Supplementary Materials for

Carrots and sticks fail to change behavior in cocaine addiction

Karen D. Ersche,* Claire M. Gillan, P. Simon Jones, Guy B. Williams, Laetitia H. E. Ward, Maartje Luijten, Sanne de Wit, Barbara J. Sahakian, Edward T. Bullmore, Trevor W. Robbins

*Corresponding author. E-mail: ke220@cam.ac.uk

Published 17 June 2016, *Science* **352**, 1468 (2016)
DOI: 10.1126/science.aaf3700

This PDF file includes:

Materials and Methods

1. Study Sample

All volunteers consented in writing before undergoing medical review and a psychiatric screening procedure using the Mini International Neuropsychiatric Inventory (MINI). Psychopathology in drug users was further evaluated using the Structured Clinical Interview for DSM-IV. A history of traumatic head injury and/or insufficient proficiency in English were applied as exclusion criteria. All participants completed the National Adult Reading Test (*NART, 25*) to provide an estimate of verbal IQ. In view of preclinical work suggesting that stress increases proneness to habit formation (*6*), we also assessed the accumulated number of stressful life events participants have experienced using the Social Readjustment Rating Scale (*SRRS, 26*). All participants further completed the Barratt Impulsiveness Scale (*BIS-11, 27*) and the Obsessive-Compulsive Inventory (*OCI-R, 28*) to assess both impulsive and compulsive personality traits, which in animal models of addiction have been linked with the transition of habitual to compulsive responding (*14*). Urine samples were tested for undeclared drugs. All samples provided by CUD tested positive for cocaine, and urine samples provided by healthy volunteers were all negative, with the exception of one who consumed multigrain bread at breakfast leading to a positive urine screen for opiates. Breath tests were also used to screen for acute alcohol intoxication. The study was approved by the National Research Ethics Committee (12/EE/0519; PI: KDE).

CUD were recruited from local drug treatment centers and by word-of-mouth. They had been actively using cocaine/crack-cocaine for an average of 16 years (±6.7 standard deviation [SD]). In addition to meeting the DSM-IV-TR criteria for cocaine dependence, 48 also met DSM-IV-TR criteria for opiate dependence,

25 for cannabis dependence and five for alcohol dependence. Twenty-six CUD were prescribed methadone (mean dose 48.7ml, SD ± 18.0) and 14 were prescribed bruprenorphine (mean dose 7.2ml, SD ± 4.8). Although significantly more CUD (94%) reported smoking tobacco compared with control volunteers (11%) (Fisher's $p < 0.001$), nicotine dependence was not assessed using the DSM-IV- TR criteria. Only a small minority of CUD received additional medication on prescription, including narcotic-like pain relief (n = 3), benzodiazepines (n = 2), antidepressants (n = 3), baclofen for the treatment of alcohol dependence (n = 1), omeprazole for heartburn (n = 1), warfarin for the prevention of thrombosis (n = 1), metformin for the treatment of diabetes (n = 1), and steroids for the treatment of asthma (n = 1).

Healthy volunteers were recruited partly by advertisement within the local community and partly from the Cambridge BioResource volunteer panel (www.cambridgebioresource.org.uk). None of the healthy control volunteers ever fulfilled the DSM-IV-TR criteria of substance dependence. Accordingly, their scores on the Drug Abuse Screening Test (DAST-20) and the Alcohol Use Identification Test (AUDIT) were low (see Tab.S1). Only one control volunteer used prescribed medication on a regular basis, i.e. mebeverine for the treatment of irritable bowel syndrome.

2. Behavioral Measures

All participants were trained and tested on two instrumental tasks, in the context of reward and avoidance learning respectively. Both tasks, which are described in detail elsewhere (*23, 24*), are based on outcome devaluation paradigms, in which the value of the outcomes that originally motivated behavior are subsequently reduced to test for habits. Nevertheless there are some difference between the tasks as the appetitive task is based on secondary conditioned reinforcement (symbols associated with money) whereas the avoidance task is based on discomfort produced by a primary aversive stimulus, an electrical shock.

Appetitive Learning

The instrumental learning task involved four stages and, in the present study, utilized images of animals as the training stimuli and outcome instead of pictures of fruit used in previous work (*18, 29*). During the discrimination training phase, participants had to learn from feedback over the course of eight blocks (96 trials in total) which one of two buttons they needed to press in response to a stimulus shown on the screen in order to earn an outcome that was worth points (Fig.1A). The points counted towards a monetary reward, which was provided to the highest scorers as a supermarket voucher. In the second stage of the task, participants were tested behaviorally on their outcome-action knowledge. Here, two of the outcomes from the discrimination training phase were presented simultaneously on the screen, where one had a red X superimposed and the other did not. Participants were instructed to make the response that had previously led to the outcome that did not have a red X superimposed (Fig.1B). In the third stage of the task, habits were tested using the slip-of-action test (Fig.1C). Participants were first shown all of the outcomes from training, two of which had a red X superimposed, indicating that those outcomes were no longer worth any points. Participants were next presented with the stimuli from the discrimination phase in rapid succession and were instructed to respond to stimuli that were associated with valued outcomes and to withhold from responding to stimuli associated with devalued outcomes. In the fourth stage, a control test was introduced (Fig.1D), which was identically structured as the slip-of-action test, except that the discriminative stimuli were devalued, rather than the outcomes. Participants were instructed to solely respond to those stimuli that that were still 'valued' and to withhold responding to stimuli that had been devalued (with the red X). Finally, participants were asked by questionnaire about their explicit knowledge of the task with respect to relationships between stimuli, outcomes and responses. The key measure of this task is the balance between participants' goal-directed and habitual responses during the slip-of-action test, which is reflected by the differences between participants' responses to rewarded (valued) and

devalued outcomes. Accuracy and latency data were also recorded at all task stages.

Avoidance Learning

The shock avoidance task also consisted of four-stages in which participants had to learn to make a response on exposure to a visual stimulus in order to avoid receiving an electrical shock. The intensity of the shocks delivered was determined by participants' levels of tolerance for unpleasantness, as measured during a prior work-up procedure. In the first part of the task, a visual stimulus appeared on the computer screen for a duration of 750 milliseconds (ms), and the presentation was paired with an electrical shock at either the left or the right wrist (Fig.2A). The inter-trial-intervals were eight seconds, and the interval between stimulus termination and shock delivery varied randomly between 350-600 ms. The presentation order of the conditioned stimuli was randomized. Participants were subsequently introduced to an operant response that would allow them to avoid receiving an electrical shock at the wrist by pressing a foot pedal during the 750 ms presentation time of the conditioned stimulus (CS) (Fig.2B). The electrical shock (outcome) and pedal-press (response) were always congruent for the right or left hand side, i.e. a left pedal press would cancel an imminent shock to the left wrist. The avoidance contingency was set to a variable ratio schedule, such that the number of presses needed to avoid a shock varied randomly between one and two on a trial-by-trial basis. Participants were told they may need to respond more than once to effectively avoid the shock and they were permitted to press as many times as they liked within the 750 ms time window. Failure to respond the required number of times within the 750 ms time window, or pressing the incorrect pedal, resulted in the participant receiving an electrical shock.

There were two training phases – a short one (12 trials) and a one long one (120 trials) – and after each training phase, participants' sensitivity to devaluation was assessed as follows: the shock predicted by one CS was devalued by disconnecting the current stimulator from the electrodes on the participants' corresponding wrist (Fig.2C). The opposite wrist was devalued in the second devaluation test and this order was counterbalanced across participants. The shock associated with the other CS, however, was not devalued as the electrode on the participant' other wrist remained connected and active. After the devaluation procedure, participants' responses were then tested in extinction (Fig.2D), in which goal-directed behavior would be reflected by greater responding to the valued CS compared to the devalued CS. In each extinction test, there were 24 trials, i.e. 12 per CS. At the end of the task participants were asked about their explicit contingency knowledge and expectancy of receiving a shock. We also asked participants to rate on a scale from 0-100 the degree of unpleasantness of the shock, feelings of compulsion to respond in a habitual manner, and attempts to suppress habitual responses. Our measure of interest was the difference in participants' responses to the valued and the devalued CS following the long training session, i.e. following the second devaluation. We also recorded the number of shocks participants received during the training phase, as an indicator of unsuccessful avoidance learning performance. To ensure that task performance was not confounded by increased implicit fear of receiving electrical shocks, we additionally recorded skin conductivity during the task.

3. <u>Physiological Recordings</u>

Skin conductance responses (SCRs) were measured at 1000 Hz during the avoidance learning task. SCRs were identified within a 0.5 to 5 seconds window after the onset of each stimulus, using two seconds before stimulus onset as baseline. Reponses in microsiemens were subject to a 0.02 umho threshold, and SCRs less than 5% of the maximum were excluded. The entire task included 188 trials and recordings were conducted during each stage of the experiment ('Snap Test' during classical conditioning = 2 trials, practice run = 6 trials, short training = 12 trials, extinction = 24 trials, extended training = 120 trials, extinction = 24 trials). Due to technical difficulties and failure of detecting SCRs, data was only available in 92

participants (43 controls, 49 CUD; $\chi^2 = 2.69$; $p = 0.101$) out of 125 recorded data sets. Demographics of participants with missing data did not significantly differ from the rest of the sample.

4. Statistical Analyses

Data were analyzed using Statistical Package for Social Sciences (SPSS) version 21. Where appropriate, demographics, questionnaire responses, and SCRs were square-root transformed before the analysis to reduce skew and stabilize variances. Pearson *r* and Spearman *rho* correlations were used, as appropriate. Chi-square or Fisher's exact tests were used for the analysis of categorical data. Group differences regarding demographic and the total scores of the questionnaire measures were analyzed using independent samples t-tests or Mann-Whitney U tests, as appropriate. For tests where more than one condition or set of conditions was present, repeated-measures ANOVA was employed with group (2 levels: control, cocaine) as the between-subject factor and valence (2 levels: valued, devalued) as the within-subject factor. Where the assumption of sphericity was violated in repeated-measures ANOVA, within-subjects degrees of freedom were adjusted using the Greenhouse-Geisser correction.

In the appetite learning task, we subsequently included three variables separately as covariates in the analysis: i) accuracy scores of the action outcome test (Fig.1B), ii) the difference score between responses to valued and devalued stimuli of the control test (Fig.1D), and iii) overall learning accuracy during the training phase (Fig.1A). We included these covariates to verify that habitual responding was not confounded by group differences in outcome-action knowledge, working-memory/inhibition performance, or discrimination learning performance during training.

For the avoidance learning task, we included the mean intensity ratings to electrical shocks as a covariate in all analyses to statistically control for group differences. To verify that performance in the habit test was not confounded by CUDs poorer instrumental learning performance, we repeated the analysis using performance accuracy during overtraining as a covariate.

Two-step hierarchical multiple regression models were performed to investigate whether habitual responding can be predicted by drug dependency (as assessed by DSM-IV-TR criteria) or by other factors that have been linked with increased habit formation such as impulsivity, compulsivity, a history of stressful life events, and discrimination performance during the training phase. Habitual responding in the appetitive task was calculated from participants' responses in the slip-of-action test to the valued outcome minus the devalued outcome. We multiplied the difference score by -1 so that higher values reflected a greater habit bias. As recommended (30), we used the predictor–participant ratio of 1:15, resulting in a total of eight predictor variables. In Step 1, we included dependency status as four separate dichotomous variables (D1: not cocaine-dependent, cocaine-dependent; D2: not opiate-dependent, opiate-dependent; D3: not cannabis-dependent, cannabis dependent; D4: not alcohol-dependent, alcohol- dependent). In Step 2 the following continuous variables were entered: BIS-11 (impulsivity), OCI-R (compulsivity), SRRS (stressful life events), and mean discrimination performance accuracy over the eight training blocks (Fig.1A).

We also investigated the influence of drug dependency on avoidance responses in extinction to the valued CS (which were significantly attenuated in CUD) using the same two-step hieratical multiple regression model, as described above. In Step 2, we replaced mean discrimination accuracy with the mean avoidance accuracy of the prolonged training phase (Fig.2B).

All statistical tests were two-tailed and a significance level was set at 0.05. Due to technical problems, behavioral data of two CUD during appetitive learning were incomplete and had to be excluded from the analysis.

Supplementary Text

1. Relationship between verbal IQ and task performance

Appetitive Learning: Although the groups differed significantly in terms of verbal IQ, we only observed correlations between verbal IQ and discrimination learning accuracy during the training phase in controls (r = 0.30, $p = 0.029$) but not in CUD (r = 0.02, $p = 0.872$). No correlations were found between IQ and performance on the devaluation test (controls: $r = 0.14$, $p = 0.331$; CUD: $r = 0.12$, $p = 0.423$), or the slip-of- action test (controls: $r = 0.17$, $p = 0.227$; CUD: $r = 0.02$, $p = 0.906$) or the disinhibition/working-memory control task (controls: $r = 0.15$, $p = 0.312$; CUD: $r = 0.24$, $p = 0.089$). Consequently, IQ was not used as a covariate in the statistical analyses of appetitive learning.

Avoidance Learning: Verbal IQ was also not correlated with avoidance learning during the training phase (controls: $r = 0.14$, $p = 0.321$; CUD: $r = 0.16$, $p = 0.188$) or in extinction (controls: $r = 0.04$, $p = 0.758$; CUD: $r = 0.10$, $p = 0.416$). IQ was therefore not used as a covariate in the analysis of avoidance learning.

2. Behavioral Performance

Appetitive Learning

All participants steadily improved their **learning performance** over the eight training blocks, as reflected by a significant main effect of block ($F_{6,685} = 43.98$, $p < 0.001$) but no group-by-block interaction ($F_{6,685} = 1.09$, $p = 0.365$). Overall, CUDs' learning accuracy fell short from that seen in the control volunteers, as indexed by a highly significant main effect of group ($F_{1,121} = 20.2$, $p < 0.001$). Although response latency during feedback-learning declined in all participants across the training blocks ($F_{4,470} = 43.33$, $p < 0.001$), the decline was significantly sharper in controls compared with CUD, as indicated by a significant group-by-block interaction ($F_{4,470} = 4.13$, $p = 0.003$). Across all the training trials, response latency did not differ between the groups ($F_{1,121} = 0.06$, $p = 0.805$).

Following **outcome devaluation**, participants' outcome-action knowledge was tested behaviorally. CUD performed significantly less accurately ($t_{88.2} = 3.83$, $p < 0.001$) but not slower ($t_{121} = -1.40$, $p = 0.177$) compared with controls.

When participants' response strategies were subsequently examined in the **slip-of-action test**, CUD showed a strong habit bias, as reflected by a highly significant group-by-value interaction ($F_{1,106} = 18.26$, $p < 0.001$). There was also a significant main effect of group ($F_{1,121} = 14.13$, $p < 0.001$), suggesting that CUD made more responses overall. Post-hoc tests further confirmed that the significant group-by-value interaction was due to CUD making significantly more responses to the devalued outcome compared with controls ($t_{121} = -4.72$, $p < 0.001$) whilst responses to the valued outcome did not differ between the groups ($t_{121} = -0.65$, $p = 0.520$). To verify that CUDs' habit bias was not driven by insufficient outcome-action knowledge, we included performance accuracy of the outcome-action test as a covariate in the analysis. Indeed both, the main effect of group ($F_{1,120} = 7.99$, $p = 0.006$) and the group-by-value interaction ($F_{1,120} = 6.42$, $p = 0.013$) remained significant. We also verified that performance on the slip-of- action test was not confounded by impairments in working memory/inhibition, so we included the performance score of the control task (i.e. responses to the valued minus responses to the devalued stimulus) as a co-variate in the analysis. Both the group-by-value interaction ($F_{1,120} = 8.79$, $p = 0.004$) and the main effect of group ($F_{1,120} = 9.56$, $p = 0.002$) remained highly significant, suggesting that the strong habit bias in CUD was not

driven by cognitive impairments. Finally, we included performance accuracy of discrimination learning during the training phase as a co-variate in the analysis, and once again, the main effect of group ($F_{1,120}$ = 8.30, p = 0.005) and the group-by-value interaction ($F_{1,120}$ = 6.53, p = 0.012) remained significant, indicating that the strong habit bias in CUD was not confounded by poor learning performance. With respect to response latency, we did not observe main effects of value ($F_{1,101}$ = 1.31, p = 0.256) or group ($F_{1,101}$ = 2.46, p = 0.120) and no group-by-value interaction ($F_{1,101}$ = 1.99, p = 0.162) in the slip-of-action test.

During the test of **working memory/disinhibition,** which was included as a control condition for our primary habit analysis above, all participants responded predominantly to stimuli associated with reward compared with stimuli that were no longer associated with gaining points, as reflected by a highly significant main effect of value ($F_{1,121}$ = 111, p < 0.001). This difference in response preference was significantly smaller in CUD than in controls, as indicated by a significant group-by-valence interaction ($F_{1,121}$ = 42.10, p < 0.001), suggesting impairments in working memory or inhibitory control. Post-hoc tests further confirmed group differences for responses to both the valued ($t_{119.4}$ = 3.38, p = 0.001) and the devalued stimuli ($t_{91.7}$ = 3.38, p < 0.001). The overall response rate was again reduced in CUD ($F_{1,121}$ = 15.27, p < 0.001). All participants responded faster to valued than to devalued stimuli ($F_{1,121}$ = 8.31, p = 0.005). The main effect of group ($F_{1,121}$ = 2.08, p = 0.152) and the group-by-value interaction ($F_{1,121}$ = 0.99, p = 0.322) were both non-significant.

Test of explicit knowledge: At the end of the task, all participants were explicitly asked by questionnaire about their outcome knowledge. Compared with control volunteers, CUD demonstrated significant deficits in explicit knowledge in terms of stimulus-outcome (mean: U = 985, p < 0.001), response-outcome (U = 1250, p = 0.007) and stimulus-response (U = 1023, p < 0.001) relationships.

Avoidance Learning

Prior to the task, participants' levels of tolerance for shock intensity were measured during a work-up procedure. Although shock intensity did not objectively differ between the two groups (t_{120} = 0.37, p = 0.713), CUD rated the shocks both before and during the task as less intense than controls, as reflected by a significant main effect of group ($F_{1,122}$ = 7.80, p = 0.006). There was no main effect of time ($F_{1,122}$ = 0.29, p = 0.593) and no time-by-group interaction ($F_{1,122}$ = 0.28, p = 0.596). Mean intensity ratings were therefore included as a covariate in all subsequent analyses.

Prior to over-training, all participants completed a practice trial and a short phase, in which both CUD exhibited a tendency of fewer successful avoidance responses (practice: $F_{1,121}$ = 2.78, p = 0.098; training: $F_{1,121}$ = 3.60, p = 0.060). CUD and controls showed equivalent performance in our **devaluation sensitivity test**. Both groups understood the implications of disconnecting the electrodes and were able to integrate this information into their decision-making such that they made more accurate avoidance responses for shocks at the connected wrist (valued shock) compared to shocks at the disconnected wrist (devalued shock), as reflected by a significant main effect of value ($F_{1,121}$ = 30.29, p < 0.001). Critically, there was no interaction group-by-value interaction ($F_{1,121}$ = 0.49, p = 0.484), indicating equivalent devaluation sensitivity. There was, however, a main effect of group ($F_{1,121}$ = 7.10, p = 0.009) on response accuracy, as CUD were overall less successful than controls in avoiding shocks, consistent with their performance during the practice and training phases.

The over-training phase revealed consistent avoidance responses in both groups across the 10 training blocks, as reflected by the absence of a main effect of block ($F_{6.6,802}$ = 0.36, p = 0.916) and block-by-group interaction ($F_{6.6,802}$ = 1.04, p = 0.401). Overall CUD were less successful in avoiding electrical shocks compared with controls, as reflected by a significant main effect of group ($F_{1,121}$ = 11.28, p = 0.001) (Fig.

2B).

Following over-training of stimulus-response associations, we conducted the **critical habit test**. We first examined only accurate responses (i.e. response made on the correct pedal). Overall, participants made more responses to avoid the valued outcome compared to the devalued outcome, reflected by a significant main effect of value ($F_{1,121} = 20.05$, $p < 0.001$). The decline in responses to the devalued outcome was, however, slightly smaller in the cocaine group compared with the control group, as indicated by a marginally significant group-by-value interaction ($F_{1,121} = 3.23$, $p = 0.075$) but there was no significant main effect of group ($F_{1,121} = 2.62$, $p = 0.108$). The group-by-value interaction became, however, significant ($F_{1,121} = 7.39$, $p = 0.008$) and the main effect of group became marginally significant ($F_{1,121} = 3.83$, $p = 0.053$) when examined response vigor by considering all the responses participants made in the habit test, i.e. not just the correct responses. This significant interaction was driven by CUD making significantly fewer responses overall to the valued compared with devalued stimulus compared with controls. Post hoc tests further confirmed that CUD responded less frequently than controls to the stimulus still associated with the shock ($F_{1,121} = 10.65$, $p = 0.001$) but responded on par with controls to the stimulus no longer associated with a shock ($F_{1,121} = 0.85$, $p = 0.771$).

To verify that performance in the habit test was not confounded by CUDs poorer instrumental learning performance, we repeated the analysis using performance accuracy during overtraining as a covariate. When taking CUDs' lower overall accuracy level into account, both groups showed equivalent performance. Critically, both group-by-value interactions of correct responses ($F_{1,121} = 1.61$, $p = 0.207$) and total response ($F_{1,121} = 0.47$, $p = 0.494$) were non-significant, indicating no abnormal habit bias in CUD for avoidance learning.

Test of explicit knowledge: To test whether the observed impairments in avoidance responses were due to insufficient contingency knowledge in CUD, we explicitly asked all participants to indicate the stimulus-action-outcome associations they experienced during the training phase. All participants demonstrated intact awareness about the task contingences (100% accuracy in both groups). Most participants reported feeling a compulsion to press the pedal even though they knew that this was not necessary because they were not connected to the stimulator (59% controls, 64% CUD; $\chi^2 = 0.37$; $p = 0.541$), but only a subgroup of these reported making an attempt to suppress this compulsion (33% controls, 49% CUD; $\chi^2 = 1.78$; $p = 0.182$).

SCRs during Avoidance Learning

There were no group differences in anticipatory SCRs to the warning stimulus during training ($F_{1,89} = 0.71$, $p = 0.401$), indicating intact fear conditioning. When avoidance behavior was tested in extinction, all participants showed increased anticipatory SCRs to the valued relative to the devalued CS ($F_{1,88} = 8.23$, $p = 0.005$). There were no differences in SCRs between the groups ($F_{1,88} = 0.29$, $p = 0.592$) and no significant value-by-group interaction ($F_{1,88} = 0.04$, $p = 0.848$).

3. Regression Analysis

Appetitive Learning

In Step 1 of the hierarchical multiple regression model, four predictors were entered: cocaine dependence, opiate dependence, cannabis dependence, alcohol dependence. This model was statistically significant ($F_{4,117}$ = 4.48, p = 0.002) and explained 13% of the variance in habitual responding (Tab.S2). Cocaine dependence was the only significant predictor variable (β = .34, p = 0.007). Following the inclusion of BIS-11, OCI-R, SRRS, and goal-directed discrimination learning accuracy to the regression model at Step 2, the total variance explained by the model as a whole was 31% ($F_{8,113}$ = 6.32; p < 0.001). The introduction of the aforementioned variables thus explained an additional variance of 18% after controlling for drug dependency (ΔR^2 = 0.18, $F_{4,113}$ = 7.21, p < 0.001). An increased number of stressful life events (β = .30, p = 0.015) and poor learning performance during training (β = -.41, p < 0.001) were the only statistically significant predictors for habit bias in the final model (see also Tab.S2).

Avoidance Learning

In Step 1, four predictors were entered: cocaine dependence, opiate dependence, cannabis dependence, alcohol dependence. This model was statistically significant ($F_{4,119}$ = 2.28, p = 0.028) and explained 9% of the variance in appropriate avoidance of the valued CS in extinction (i.e. pressing the correct pedal in responses to the CS that was still associated with an electrical shock). Cocaine dependence was a significant negative predictor variable (β = -.30, p = 0.021) in the model. The inclusion of BIS-11, OCI-R, SRRS, and avoidance accuracy during training at Stage 2 explained an additional 44% of the variance (ΔR2 = 0.44, $F_{4,115}$ = 26.5, p < 0.001). Better avoidance accuracy during the training (β = .67, p < 0.001) and low levels impulsivity (β = -.18, p = 0.047) were the only two statically significant variables in the final model, predicting 52% of the variance of appropriate avoidance (R^2 = 0.52, $F_{4,115}$ = 15.8, p < 0.001) (see also Tab.S2).

Supplementary Tables

Table S1: Group characteristics of the full sample in means and standard deviations in parentheses.

	Control Group (n=53)	**Cocaine Group** (n=72)	**Group Comparison**	
	Mean (Std.)	Mean (Std.)	t or χ^2	*p*
Gender (% male)	94%	90%	0.68	0.408
Age (years)	41.3 (±10.5)	38.0 (±8.6)	1.74	0.084
Verbal IQ (NART score)	114.2 (±7.4)	101.6 (±8.1)	8.83	<0.001
Impulsivity (BIS-11 score)	59.3 (±7.6)	76.6 (±9.6)	-10.94	<0.001
Compulsivity (OCI-R score)	9.7 (±6.6)	17.3 (±13.4)	-3.37	0.001
Stressful events lifetime (SRRS score)	3.2 (±1.7)	8.2 (±1.9)	-15.24	<0.001
Systolic blood pressure (mm Hg)	128.6 (±13.6)	125.8 (±14.1)	1.13	0.261
Diastolic blood pressure (mm Hg)	77.4 (±10.8)	74.7 (±10.6)	1.38	0.169
Pulse (BPM)	66.2 (±11.8)	67.8 (±11.2)	-0.80	0.424
Alcohol use (AUDIT score)	4.2 (±2.0)	4.2 (±4.8)	-0.04	0.964
Drug use (DAST-20 score)	0.4 (±0.6)	----		
Cigarette use (number/day)	5.0 (±4.7)	10.8 (±6.1)	-2.29	0.025

Table S2: Summary of hierarchical regression analyses for variables predicting habitual responding during appetitive learning (upper part of the table). Habitual responding was indexed by the difference in participants' responses in the slip-of-action test to the valued minus the devalued outcome, multiplied by -1. Summary of hierarchical regression analyses for variables predicting impaired avoidance responses in extinction to the CS that was still associated with a shock (lower part of the table). (R: Pearson product-moment correlation coefficients; B: β-coefficients; SEB: standard error of the β coefficients; β: standardized β coefficients for the regression model; R^2: coefficients of determination; Δ: change; $p < 0.05$*; $p < 0.001$**).

	Appetitive Learning (Habit Index)						
Predicting variables	B	SEB	ß		B	SEB	ß
Step 1				*Step 2*			
Cocaine dependence	15.32	5.62	0.34*	Cocaine dependence	0.39	7.19	0.01
Opiate dependence	0.58	5.45	0.01	Opiate dependence	-4.91	5.05	-0.11
Cannabis dependence	0.47	5.32	0.01	Cannabis dependence	-2.51	4.96	-0.05
Alcohol dependence	4.02	9.89	0.04	Alcohol dependence	4.52	9.02	0.04
ΔR2			0.13**	Impulsivity (BIS-11)	2.46	3.38	0.08
				Compulsivity (OCI-R)	-0.12	1.22	-0.01
				Stressful life events (SRRS)	8.96	3.61	0.30*
				Discrimination accuracy (training phase)	-0.09	0.02	-0.41**
				ΔR2			0.18**
	Avoidance Learning (attenuated avoidance responses to the valued CS using the correct pedal)						
Predicting variables	B	SEB	ß		B	SEB	ß
Step 1				*Step 2*			
Cocaine dependence	-5.52	2.36	0.30*	Cocaine dependence	0.64	2.42	-0.03
Opiate dependence	-0.14	2.27	0.01	Opiate dependence	1.11	1.68	-0.06
Cannabis dependence	0.45	2.22	-0.02	Cannabis dependence	0.07	1.65	0.00
Alcohol dependence	2.53	4.16	-0.05	Alcohol dependence	-2.63	3.12	0.06
ΔR2			0.09*	Impulsivity (BIS-11)	-2.28	1.14	0.18*
				Compulsivity (OCI-R)	0.35	0.41	-0.06
				Stressful life events (SRRS)	-1.13	1.23	0.09
				Avoidance accuracy (training phase)	0.29	0.03	-0.67**
				ΔR2			0.44**

References and Notes

1. B. J. Everitt, T. W. Robbins, Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nat. Neurosci.* **8**, 1481–1489 (2005). Medline doi:10.1038/nn1579

2. F. J. Miles, B. J. Everitt, A. Dickinson, Oral cocaine seeking by rats: Action or habit? *Behav. Neurosci.* **117**, 927–938 (2003). Medline doi:10.1037/0735-7044.117.5.927

3. A. Dickinson, Actions and habits: The development of behavioural autonomy. *Philos. Trans. R. Soc. London Ser. B* **308**, 67–78 (1985). doi:10.1098/rstb.1985.0010

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

5. L. H. Corbit, B. C. Chieng, B. W. Balleine, Effects of repeated cocaine exposure on habit learning and reversal by *N*-acetylcysteine. *Neuropsychopharmacology* **39**, 1893–1901 (2014). Medline doi:10.1038/npp.2014.37

6. E. Dias-Ferreira, J. C. Sousa, I. Melo, P. Morgado, A. R. Mesquita, J. J. Cerqueira, R. M. Costa, N. Sousa, Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* **325**, 621–625 (2009). Medline doi:10.1126/science.1171203

7. G. Schoenbaum, B. Setlow, Cocaine makes actions insensitive to outcomes but not extinction: Implications for altered orbitofrontal-amygdalar function. *Cereb. Cortex* **15**, 1162–1169 (2005). Medline doi:10.1093/cercor/bhh216

8. Supplementary materials are available on *Science* Online.

9. K. D. Ersche, P. S. Jones, G. B. Williams, A. J. Turton, T. W. Robbins, E. T. Bullmore, Abnormal brain structure implicated in stimulant drug addiction. *Science* **335**, 601–604 (2012). Medline doi:10.1126/science.1214463

10. R. Z. Goldstein, N. D. Volkow, Dysfunction of the prefrontal cortex in addiction: Neuroimaging findings and clinical implications. *Nat. Rev. Neurosci.* **12**, 652–669 (2011). Medline doi:10.1038/nrn3119

11. R. J. Beninger, S. T. Mason, A. G. Phillips, H. C. Fibiger, The use of conditioned suppression to evaluate the nature of neuroleptic-induced avoidance deficits. *J. Pharmacol. Exp. Ther.* **213**, 623–627 (1980). Medline

12. J. D. Salamone, M. Correa, Motivational views of reinforcement: Implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* **137**, 3–25 (2002). Medline doi:10.1016/S0166-4328(02)00282-6

13. J. W. Dalley, T. D. Fryer, L. Brichard, E. S. Robinson, D. E. Theobald, K. Lääne, Y. Peña, E. R. Murphy, Y. Shah, K. Probst, I. Abakumova, F. I. Aigbirhio, H. K. Richards, Y. Hong, J. C. Baron, B. J. Everitt, T. W. Robbins, Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science* **315**, 1267–1270 (2007). Medline doi:10.1126/science.1137073

14. D. Belin, A. C. Mar, J. W. Dalley, T. W. Robbins, B. J. Everitt, High impulsivity predicts the switch to compulsive cocaine-taking. *Science* **320**, 1352–1355 (2008). Medline doi:10.1126/science.1158136

15. N. D. Volkow, G. J. Wang, J. S. Fowler, J. Logan, S. J. Gatley, R. Hitzemann, A. D. Chen, S. L. Dewey, N. Pappas, Decreased striatal dopaminergic responsiveness in detoxified cocaine-dependent subjects. *Nature* **386**, 830–833 (1997). Medline doi:10.1038/386830a0

16. D. Martinez, K. Greene, A. Broft, D. Kumar, F. Liu, R. Narendran, M. Slifstein, R. Van Heertum, H. D. Kleber, Lower level of endogenous dopamine in patients with cocaine dependence: Findings from PET imaging of D_2/D_3 receptors following acute dopamine depletion. *Am. J. Psychiatry* **166**, 1170–1177 (2009). Medline doi:10.1176/appi.ajp.2009.08121801

17. M. J. Frank, L. C. Seeberger, R. C. O'Reilly, By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science* **306**, 1940–1943 (2004). Medline doi:10.1126/science.1102941

18. S. de Wit, H. R. Standing, E. E. Devito, O. J. Robinson, K. R. Ridderinkhof, T. W. Robbins, B. J. Sahakian, Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. *Psychopharmacology* **219**, 621–631 (2012). Medline doi:10.1007/s00213-011-2563-2

19. K. Wunderlich, P. Smittenaar, R. J. Dolan, Dopamine enhances model-based over model-free choice behavior. *Neuron* **75**, 418–424 (2012). Medline doi:10.1016/j.neuron.2012.03.042

20. Z. Sjoerds, S. de Wit, W. van den Brink, T. W. Robbins, A. T. F. Beekman, B. W. J. H. Penninx, D. J. Veltman, Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Transl. Psychiatry* **3**, e337 (2013). doi:10.1038/tp.2013.107

21. J. M. Barker, J. R. Taylor, Habitual alcohol seeking: Modeling the transition from casual drinking to addiction. *Neurosci. Biobehav. Rev.* **47**, 281–294 (2014). Medline doi:10.1016/j.neubiorev.2014.08.012

22. C. M. Gillan, M. Papmeyer, S. Morein-Zamir, B. J. Sahakian, N. A. Fineberg, T. W. Robbins, S. de Wit, Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am. J. Psychiatry* **168**, 718–726 (2011). Medline doi:10.1176/appi.ajp.2011.10071062

23. C. M. Gillan, M. Kosinski, R. Whelan, E. A. Phelps, N. D. Daw, Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* **5**, e11305 (2016). Medline doi:10.7554/eLife.11305

24. C. M. Gillan, S. Morein-Zamir, G. P. Urcelay, A. Sule, V. Voon, A. M. Apergis-Schoute, N. A. Fineberg, B. J. Sahakian, T. W. Robbins, Enhanced avoidance habits in obsessive-compulsive disorder. *Biol. Psychiatry* **75**, 631–638 (2014). Medline doi:10.1016/j.biopsych.2013.02.002

25. H. E. Nelson, *National Adult Reading Test Manual* (NFER-Nelson, 1982).

26. J. Hayaki, M. D. Stein, J. A. Lassor, D. S. Herman, B. J. Anderson, Adversity among drug users: Relationship to impulsivity. *Drug Alcohol Depend.* **78**, 65–71 (2005). Medline doi:10.1016/j.drugalcdep.2004.09.002

27. J. H. Patton, M. S. Stanford, E. S. Barratt, Factor structure of the Barratt impulsiveness scale. *J. Clin. Psychol.* **51**, 768–774 (1995). Medline doi:10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

28. E. B. Foa, J. D. Huppert, S. Leiberg, R. Langner, R. Kichic, G. Hajcak, P. M. Salkovskis, The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychol. Assess.* **14**, 485–496 (2002). Medline doi:10.1037/1040-3590.14.4.485

29. Y. Worbe, G. Savulich, S. de Wit, E. Fernandez-Egea, T. W. Robbins, Tryptophan depletion promotes habitual over goal-directed control of appetitive responding in humans. *Int. J. Neuropsychopharmacol.* **18**, 1–9 (2015).

30. J. Stevens, *Applied Multivariate Statistics for the Social Sciences* (Lawrence Erlbaum Associates, 1996).