

# Phasic transition from goal-directed to habitual control over drug-seeking produced by conflicting reinforcer expectancy

Lee Hogarth<sup>1</sup>, Matt Field<sup>2</sup> & Abigail K. Rose<sup>2</sup>

School of Psychology, University of New South Wales, Australia<sup>1</sup> and School of Psychology, University of Liverpool, UK<sup>2</sup>

## ABSTRACT

The transition from goal-directed to habitual control over drug-seeking has been experimentally demonstrated in animals, but there have been no comparable reports in humans. Following a recent animal design, the current study employed an outcome-devaluation procedure to test whether goal-directed control over tobacco seeking would be abolished by alcohol expectancy. Eighty smokers first learned that two responses earned tobacco or chocolate points, respectively, before tobacco was devalued by health warnings and smoking satiety. Participants were then presented with either a glass of beer/wine or water with instructions that this item could be consumed after the task (alternative reward). Then choice between the tobacco and chocolate response was measured in extinction to assess goal-directed control of tobacco seeking, in a nominal Pavlovian to instrumental transfer (PIT) test to assess stimulus control of tobacco seeking, and in a reacquisition test to assess the impact of direct feedback from the outcomes. The results showed that alcohol expectancy selectively abolished goal-directed control of tobacco seeking but not stimulus control or the impact of feedback from outcomes. These data suggest that 'endogenous' retrieval of low drug value governing goal-directed regulation of drug seeking is disrupted by conflicting appraisal of an alternative reinforcer, promoting habitual control, which may play a role in relapse.

**Keywords** Addiction, goal, habit, human, learning.

*Correspondence to:* Lee Hogarth, School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia. E-mail: l.hogarth@unsw.edu.au

## INTRODUCTION

Animal learning theorists have long expounded a dual-process model of action selection wherein reward seeking may be goal-directed (i.e. governed by an expectation of the current value of the reward) or habitual (i.e. elicited directly by contextual cues that have been reliably associated that the response in the past) (Dickinson & Balleine 2002). The key method for identifying the contribution of these two controllers to reward seeking is the outcome-devaluation protocol. Here, animals first learn that two responses earn different rewarding outcomes (e.g. pellets and sucrose), before one outcome is devalued by specific satiety of taste aversion learning. The animal is then given the opportunity to perform the two responses in extinction to determine the impact of the devaluation treatment on choice without direct feedback from the

outcomes themselves. Any reduction in responding for the devalued outcome in the extinction test must be mediated by animal's knowledge of the response–outcome contingencies combined with knowledge of the current value of the outcomes, i.e. reflects goal-directed control. By contrast, no effect of devaluation on responding suggests that the animal's behaviour is not governed by knowledge of the consequences (goal-directed) but instead is elicited directly by the test context that has become associated with the performance of the responses through stimulus–response/reinforcement or habit learning. In this case, a reacquisition test identical to initial training is typically conducted to confirm that the animal can modify responding for the devalued outcome given direct feedback from the outcomes through S–R/reinforcement learning. Thus, whereas goal-directed control is marked by a devaluation effect in

This work was carried out at the University of Liverpool, School of Psychology, Eleanor Rathbone Building, Bedford Street South, Liverpool, L69 7ZA, UK.

extinction, habitual control is marked by no devaluation effect in extinction combined with a devaluation effect in reacquisition.

Several animal studies have utilised this outcome-devaluation protocol to demonstrate the transition between goal-directed and habitual control over drug seeking. Whereas two of these designs have demonstrated goal-directed control of drug seeking (Hutcheson *et al.* 2001; Olmstead *et al.* 2001), two others have demonstrated habitual control of drug seeking (Dickinson, Wood & Smith 2002; Miles, Everitt & Dickinson 2003). Precisely what difference between these protocols promoted goal-directed action versus habit remains unclear, but the position of the response within the instrumental sequence or chain (Balleine *et al.* 1995; Daw, Niv & Dayan 2005; Dezfouli & Balleine 2012) or the amount of training (Dickinson *et al.* 1995; Killcross & Coutureau 2003) may have been relevant. Consistent with the latter claim, two further studies have demonstrated that drug seeking shifts from being goal-directed in early training to being habitual following extended training (Zapata, Minney & Shippenberg 2010; Corbit, Nie & Janak 2012). Thus, the outcome-devaluation protocol has proven validity in as an index of the differential governance of drug seeking by the goal-directed and habitual controllers at different stages of training.

More recently, phasic transitions from goal-directed to habitual control over reward seeking has been produced by various 'acute' manipulations. Specifically, using the outcome-devaluation protocol with natural rewards, abolition of goal-directed control over action selection in the extinction test has been produced by stress induction prior to instrumental training (Schwabe & Wolf 2009) or prior to the extinction test (Schwabe & Wolf 2010) in humans, by conducting the extinction test in an alcohol paired context in rats (Ostlund, Maidment & Balleine 2010) and by administration of an acute dose of alcohol prior to instrumental training in humans (Hogarth *et al.* 2012a). Importantly, in the latter two studies, the devaluation effect was corrected by direct feedback from the outcomes in a reacquisition test confirming that these acute manipulations produced a selective impairment in goal-directed control. One interpretation of these effects is that appraisal of alternative reinforcement during the extinction test (i.e. thinking about the stress procedure, alcohol expectancy or alcohol intoxication) interferes with capacity to retrieve representations of the specific instrumental outcomes and their values, which is required for goal-directed control over action selection in the extinction test. By contrast, in the reacquisition test, feedback from the outcomes can modify action selection directly through S-R/reinforcement learning, thereby correcting the devaluation effect.

The implication of the foregoing work is that appraisal of an alternative reinforcer might acutely impair goal-directed regulation of drug-seeking behaviour, promoting a transition to habitual control over this behaviour, which may play a role in relapse. The current study sought to test this prediction directly using an outcome-devaluation protocol established for human smokers (Hogarth & Chase 2011; Hogarth 2012). Eighty smokers were first trained on a concurrent choice procedure in which two responses earned tobacco or chocolate points, respectively. Tobacco was then devalued by having all participants rate health warning against smoking, for instance, 'Smoking causes fatal lung cancer', before smoking a cigarette to satiety. Then, to induce conflicting reinforcer appraisal, participants were poured a glass of beer/wine or water and were told that they could consume this item after the test that followed. As the study was conducted in a bar lab (a large room decorated and furnished to convincingly mimic a typical British pub), these drinks should evoke a compelling representation of their consumption and thus command retrieval capacity. In the first test phase that followed, choice between the tobacco and chocolate response was tested in nominal extinction, where outcomes were no longer displayed. Our hypothesis was that alcohol versus water expectancy would abolish the tobacco devaluation effect on tobacco choice in this extinction test demonstrating impaired goal-directed regulation of drug seeking.

To evaluate specificity of this impairment, a nominal Pavlovian to instrumental transfer (PIT) test was then conducted in which a picture of a cigarette, a picture of a chocolate bar or a blank picture was presented just before participants made a choice in extinction. Two analyses were derived from this PIT test. First, the overall choice of the tobacco and chocolate response (collapsed across stimulus conditions) should be equivalent to choice in the extinction test and thus replicate any impairment in goal-directed control produced by alcohol expectancy. Second, the tobacco and chocolate stimuli should selectively enhance choice of the response that earned the same outcome (Hogarth *et al.* 2007). This selective PIT effect is thought to be mediated by the stimuli retrieving a representation of their associated outcome (S-O), which in turn retrieves responses associated with that same outcome (O-R). Thus, the selective PIT effect can be used to evaluate whether alcohol expectancy impaired the retrieval of these S-O or O-R associations, or their integration (de Wit & Dickinson 2009; Hommel 2009). Our prediction was that alcohol expectancy would abolish the devaluation effect in extinction and overall choice of the PIT test, but would have no effect on the specific PIT effect. Such data would suggest that alcohol expectancy selectively impaired 'endogenous' retrieval of specific outcome values required for goal-directed control of

action selection but had no effect cue-elicited (or exogenous) outcome retrieval required for specific PIT.

Finally, a reacquisition test followed in which choice between the two responses was again measured but with the respective outcomes now being earned as in the initial concurrent training phase. This reacquisition test allows choice to be modified by direct experience of the instrumental outcomes through S–R/reinforcement learning, i.e., experience of the devalued tobacco points should serve to decrease the ability of the procedural cues to prime that response on future trials. Our prediction was that alcohol expectancy would abolish the devaluation effect in the extinction and overall PIT test, but not in the reacquisition test, corroborating the animal design of Ostlund *et al.* (2010). These data would suggest that alcohol expectancy impairs retrieval of outcome values governing goal-directed action, but does not influence S–R/reinforcement learning, or counteract the efficacy of the devaluation treatment. Finally, to confirm this latter point, subjective craving to smoke (Cox, Tiffany & Christen 2001) was measured at the beginning and end of the procedure, and it was expected that alcohol expectancy would not modify the effect of the tobacco devaluation treatment on reducing craving.

## MATERIALS AND METHOD

### Participants

Eighty smokers (half male) aged between 17 and 65 years (mean 26.5) were recruited. Participants were

randomly allocated to the alcohol and water expectancy group, balancing for gender and response–outcome assignment in the choice task within each group. The distribution of units of alcohol consumed per week (assessed by timeline follow back) was bimodal with significant skew at the high end ( $P = 0.007$ ), driven by nine outlying participants, six of whom were in the alcohol expectancy group. These nine outliers were excluded to achieve normality in the distribution of alcohol units per week ( $P = 0.48$ ) and match the alcohol and water expectancy groups with respect to alcohol use and dependence criteria shown in Table 1. Ethical approval was obtained from the University of Liverpool, School of Psychology Research Ethics Committee. Participants gave written informed consent, and participation was voluntary.

### Apparatus and materials

All computer tasks and questionnaires were completed by the participant seated on a stool at the bar of Liverpool Psychology bar lab. At the outset of the experiment, participants reported age and gender before completing the brief questionnaire of smoking urges (QSU—Cox *et al.* 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 QSU was completed at the beginning and end of the experiment to assess the effectiveness of the tobacco devaluation treatment in reducing desire for tobacco.

**Table 1** Characteristics of the alcohol and water expectancy groups.

	Expectancy group		P
	Alcohol (n = 34)	Water (n = 37)	
Years of age	25 (6.5, 18–45)	27.9 (8.6, 17–65)	0.07
Smoking days per week	6.4 (1.1, 4–7)	6.5 (1.1, 2–7)	0.87
Cigarettes on smoking days	10.0 (4.9, 2–25)	10.8 (5.7, 2–25)	0.58
Time since a cigarette (min)	35.6 (46.3, 7.2–230)	30.5 (35.7, 5–180)	0.97
Smoking years	7.6 (5.9, 1–20)	10.2 (8.8, 1–45)	0.21
Age of smoking onset	17.3 (2.3, 14–25)	18.1 (2.8, 14–28)	0.19
DSM nicotine dependence total score	4.69 (1.4, 2–7)	4.5 (1.5, 1–7)	0.94
Fagestrom nicotine dependence	3.1 (1.8, 0–7)	3.1 (1.7, 0–6)	0.76
Cigarette dependence scale	16.3 (3.5, 9–22)	14.9 (4.4, 5–22)	0.18
Mean alcohol units per week	12.1 (7.2, 3–28)	13.1 (6.8, 2.5–27)	0.63
Mean binge drinking occasions	1.2 (1, 0–3)	1.3 (1.1, 0–4)	0.95
Age of alcohol onset	18.1 (1.9, 13–22)	17.7 (1.9, 14–24)	0.32
Alcohol use disorders inventory	9.0 (3.9, 2–19)	9.5 (4.4, 2–21)	0.56
Cognitive emotional preoccupation	27.4 (9.6, 9–47)	25.9 (9.7, 9–56)	0.64
Cognitive behavioural control	16.9 (7.3, 6–31)	17.9 (8.8, 6–33)	0.65
Substance misuse	2.1 (2.5, 0–8)	1.9 (2.2, 0–7)	0.82
Barratt impulsivity total score	72.2 (10.5, 55–91)	72.8 (10.4, 46–91)	0.84
Rating of smoking health warnings	5.1 (1.0, 2.6–7.4)	5.3 (1.0, 3.6–7.7)	0.48

At the end of the experiment, a battery of validated questionnaires were given to measure smoking/drinking behaviour and impulsivity: the cigarette dependence scale (CDS-5—Etter, Le Houezec & Perneger 2003); the Fagerström questionnaire of nicotine dependence (Fagerström 1978); the Diagnostic and Statistical Manual of Mental disorders tobacco dependence questionnaire (Grant *et al.* 2003); the substance misuse scale (Willner 2000); the Barratt impulsivity scale (BIS-11 Stanford *et al.* 2009); the World Health Organisation alcohol use disorders inventory (AUDIT—Babor *et al.* 2001); the alcohol use timeline follow-back questionnaire (Sobell & Sobell 1992), which estimates units of alcohol and binge consumption in the last 2 weeks; and age of drinking onset. Experimental generation was with E-prime software (Psychology Software Tools, Inc.) on a standard laptop.

## Procedure

### *Concurrent choice training*

The purpose of concurrent choice training was to establish two instrumental responses that earned tobacco and chocolate reward, respectively. Participants had access to the laptop placed on the bar of the bar lab. On-screen instructions stated, 'This is a game in which you can earn cigarettes and chocolate. In each trial, press the D or H key to try and win these rewards. You will only win on some trials. Press the space bar to begin'.

Each trial began with the centrally presented text, 'Choose a key', which remained until either the D or H key was pressed (i.e. a two-key forced-choice task). A response on one key replaced this text with the outcome, 'You win  $\frac{1}{4}$  of a cigarette', whereas a response on the other key produced the outcome, 'You win  $\frac{1}{4}$  of chocolate bar'. Only one 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 key for the unscheduled outcome was pressed), the text 'You win nothing' was presented. These three outcome texts were presented for 1500 milliseconds, followed by a random inter-trial interval (ITI) between 1000 and 2000 milliseconds 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. The percent choice of the tobacco versus the chocolate key was recorded as the dependent measure.

### *Devaluation treatment*

Following concurrent choice training, the following instructions were presented: In this part of the task, we would like to assess how unpleasant you find statements

concerning the adverse consequences of smoking. Please read each statement carefully. Then report how unpleasant you find each statement by pressing a number key between 1 and 9, where: 1 = Not at all unpleasant, 5 = mildly unpleasant and 9 = extremely unpleasant. Press the space bar to begin'. Each statement was presented for 5 s before the question was presented underneath: 'How unpleasant do you find this statement? Choose a number between 1–9, where: 1 = Not at all unpleasant, 5 = mildly unpleasant and 9 = extremely unpleasant'. Responding launched a random ITI between 1000 and 2000 milliseconds prior to the next statement. There were 16 statements, e.g. 'Smoking causes fatal lung cancer' selected at random (for a full list see appendix 1 of Hogarth & Chase 2011). Following this, participants were taken outside and allowed a fixed 10 minute period in which to smoke as much or as little as they wished. These health warning and smoking satiety procedures have been separately demonstrated to be effective devaluation treatments in previous studies (Hogarth & Chase 2011) and were combined in the current design with the aim of producing a decisive devaluation effect.

### *Expectancy manipulation*

Upon returning to the bar lab, participants were again seated at the bar in front of the laptop. The alcohol expectancy group were presented with a cold bottle of Becks lager (275 ml) and a bottle of Villa Radiosa wine (750 ml) and asked to choose which drink they would prefer. Their chosen drink was then opened and poured into a glass (either a 284 ml half pint glass or 175 ml wine glass) and placed next to the participant along with the verbal instructions that they could drink this after the computer task was finished. As participants were seated in a large bar lab (~160 square feet) furnished to look like a comfortable British pub, these drinks were likely to evoke a powerful expectation of their consumption. By contrast, participants in the water expectancy group saw the experimenter pour a standard glass (284 ml) of bottled water, which was placed by the participant with the instructions that they could drink this after the computer task was finished. The aim of this design was to generate effective alcohol expectancy arguably akin to rats in Ostlund *et al.* (2010).

### *Extinction test*

The extinction test following devaluation comprised a single block of 24 trials identical to concurrent training apart from the outcomes being omitted. The purpose of this test was to measure the impact of the tobacco devaluation treatment on choice between the two keys in the absence of feedback from the outcomes. Participants were presented with the on-screen instructions: 'In this

part of the task, you can earn cigarettes and chocolate by pressing the D or H keys in the same way as during the first part of the experiment. However, you will only be told how many of each reward you have earned at the end of the experiment. Press the space bar to begin'. Each trial began with the prompt, 'Choose a key', whereupon the participant pressing the D or H key launched a random ITI between 1000 and 2000 milliseconds prior to the next trial. Totalizer screens were also omitted between blocks to avoid feedback from earned outcomes. The question at stake was whether tobacco devaluation would reduce tobacco choice relative to concurrent training indicative of goal-directed control, and the differences between expectancy groups in this effect.

#### PIT test

The PIT test that followed extinction examined the impact of tobacco and chocolate cues on choice between the two responses. On-screen instructions stated: 'In this part of the task, you can earn cigarettes and chocolate by pressing the D or H keys in the same way as during the first part of the experiment. However, you will sometimes be shown pictures before you choose which key to press. Press the space bar to begin'. Each trial began with the prompt 'Choose a key', which was compounded with either a cigarette picture (two cigarettes on a white background) or chocolate picture (a single 49g Cadbury Dairy Milk Chocolate bar on a white background) presented directly above the prompt, intermixed with trials containing no stimulus (these pictures are shown in fig. 1 of Hogarth 2012; Hogarth & Chase 2011). Pressing the D or H key launched a random ITI between 1000 and 2000 milliseconds prior to the next trial. Thus, the trial structure was identical to extinction trials, apart from the stimulus presented with the prompt. The PIT test totalled 48 trials, comprising four cycles of 12 trials where each cycle presented the cigarette, chocolate and no stimulus four times each, in random order. The totalizer screen between blocks was again omitted to avoid feedback from outcomes.

The PIT data were analysed to address three points. First, if tobacco devaluation reduced tobacco choice in the PIT test overall (collapsed across cues) relative to concurrent training, this would replicate the devaluation effect seen in the extinction test relative to concurrent training. Second and most importantly, if alcohol expectancy attenuated the devaluation effect in both the extinction test and the overall PIT data, this would provide a within-experiment replication of the key hypothesis of the study, greatly substantiating our claim to have demonstrated a phasic impairment in goal-directed control. Finally, PIT data were also examined to determine whether stimuli enhanced responding for the same outcome, and whether expectancy groups differed in

such stimulus control of action selection to address the specificity of the alcohol expectancy effect.

#### Reacquisition test

The reacquisition test was identical to concurrent training and matched extinction trials except that outcomes—chocolate and tobacco points—were again delivered upon their respective response. The reacquisition test comprised 32 trials broken into two blocks of 16 trials, wherein the tobacco and chocolate outcome available in each trial was selected at random. The 'totalizer' screen following each block again reported the quantity of each reward type earned. The question at stake was whether experience of the outcomes contingent upon the response would engage a tobacco devaluation effect on choice, which differed from the extinction test.

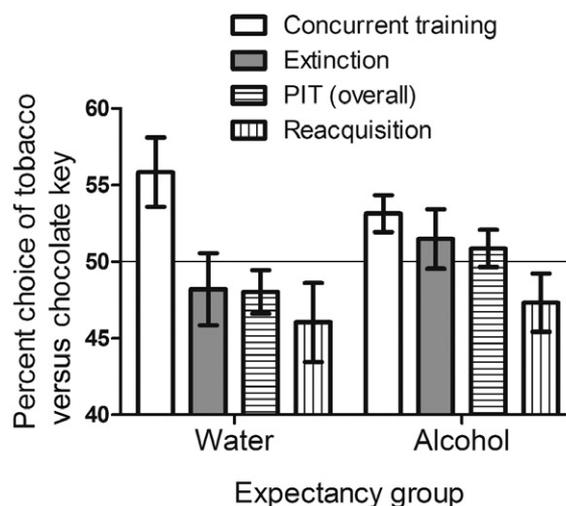
## RESULTS

### Participants

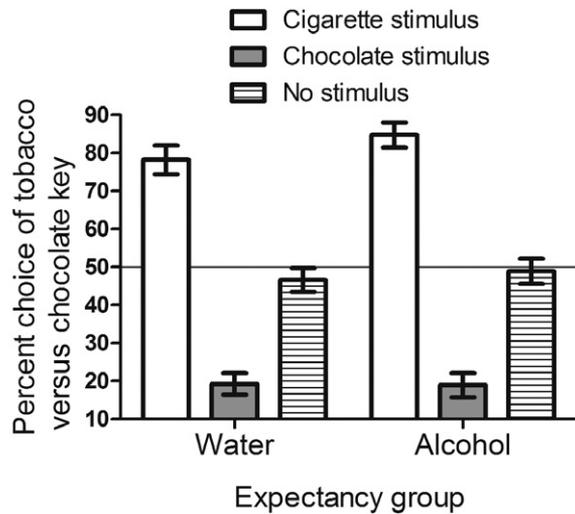
Table 1 shows that the characteristics of the alcohol and water expectancy group were matched (although there was a trend towards a difference in age). The two groups were also matched for their ratings of the unpleasantness of the smoking health warnings during the devaluation treatment.

### Choice procedures

Figure 1 shows the percent choice of the tobacco versus chocolate key during concurrent training, extinction test, PIT test (collapsed across cue conditions) and reacquisi-



**Figure 1** Percent choice of the tobacco versus chocolate response in concurrent training, extinction test, PIT test (collapsed across the cigarette, chocolate and no stimulus condition) and reacquisition test, for the alcohol and water expectancy group



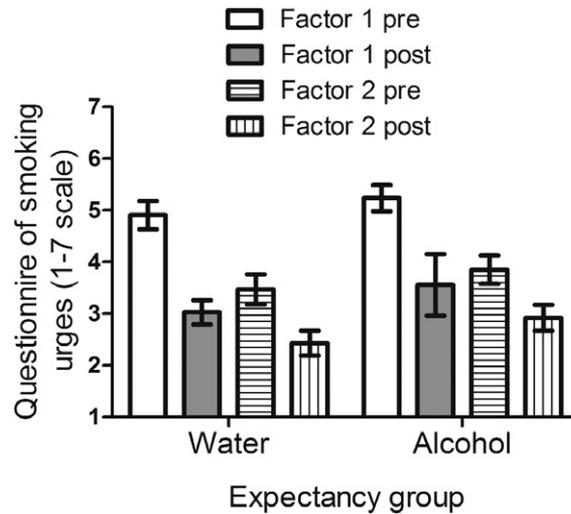
**Figure 2** Percent choice of the tobacco versus chocolate response in the cigarette, chocolate and no stimulus condition of the PIT test, for the alcohol and water expectancy group

tion test, for the alcohol and water expectancy group. Consistent with our hypothesis, alcohol expectancy abolished the devaluation effect on tobacco choice in the extinction test and in the PIT test overall, but not in the reacquisition test.

This description was substantiated by ANOVA on the concurrent and extinction data, which yielded a significant interaction between group (alcohol, water) and block (concurrent, extinction),  $F(1,69) = 4.54$ ,  $P < 0.05$ , where the block effect was significant in the water group,  $F(1,36) = 14.58$ ,  $P = 0.001$ , but not the alcohol group,  $F < 1$ . Similarly, ANOVA incorporating overall choice in the PIT test and concurrent training phase yielded a significant interaction between group and block,  $F(1,69) = 4.68$ ,  $P < 0.05$ , where the main effect of block was reliable in the water group,  $F(1,36) = 14.58$ ,  $P = 0.001$ , and marginal in the alcohol group,  $F(1,33) = 3.60$ ,  $P = 0.066$ . Thus, the alcohol expectancy group's impairment in goal-directed control over choice in the extinction test was maintained in overall choice in the PIT test that followed, replicating the impairment.

By contrast, when concurrent and reacquisition data were entered into ANOVA, there was no reliable group by block interaction,  $F < 1$ , but instead, the block effect was reliable overall,  $F(1,69) = 10.95$ ,  $P = 0.001$ , and in both the water,  $F(1,36) = 6.03$ ,  $P < 0.05$ , and alcohol group,  $F(1,33) = 6.33$ ,  $P < 0.05$ , in isolation. Thus, whereas alcohol expectancy impaired goal-directed control over drug seeking in the extinction test, both groups decreased tobacco choice following experience of the outcomes in the reacquisition test, corroborating Ostlund *et al.* (2010).

Figure 2 shows that alcohol expectancy had no effect on the ability of cues to drive choice of the same outcome



**Figure 3** Questionnaire of smoking urges factor 1 and factor 2 recorded at the beginning and end of the experiment (pre- and post-tobacco devaluation treatment), for the alcohol and water expectancy group

in the PIT test. ANOVA on these data incorporating the variables group and stimulus (3), yielded a main effect of stimulus,  $F(2,138) = 145.87$ ,  $P < 0.001$ ; no main effect of group,  $F(1,69) = 2.23$ ,  $P = 0.14$ ; and no group by stimulus interaction,  $F < 1$ . Finally, the extent to which the tobacco and chocolate stimuli elicited responding for the same outcome relative to the no stimulus condition was statistically equivalent,  $F(1,70) = 1.14$ ,  $P = 0.29$ , indicating that these cues were equally effective at priming choice of the same outcome.

### Craving

Figure 3 shows factor 1 and 2 craving scores obtained at the beginning and end of the experiment. Craving declined from pre to post devaluation, in both factors, and the magnitude of these declines was equivalent between the alcohol and water expectancy groups. This impression was confirmed by ANOVA on Fig. 3, which yielded an effect of time,  $F(1,69) = 20.62$ ,  $P < 0.001$ ; factor,  $F(1,70) = 57.45$ ,  $P < 0.001$ ; and time by factor interaction,  $F(1,69) = 11.17$ ,  $P = 0.008$ , but no other reliable effects or interactions,  $F_s < 1.92$ ,  $P_s > 0.17$ . Thus, the two groups were matched in the tobacco devaluation effect on subjective craving, and factor 1 craving scores were more sensitive to this devaluation effect than factor 2 scores.

### DISCUSSION

The current study is the first to demonstrate a transition from goal-directed to habitual control over drug seeking in humans, using analogous methods to those employed

with animals. Overall, the study suggests that appraisal of an alternative reinforcer (alcohol) selectively impaired capacity for 'endogenous' retrieval of drug (tobacco) value required for goal-directed control over responding for tobacco, rendering this behaviour prone to habitual control by contextual cues. Let us consider each effect in turn to specify this claim. The primary finding was that alcohol expectancy compared to water expectancy abolished the impact of tobacco devaluation on reducing tobacco choice in the extinction test and overall in the PIT test, relative to concurrent training, demonstrating the stability of the impairment in goal-directed control across these two test phases. The remaining results address the specificity of this impairment. First, alcohol expectancy did not modify the impact of tobacco devaluation on reducing tobacco choice in the reacquisition test, suggesting that capacity for S–R/reinforcement learning was intact, that is, direct experience of the devalued tobacco outcome was able to reduce the propensity of procedural cues to prime this choice on future trials. Second, the finding that alcohol expectancy did not modify the impact of tobacco devaluation on reducing tobacco choice in the reacquisition test or on reducing subjective craving indicates that alcohol expectancy did not counteract the efficacy of tobacco devaluation treatment, which might be expected given cross-priming effects between alcohol and tobacco (Mintz *et al.* 1985; Burton & Tiffany 1997). Third, alcohol expectancy did not modify the extent to which tobacco and chocolate stimuli enhanced choice of the same outcome in the PIT test, indicating that stimulus-induced retrieval of S–O or O–R associations, or their integration, were not impaired by alcohol expectancy. Finally, alcohol expectancy did not modify reacquisition and selective PIT performance indicating that alcohol expectancy did not produce a general disengagement from the task or loss of response selectivity.

The current findings bear a striking resemblance to three other designs, which together support broader claims regarding the impairment. First, Ostlund *et al.* (2010) found that conducting the test phase in an alcohol paired context abolished the devaluation effect on natural reward seeking in extinction but not reacquisition, demonstrating a comparable impairment in goal-directed control across species under similar conditions. Second, Hogarth *et al.* (2012a) found that administering alcohol prior to training abolished the devaluation effect on natural reward seeking in extinction but not reacquisition, suggesting a comparability between alcohol intoxication and alcohol expectancy in impairing goal-directed control. Finally, Schwabe and colleagues found that stress induction either prior to training (Schwabe & Wolf 2009) or test (Schwabe & Wolf 2010) abolished the devaluation effect on natural reward seeking in the extinction test, suggesting a comparability between alcohol expectancy,

alcohol intoxication and stress in impairing goal-directed control. Our claim is that all three manipulations, alcohol expectancy, alcohol intoxication and stress induction, abolished goal-directed control because they induced a strong appraisal that limited capacity to retrieve specific instrumental outcome values required for goal-directed control. In the present study, the alcohol drink may have engaged a stronger appraisal than water because of the richer and potentially ambivalent consequence of intoxication, and/or because this drink was compounded by the evocative environment of the bar lab.

The idea that appraisal of alternative reinforcers competes with outcome retrieval underpinning goal-directed action garners support from other domains. First, many theories of value based decision making propose that outcome values are represented sequentially before the relatively highest value response is selected, suggesting limited capacity in value based representational space (Vlaev *et al.* 2011). Second, it has been shown that when rats and humans are faced with a learning task in which the same event must be encoded as both a stimulus and an outcome, they favour a solution based upon S–R over goal-directed learning, suggesting a delegation to S–R under heavy cognitive demands or conflict (de Wit *et al.* 2007). Third, a number of studies have shown that tasks in which participants must choose between incommensurable rewards (e.g. food versus shoes: FitzGerald, Seymour & Dolan 2009; Guitart-Masip, Talmi & Dolan 2010) activate similar frontal cortical regions that are activated during goal-directed action selection (Sugrue, Corrado & Newsome 2005; Valentin, Dickinson & O'Doherty 2007; de Wit *et al.* 2009; Rangel & Hare 2010; Balleine, Leung & Ostlund 2011), suggesting that value-based representations in these two instances occupy common neural resources and thus may interfere.

Although the discussion thus far as focused on phasic impairments in goal-directed control at test, substantial evidence has shown that a seemingly equivalent impairment can arise from chronic psychiatric or spectrum traits (Klossek, Russell & Dickinson 2008; de Wit *et al.* 2011; Gillan *et al.* 2011; Hogarth, Chase & Baess 2012b), from chronic drug exposure prior to training (Dickinson *et al.* 2002; Miles *et al.* 2003; Nelson & Killcross 2006; Zapata *et al.* 2010; Corbit *et al.* 2012), brain lesions prior to training (Corbit & Balleine 2003; Killcross & Coutureau 2003; Yin *et al.* 2005) and from procedural variables that operate exclusively during training (Balleine *et al.* 1995; Dickinson *et al.* 1995; Kosaki & Dickinson 2010). One reconciliation of the collected action–habit literature is to assume that variables which either reduce the perception of the response–outcome contingency during training (Tanaka, Balleine & O'Doherty 2008; Kosaki & Dickinson 2010) or reduce capacity to retrieve specific outcome values at test, converge in producing a

common impairment in goal-directed action favouring S–R control. Indeed, such variables may be additive in producing the impairment, but this remains to be tested.

Examination of why the devaluation effect in reacquisition was robust against conflicting alcohol expectancy qualifies the nature of the impairment in extinction. Four explanations of the impact of reacquisition are possible. The outcomes may have (1) reminded participants of the outcome values enabling goal-directed control, or (2) modified the propensity of contextual cues to elicit the two responses through S–R/reinforcement learning. Alternatively, the time elapsing between the extinction and reacquisition test may have (3) enabled the alcohol expectancy group to catch up in their goal-directed control, or (4) allowed the alcohol expectancy to decay sufficiently for goal-directed control to be reasserted. These interpretations may be constrained by following observations. First, if the outcomes acted as a reminder (explanation 1), the cigarette and chocolate stimuli in the PIT test might be expected have similarly normalised the devaluation effect. Indeed, the alcohol expectancy group did show a trend towards a devaluation effect in overall choice of the PIT test, but this was significantly smaller than the water expectancy group, suggesting a weak cue reminder effect that cannot readily explain the full normalisation of the devaluation effect in the reacquisition test. Indeed, this claim that cue reminder plays little part in correction achieved by the reacquisition test is supported by Klossek *et al.* (2008; experiment 3), where a cue reminder test similarly failed to normalise a devaluation effect in contrast to a reacquisition test.

The possibility that time enabled goal-directed control to catch up (explanation 3) or the alcohol expectancy to decay (explanation 4) may be addressed by pointing out that the alcohol expectancy group showed a sudden correction of their devaluation effect in the reacquisition test, whereas an explanation based upon time might predict a more linear correction across the three tests (extinction, PIT, reacquisition). Correspondingly, Ostlund *et al.* (2010) found that impaired goal-directed control in an alcohol paired context did not show a linear correction across fine-grained time bins of testing but rather showed a sudden correction upon institution of the reacquisition conditions. In addition, participants anticipated drinking at the end of the computer task; therefore, if anything, alcohol expectancy should have increased over time. This analysis leaves only explanation 2 intact, i.e. that response contingent outcomes normalised the alcohol expectancy group's devaluation effect by modifying the propensity to make each response directly through S–R/reinforcement habit learning.

Finally, we must consider why alcohol expectancy abolished the devaluation effect in extinction but left the selective PIT effect intact. It is a convenient heuristic to

interpret this difference as suggesting that alcohol expectancy produced a selective impairment in endogenous but not exogenous outcome retrieval. However, this position must be qualified. According to the current thinking, goal-directed action is initiated by contextual stimuli, which retrieve a representation of available response options that have previously been reinforced in that context, and these response representations in turn retrieve their respective outcome values, which weigh that response for motor performance accordingly (de Wit & Dickinson 2009). By contrast, the selective PIT effect is thought to be mediated by stimuli retrieving a representation of their associated outcome, which in turn elicits the response that was associated with that outcome. Moreover, as the PIT effect is not modified by outcome devaluation (Colwill & Rescorla 1990; Rescorla 1994; Holland 2004; Corbit, Janak & Balleine 2007; Hogarth, Dickinson & Duka 2010; Hogarth & Chase 2011; Hogarth 2012), stimuli are thought to retrieve the perceptual identity but not value of the outcome. Indeed, the insensitivity of specific PIT to devaluation can be observed in the current data. Despite the two expectancy groups showing a differential devaluation effect in overall choice of the PIT test, they showed an equivalent impact of the tobacco stimulus in eliciting tobacco choice, demonstrating that this cueing effect was unaffected by devaluation. Taking these arguments together, therefore, it is probably more accurate to claim that goal-directed control but not selective PIT was abolished by alcohol expectancy because the outcome representation underpinning goal-directed control encodes current value and is more weakly (non-differentially) associated with external stimuli at the choice point, whereas the outcome representation underpinning specific PIT does not encode current value and is strongly (differentially) associated with external stimuli at the choice point making this form of control more robust against alternative reinforcer appraisal. It remains to be explored whether value encoding or the strength of the cue–outcome contingency was critical for the differential sensitivity of these two tests to alcohol expectancy.

To conclude, alcohol expectancy selectively abolished goal-directed control of tobacco seeking in the extinction test, but did not modify stimulus control of tobacco seeking in the PIT test or the impact of direct feedback from outcomes on tobacco seeking in the reacquisition test, or the devaluation effect on subjective craving. We have favoured an interpretation wherein strong appraisal of alternative reinforcers occupies limited resources required for 'endogenous' retrieval of drug values underpinning goal-directed regulation of drug seeking, thus rendering drug seeking prone to habitual control. By contrast, strong appraisal of alternative reinforcers leaves behavioural control driven by external stimuli intact,

specifically, leaving cued outcome retrieval underpinning stimulus control of drug seeking and S–R/reinforcement learning underpinning habitual control of drug seeking. The key message is that goal-directed and habitual control over drug-seeking coexist even in relatively undertrained cohorts/responses like the present, and that transitions to habit can occur phasically driven by competing cognitive demands, in addition to advancing with practice or trait vulnerability as previously emphasised, which may play a role in relapse.

### Acknowledgements

This work was supported by the Economic and Social Research Council (RES-000-22-4365: AR/LH) and the Medical Research Council (G0701456: LH). The authors would like to thank Ian Polanowski and John Dee for assistance with data collection.

### Authors Contribution

All authors made a substantial contribution to the design, implementation and reportage of the study.

### References

- Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG (2001) AUDIT: The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care. Dependence DoMHaS, ed. London: World Health Organization.
- Balleine BW, Garner C, Gonzalez F, Dickinson A (1995) Motivational control of heterogeneous instrumental chains. *J Exp Psychol Anim Behav Process* 21:203–217.
- Balleine BW, Leung BK, Ostlund SB (2011) The orbitofrontal cortex, predicted value, and choice. *Ann N Y Acad Sci* 1239:43–50.
- Burton SM, Tiffany ST (1997) The effect of alcohol consumption on craving to smoke. *Addiction* 92:15–26.
- Cappelleri JC, Bushmakin AG, Baker CL, Merikle E, Olufade AO, Gilbert DG (2007) Multivariate framework of the brief questionnaire of smoking urges. *Drug Alcohol Depend* 90:234–242.
- Colwill RM, Rescorla RA (1990) Effects of reinforcer devaluation on discriminative control of instrumental behavior. *J Exp Psychol Anim Behav Process* 16:40–47.
- Corbit LH, Balleine BW (2003) The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* 146:145–157.
- Corbit LH, Janak PH, Balleine BW (2007) General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *Eur J Neurosci* 26:3141–3149.
- Corbit LH, Nie H, Janak PH (2012) Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. *Biol Psychiatry* 72:389–395.
- Cox LS, Tiffany ST, Christen AG (2001) Evaluation of the brief questionnaire of smoking urges (QSU-brief) in the laboratory and clinical settings. *Nicotine Tob Res* 3:7–16.
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711.
- de Wit S, Barker RA, Dickinson AD, Cools R (2011) Habitual versus goal-directed action control in Parkinson disease. *J Cogn Neurosci* 23:1218–1229.
- de Wit S, Corlett PR, Aitken MR, Dickinson A, Fletcher PC (2009) Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J Neurosci* 29:11330–11338.
- de Wit S, Dickinson A (2009) Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychol Res* 73:463–476.
- de Wit S, Niry D, Wariyar R, Aitken MRF, Dickinson A (2007) Stimulus-outcome interactions during instrumental discrimination learning by rats and humans. *J Exp Psychol Anim Behav Process* 33:1–11.
- Dezfouli A, Balleine BW (2012) Habits, action sequences and reinforcement learning. *Eur J Neurosci* 35:1036–1051.
- Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes RA (1995) Motivational control after extended instrumental training. *Anim Learn Behav* 23:197–206.
- Dickinson A, Balleine BW (2002) The role of learning in the operation of motivational systems. In: Gallistel CR, ed. *Stevens' Handbook of Experimental Psychology, Vol 3, Learning, Motivation and Emotion*, 3rd edn, pp. 497–533. New York: Wiley.
- Dickinson A, Wood N, Smith JW (2002) Alcohol seeking by rats: action or habit? *Q J Exp Psychol B* 55:331–348.
- Etter JF, Le Houezec J, Perneger TV (2003) A self-administered questionnaire to measure dependence on cigarettes: the Cigarette Dependence Scale. *Neuropsychopharmacology* 28:359–370.
- Fagerström K (1978) Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addict Behav* 3:235–241.
- FitzGerald THB, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci* 29:8388–8395.
- Gillan CM, Pappmeyer M, Morein-Zamir S, Sahakian BJ, Fineberg NA, Robbins TW, de Wit S (2011) Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am J Psychiatry* 168:718–726.
- Grant BF, Dawson DA, Stinson FS, Chou PS, 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 Alcohol Depend* 71:7–16.
- Guitart-Masip M, Talmi D, Dolan R (2010) Conditioned associations and economic decision biases. *NeuroImage* 53:206–214.
- Hogarth L (2012) Goal-directed and transfer-cue-elicited drug-seeking are dissociated by pharmacotherapy: evidence for independent additive controllers. *J Exp Psychol Anim Behav Process* 38:266–278.
- Hogarth L, Attwood AS, Bate HA, Munafò MR (2012a) Acute alcohol impairs human goal-directed action. *Biol Psychol* 90:154–160.
- Hogarth L, Chase HW (2011) Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. *J Exp Psychol Anim Behav Process* 37:261–276.
- Hogarth L, Chase HW, Baess K (2012b) Impaired goal-directed behavioural control in human impulsivity. *Q J Exp Psychol* 65:305–316.

- Hogarth L, Dickinson A, Duka T (2010) The associative basis of cue elicited drug taking in humans. *Psychopharmacology* 208:337–351.
- Hogarth L, Dickinson A, Wright A, Kouvaraki M, Duka T (2007) The role of drug expectancy in the control of human drug seeking. *J Exp Psychol Anim Behav Process* 33:484–496.
- Holland PC (2004) Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process* 30:258–258.
- Hommel B (2009) Action control according to TEC (theory of event coding). *Psychol Res* 73:512–526.
- Hutcheson DM, Everitt BJ, Robbins TW, Dickinson A (2001) The role of withdrawal in heroin addiction: enhances reward or promotes avoidance? *Nat Neurosci* 4:943–947.
- Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13:400–408.
- Klossek UMH, Russell J, Dickinson A (2008) The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *J Exp Psychol Gen* 137:39–51.
- Kosaki Y, Dickinson A (2010) Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *J Exp Psychol Anim Behav Process* 36:334–342.
- Miles FJ, Everitt BJ, Dickinson A (2003) Oral cocaine seeking by rats: action or habit? *Behav Neurosci* 117:927–938.
- Mintz J, Boyd G, Rose JE, Charuvastra VC, Jarvik ME (1985) Alcohol increases cigarette smoking: a laboratory demonstration. *Addict Behav* 10:203–207.
- Nelson A, Killcross S (2006) Amphetamine exposure enhances habit formation. *J Neurosci* 26:3805–3812.
- Olmstead MC, Lafond MV, Everitt BJ, Dickinson A (2001) Cocaine seeking by rats is a goal-directed action. *Behav Neurosci* 115:394–402.
- Ostlund SB, Maidment NT, Balleine BW (2010) Alcohol-paired contextual cues produce an immediate and selective loss of goal-directed action in rats. *Front Integr Neurosci* 4:19. doi:10.3389/fnint.2010.00019.
- Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20:262–270.
- Rescorla RA (1994) Transfer of instrumental control mediated by a devalued outcome. *Anim Learn Behav* 22:27–33.
- Schwabe L, Wolf OT (2009) Stress prompts habit behavior in humans. *J Neurosci* 29:7191–7198.
- Schwabe L, Wolf OT (2010) Socially evaluated cold pressor stress after instrumental learning favors over goal-directed action. *Psychoneuroendocrinology* 35:977–986.
- Sobell LC, Sobell MB (1992) Timeline followback: a technique for assessing self-reported alcohol consumption. In: Litten RZ, Allen JP, eds. *Measuring Alcohol Consumption: Psychosocial and Biological Methods*, pp. 41–72. Totowa, NJ: Humana Press.
- Stanford MS, Mathias CW, Dougherty DM, Lake SL, Anderson NE, Patton JH (2009) Fifty years of the Barratt Impulsiveness Scale: an update and review. *Pers Individ Differ* 47:385–395.
- Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363–375.
- Tanaka SC, Balleine BW, O'Doherty JP (2008) Calculating consequences: brain systems that encode the causal effects of actions. *J Neurosci* 28:6750–6755.
- Valentin V, Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human brain. *J Neurosci* 27:4019–4026.
- Vlaev I, Chater N, Stewart N, Brown GDA (2011) Does the brain calculate value? *Trends Cogn Sci* 15:546–554.
- Willner P (2000) Further validation and development of a screening instrument for the assessment of substance misuse in adolescents. *Addiction* 95:1691–1698.
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005) The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* 22:513–523.
- Zapata A, Minney VL, Shippenberg TS (2010) Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *J Neurosci* 30:15457–15463.