nicotine: Implications for genetic research on nicotine dependence. *Behavior Genetics, 25*(2), 161-177. doi: 10.1007/bf02196925

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545-556. doi: 10.1038/nrn2357

Rescorla, R. A. (1991). Associative relations in instrumental learning - The 18 Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 43*(1), 1-23.

Rescorla, R. A. (1994). Transfer of instrumental control mediated by a devalued outcome. *Animal Learning & Behavior, 22*(1), 27-33. doi: 10.3758/BF03199953

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of drug addiction. *Brain Research Reviews, 18*, 247-291. doi: 10.1016/0165-0173(93)90013-P

Rohsenow, D. J., Monti, P. M., Hutchison, K. E., Swift, R. M., MacKinnon, S. V., Sirota, A. D., & Kaplan, G. B. (2007). High-dose transdermal nicotine and naltrexone: Effects on nicotine withdrawal, urges, smoking, and effects of smoking. *Experimental and Clinical Psychopharmacology, 15*(1), 81-92. doi: 10.1037/1064-1297.15.1.81

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *The Journal of Neuroscience, 29*(22), 7191-7198. doi: 1523/JNEUROSCI.0979-09.2009

Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. *Behavioural Brain Research, In Press, Uncorrected Proof*.

Shiffman, S., Shadel, W. G., Niaura, R., Khayrallah, M. A., Jorenby, D. E., Ryan, C. F., & Ferguson, C. L. (2003). Efficacy of acute administration of nicotine gum in relief of cue-provoked cigarette craving. *Psychopharmacology, 166*(4), 343-350. doi: 10.1007/s00213-002-1338-1

Sutherland, G., Russell, M. A. H., Stapleton, J., Feyerabend, C., & Ferno, O. (1992). Nasal nicotine spray: a rapid nicotine delivery system. *Psychopharmacology, 108*(4), 512-518.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behaviour: Role of automatic and nonautomatic processes. *Psychological Review, 97*(2), 147-168. doi: 10.1037/0033-295X.97.2.147

Tiffany, S. T., Cox, L. S., & Elash, C. A. (2000). Effects of transdermal nicotine patches on abstinence-induced and cue-elicited craving in cigarette smokers. *Journal of Consulting and Clinical Psychology, 68*(2), 233-240. doi: 10.1037//0022-006x.68.2.233

Townshend, J. M., & Duka, T. (2005). Binge drinking, cognitive performance and mood in a population of young social drinkers. *Alcoholism-Clinical and Experimental Research, 29*(3), 317-325.

Trick, L., Hogarth, L., & Duka, T. (2011). Prediction and uncertainty in human Pavlovian to instrumental transfer *Journal of Experimental Psychology: Learning Memory and Cognition, 37*(3), 757-765. doi: doi: 10.1037/a0022310

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience, 29*(11), 2225-2232. doi: 10.1111/j.1460-9568.2009.06796.x

Valentin, V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *The Journal of Neuroscience, 27*(15), 4019-4026.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Sciences, 15*(11), 546-554. doi: 10.1016/j.tics.2011.09.008

Waters, A. J., Shiffman, S., Sayette, M. A., Paty, J. A., Gwaltney, C. J., & Balabanis, M. H. (2004). Cue-provoked craving and nicotine replacement therapy in smoking cessation. *Journal of Consulting and Clinical Psychology, 72*(6), 1136-1143. doi: 10.1037/0022-006x.72.6.1136

West, R., Baker, C. L., Cappelleri, J. C., & Bushmakin, A. G. (2008). Effect of varenicline and bupropion SR on craving, nicotine withdrawal symptoms, and rewarding effects of smoking during a quit attempt. *Psychopharmacology, 197*(3), 371-377. doi: 10.1007/s00213-007-1041-3

Willner, P. (2000). Further validation and development of a screening instrument for the assessment of substance misuse in adolescents. *Addiction, 95*(11), 1691-1698. doi: 10.1046/j.1360-0443.2000.951116919.x

Wills, A. J., & Pothos, E. M. (2012). On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin, 138*(1), 102-125. doi: 10.1037/a0025715