Impaired Pavlovian conditioned inhibition in offenders with personality disorders

Zhimin He [a] , Helen J. Cassaday [a] , Richard C. Howard [b] , Najat Khalifa [c] & Charlotte Bonardi [a]

[a] School of Psychology, University of Nottingham, Nottingham, NG7 2RD, UK

[b] Community Health Sciences, Division of Psychiatry, University of Nottingham, Nottingham, NG7 2RD, UK

[c] Nottinghamshire Healthcare NHS Trust, The Wells Rd Centre, The Wells Rd, Nottingham, NG2 7UA, UK

Available online: 31 Aug 2011

PLEASE SCROLL DOWN FOR ARTICLE

Ψ Psychology Press
Taylor & Francis Group

Impaired Pavlovian conditioned inhibition in offenders with personality disorders

Zhimin He[1], Helen J. Cassaday[1], Richard C. Howard[2], Najat Khalifa[3], and Charlotte Bonardi[1]

[1]School of Psychology, University of Nottingham, Nottingham, NG7 2RD, UK
[2]Community Health Sciences, Division of Psychiatry, University of Nottingham, Nottingham, NG7 2RD, UK
[3]Nottinghamshire Healthcare NHS Trust, The Wells Rd Centre, The Wells Rd, Nottingham, NG2 7UA, UK

Certain types of violent offending are often accompanied by evidence of personality disorders (PDs), a range of heterogeneous conditions characterized by disinhibited behaviours that are generally described as impulsive. The tasks previously used to show impulsivity deficits experimentally (in borderline personality disorder, BPD) have required participants to inhibit previously rewarded responses. To date, no research has examined the inhibition of responding based on Pavlovian stimulus–stimulus contingencies, formally "conditioned inhibition" (CI), in PDs. The present study used a computer-based task to measure excitatory and inhibitory learning within the same CI procedure in offenders recruited from the "personality disorder" and the "dangerous and severe personality disorder" units of a high-security psychiatric hospital. These offenders showed a striking and statistically significant change in the expression of inhibitory learning in a highly controlled procedure: The contextual information provided by conditioned inhibitors had virtually no effect on their prepotent associations. Moreover, this difference was not obviously attributable to nonspecific cognitive or motivational factors. Impaired CI would reduce the ability to learn to control associative triggers and so could provide an explanation of some types of offending behaviour.

Keywords: Associative learning; Conditioned inhibition; Personality disorders; Forensic.

Certain types of violent offending are often accompanied by evidence of personality disorders (PDs). The characteristic PDs include a set of heterogeneous conditions that have in common a tendency to be deviant, troublesome, and persistent (Fazel & Danesh, 2002; Hiscoke, Langstrom, Ottosson, & Grann, 2003). Clinical accounts of antisocial (ASPD) and borderline (BPD) offenders confirm that impulsive and violent behaviours are typical. However, the personality profile of offenders is not clear-cut, and there is a high degree of comorbidity between ASPD and BPD.

Another psychological profile often identified with violent offenders is psychopathy (Hare, 1991). This condition shows clear overlap with ASPD (Blackburn & Coid, 1998; Coid &

Ullrich, 2010; Hare, Hart, & Harpur, 1991; Hart & Hare, 1996; Kosson, Lorenz, & Newman, 2006). Moreover, *DSM–IV* (*Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition*; American Psychiatric Association, 1994) describes features of psychopathy as "particularly distinguishing of Antisocial Personality Disorder in prison or forensic settings" (*DSM–IV*, p. 647). Notably both psychopathy and ASPD are characterized by a disinhibited lifestyle and a tendency to transgress social norms and legal rules.

Thus some have argued that a unifying feature in the psychological profile of forensic populations is poor impulse control, similarly recognized as a feature of psychopathy (Johansson, Kerr, & Andershed, 2005; Lesch & Merschdorf, 2000; Prichard, 1837) and BPD (Stein, Hollander, & Liebowit 1993; Stein, Towney, & Hollander, 1995). Impulsivity, a complex and multidimensional concept, spans lack of inhibitory control, actions without forethought, novelty seeking, and inability to delay gratification (Barratt, 1985, 1994). Impulsivity in psychiatric disorders is obviously more extreme than that in healthy individuals and has been described as: "a predisposition toward rapid, unplanned reactions to internal or external stimuli without regard to the negative consequences of these reactions to the impulsive individual or to others" (Moeller, Barratt, Dougherty, Schmitz, & Swann, 2001, p. 1784).

But despite its obvious importance, the exact nature of impulsivity has not been unambiguously specified, making further analysis difficult. For example, some have argued that impulsive behaviour results from lack of inhibitory control (Buss & Plomin, 1975), others that it stems from an inability to tolerate delays of reinforcement (e.g., Logue, 1988; Logue et al., 1992; Thiébot, Le Bihan, Soubrié, & Simon, 1985). In any event, individuals with high impulsivity fail to inhibit unwanted actions, and thus behavioural measures of impulsivity include a range of established laboratory behavioural tasks designed to measure this deficit, such as the go/no-go task, the stop-signal task, and antisaccadic eye movement procedures. These tasks can be characterized by the fact that the subject is normally required to make a specific

response, but on certain occasions is required to withhold it, and a failure to do so is taken as an index of impulsivity; for example, in a stop-signal task the subject is required to respond unless the stop signal is presented. Moreover, deficits in the performance of such tasks have been demonstrated in participants with BPD and psychopathy (e.g., Grootens et al., 2008; Kiehl, Smith, Hare, & Liddle, 2000; Newman & Schmitt, 1998; Nigg, Silk, & Stavor, 2005; Rentrop et al., 2007; Rubio et al., 2007; Ruchsow et al., 2008).

However, inhibition is a broad construct and should not be too narrowly identified with any one behavioural paradigm. A response may occur not only because it has been explicitly rewarded, as in the traditional tests of impulsivity, but also as a *conditioned* response to an environmental stimulus that has consistently signalled some outcome of motivational value (as a result of Pavlovian stimulus–stimulus, S–S, contingencies rather than response–stimulus, R–S, learning). Moreover, in the chain of cause and effect that ultimately results in unwanted actions, environmental cues that trigger associated thoughts and emotions through such S–S associations can be primary. The so-called weapons effect—that the presence of a weapon can heighten the level of aggression displayed—provides the classic example of how potentially antisocial responses can be triggered by environmental cues, in this case guns (Berkowitz & LePage, 1967). Moreover, such situations are typically constituted by more complex configural cues: For example, the sight of the gun generally precedes the sensory stimuli provided by grasping the weapon. Thus, the inhibition of S–S associations could play a critical role in suppressing unwanted behaviours. As it cannot be assumed that the same psychological mechanisms inhibit S–S and R–S associations (Nigg, 2000), the inhibition of responding based on Pavlovian S–S contingencies—formally "conditioned inhibition" (CI)—should also be examined.

In CI procedures, the expectation of an outcome is inhibited by the presence of a qualifying stimulus. Building on the basic classical conditioning procedure in which a conditioned stimulus (CS) signals an outcome (unconditioned stimulus, US),

a further stimulus (the conditioned inhibitor) signals the omission of the expected US. For example if, after a number of pairings of a CS with a US (A+), A is paired with another stimulus B and the US is omitted (AB−), subjects learn that B signals the absence of the US—that is, it becomes a conditioned inhibitor (Pavlov, 1927). While a few studies have successfully demonstrated CI in humans (Grings, Carey, & Schell, 1974; Migo et al., 2006; Neumann, Lipp, & Siddle, 1997; Wilkinson, Lovibond, Siddle, & Bond, 1989), to date very few studies have explored CI in relation to individual differences, and these used quite different tasks and measures (Kantini, Cassaday, Hollis, & Jackson, 2011; Migo et al., 2006); moreover, to our knowledge, to date none have demonstrated CI deficits in relation to psychological or psychiatric disorder. If the effective disinhibition of responding based on undesirable S–S contingencies (such as the sight followed by the feel of a weapon; Berkowitz & LePage, 1967) contributes to the well-documented inability of offenders to inhibit punished responses, we might expect to see a deficit in CI in this subject group. Thus the present study was devised as a first test of the prediction that offenders with PD (in the absence of comorbid schizophrenia or learning disability) would show impaired CI.

To date, learning differences and attentional anomalies have been most extensively investigated specifically in relation to psychopathy, rather than PDs as defined by *DSM–IV* criteria. Such studies have also tended to use learning tasks in which outcomes can be negative, as a defining feature of impulsivity is heedless action despite aversive consequences (Cleckley, 1964; Flor, Birbaumer, Hermann, Ziegler, & Patrick, 2002; Gorenstein & Newman, 1980; Kosson & Newman, 1986; Kosson, Smith, & Newman, 1990; Lykken, 1957; Newman, 1987; Newman & Kosson, 1986; Newman, Patterson, & Kosson, 1987; Newman & Schmitt, 1998; Schmauk, 1970; Thornquist & Zuckerman, 1995; Vitale & Newman, 2001). However, as discussed above, psychopathy has been argued to be on a continuum with ASPD (Blackburn & Coid, 1998; Coid & Ullrich, 2010; Kosson et al., 2006). Thus, since a high proportion

of the present sample were expected to be psychopathic and therefore poor at aversively motivated learning, we used a positively motivated learning task. This also allows us to rule out any contribution of insensitivity to aversive outcomes—a true deficit in CI would be evident regardless of the valence of the omitted outcome. Finally, in forensic populations of this type, poor impulse control may contribute to general learning deficits, whenever unwanted actions interfere with task performance (Ávila & Parcet, 2001; Gullo, Jackson, & Dawe, 2010). Therefore the task developed for use in this population measured both inhibitory and excitatory learning within the same procedure, to simultaneously monitor any nonspecific, general learning deficits.

Method

Participants

A total of 26 offenders volunteered to participate in the experiment, 2 of whom were subsequently excluded for a priori procedural reasons (see below). They were all male inpatients at Rampton Hospital, a high-security psychiatric hospital in the UK. All the participants recruited to the present study had a history of criminal offending behaviour, including convictions for violent and sexual offences.

Clinical assessments were based on the International Personality Disorder Examination (IPDE; Loranger et al., 1994) and the Psychopathy Check List–Revised (PCL–R; Hare, 1991). The IPDE provides categorical and dimensional scores for PDs in the form of 10 subscales that relate to the 3 clusters identified by *DSM–IV* (American Psychiatric Association, 1994). Although the PCL–R is a dimensional instrument, for research purposes, scores of 25 or above are generally taken to define psychopaths in European samples (Cooke & Michie, 1999; Dolana & Doylea, 2007; Grann, Langstrom, Tengstrom, & Stalenheim, 1998; Harris, Rice, & Cormier, 1991; Harris, Rice, & Quinsey, 1993; Howells, Krishnan, & Daffern, 2007; Langstrom et al., 1999; Rice, Harris, & Cormier, 1992). The IPDE and PCL–R had been completed following

admission, in the course of their initial assessment by qualified staff at Rampton Hospital.

Assessment of violence was based on offending history using a Severity of Violence Rating Scale (SVRS) adapted from that originally developed (Gunn & Robertson, 1976) and validated in hospitalized forensic patients (Wong, Lumsden, Fenton, & Fenwick, 1993). The SVRS was modified in house for use at Rampton by adding a measure of institutional behaviour. Thus, there were three subscales—Violence in Index Offence, Violence in Criminal Record, and Violence in the Institution. Each was rated on a 4-point scale (0 = minimal/no violence, 4 = severe violence, indicating someone's life or health was seriously endangered), yielding a total score out of 12. Patients' history of violent offending was derived from their case files, and this information was supplemented by self-report data.

Of those who completed the study, 8 participants were in the Personality Disorder Unit (PDU) and 16 in the Dangerous and Severe Personality Disorder (DSPD) Unit. To meet the criteria for severe PD justifying admission to the DSPD unit, an offender must: (a) score 30 or more on the PCL–R; (b) score between 25 and 30 on the PCL–R plus have at least one *DSM–IV* PD diagnosis other than ASPD; or (c) have two or more *DSM–IV* PD diagnoses (Howells et al., 2007).

According to the available IPDE categorical diagnoses (n = 24), of those participants that were finally included, there were definite (n = 18) or probable (n = 4) diagnoses of ASPD and definite (n = 13) or probable (n = 1) diagnoses of BPD, and 1 participant had a confirmed obsessive compulsive personality disorder but was comparable with others in the sample in that this was comorbid with BPD. The one offender without any PD diagnosis according to the IPDE had nonetheless met the criteria for admission to the DSPD (he was diagnosed as ASPD with conduct disorder according to *DSM–IV* assessment). Psychiatric assessments confirmed that none of the offenders approached in connection with the study had comorbid schizophrenia or learning disability. Fourteen of the offenders were on psychotropic

medication, including typical (n = 1) and atypical antipsychotics (n = 6), antidepressants (n = 7), anxiolytics (n = 7), and anticonvulsants (n = 1).

The healthy controls were a community-based sample of 27 participants, 3 of whom were subsequently excluded for *a priori* procedural reasons (see below), leaving 24 to complete the study. These controls lived in the same county and were recruited at the University of Nottingham (ancillary staff), Nottingham National Ice Centre, and the Nottingham Trent FM Arena. They were all without a higher education; some were unemployed, and others reported having jobs such as swimming instructor, driver, waiter, shop assistant, and university support staff. None reported or showed any indication of mental illness or substance abuse. Table 1 shows the details of participants' age, gender, ethnicity, and educational level. Controls were matched in age to the offender group, $t(46) = 1.35$, $p = .18$. Allocation to the counterbalanced experimental conditions was identical for the offender and control groups. Other than the fact that they were not incarcerated, control participants were tested under comparable quiet environmental conditions in the same way by the same experimenter (Z.H.).

This study was approved by NHS Research Ethics (Derbyshire Research Ethics Committee, Reference No. 08/H0401/65, September 2008) with amendment to study offenders (May 2009). Procedures for testing the control participants were approved by the University of Nottingham, School of Psychology Ethics Committee. Control

Table 1. *Summary details of the final sample of participants*

	Offenders (n = 24)	Control (n = 24)
Age (years)	39.5	34.92
Range of age (years)	25–58	19–56
Gender	All males	All males
Education level	Up to A level[a]	Up to A level[a]
Ethnicity	23 White and 1 Black	23 White and 1 Black

[a]In the UK, the number of years in education required to achieve A level is 14.

participants received an inconvenience allowance of £5 cash to cover their travel expenses. No such payment was possible in the case of the offender group, but they had no travel expenses or loss of earnings in consequence of participation.

Stimuli

Lego block pictures ($n = 9$) were used as the CSs in the present study. The USs were selected by a pilot study from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005). The IAPS provides a set of images, standardized on the basis of participants' ratings, on the dimensions of valence and arousal from 1 to 9, 1 representing a low rating on each dimension and 9 a high rating (i.e., 1 as low pleasure, low arousal). Based on standardized ratings, the USs in the present study were provided by 10 IAPS "positive" images and 10 IAPS "neutral" images; because of the nature of the sample, pictures depicting children and erotic images were excluded from the present study.[1] Whenever a "reinforced" trial was required, a positive US was presented, whereas neutral USs followed the CS on "nonreinforced" trials. Conditioning was measured using a rating scale: Participants were asked to guess or predict what kind of picture would follow presentation of the Lego blocks using a rating scale from 1 (neutral) to 9 (positive), with the rating 5 to reflect uncertainty.

Design and analysis

The design of the experiment is shown in Table 2. There were three stages: (a) preexperimental, (b) training with elemental and compound stimuli, and (c) test. In the preexperimental stage, the participants rated the stimuli and stimulus compounds to be used in the test stage, to preclude the possibility that differences in responding to the stimuli at test could be due to biases present before the start of training.

In the elemental training stage, two of the neutral CSs, A and C, were paired with reinforcement, positive IAPS pictures (A+ and C+ trials),

while a further two, U and V, were paired with nonreinforcement, neutral IAPS pictures (U– and V– trials). This training provided a measure of participants' simple associative learning. It also established A and C as excitatory CSs (i.e., signalling a positive outcome), which facilitated the subsequent establishment and detection of CI.

During the compound training stage, the AZ compound signalled reinforcement (AZ+), whereas the AP compound signalled nonreinforcement (AP–). As A had been paired with reinforcement in the previous stage, presenting AP allowed P to signal the absence of the reinforcement otherwise indicated by A and was thus expected to establish P as a conditioned inhibitor. Two additional stimulus compounds, CY and BX, were reinforced and nonreinforced, respectively.

Although successful discrimination between AZ and AP would be consistent with the proposal that P was a conditioned inhibitor, it was not sufficient; for example, participants might respond more to AZ simply because Z was reinforced on every trial. In order to establish unequivocally that P was a conditioned inhibitor, we conducted a summation test—more specifically, we examined whether P would suppress responding to a different excitatory stimulus more than would a suitable control stimulus (cf. Rescorla, 1969). The continued excitatory training with C on CY+ trials (C had also been reinforced in the previous stage) means it provided an excitatory test stimulus against which the inhibitory effects of P could be evaluated. The BX– trials were designed to establish X as a control stimulus, which was presented the same number of times as P and in a similar manner (in compound with another stimulus and paired with nonreinforcement); however, the stimulus with which X was presented was novel, so that X, unlike P, did not signal the absence of reinforcement during this training stage and so should not have acquired any inhibitory properties.

The test stage compared the ratings given to the trained stimuli and stimulus compounds that had signalled reinforcement (A, C, AZ, CY) and non-

---

[1] Positive images: 1440, 1610, 1750, 1920, 8370, 8380, 2395, 5480, 7330, 7502; Neutral images: 2393, 2396, 2512, 2890, 7006, 7055, 7175, 7185, 7187, 6150 (see also Supplementary Materials Online).

Table 2. *Design of the experiment*

Preexperimental phase		Elemental training phase		Compound training phase		Test phase	
CSs	Trials (N)	CSs	Trials (N)	CSs +/− outcome	Trials (N)	CSs	Trials (N)
A	2	A+	12	AZ+	8	A	2
C	2	U−	12	AP−	12	C	2
AZ	2	V−	12	BX−	12	AZ	2
AP	2	C+	12	CY+	8	AP	2
BX	2					BX	2
CY	2					CY	2
CP	2					CP	4
CX	2					CX	4

Note: CS = conditioned stimulus. Letters denote the 9 conditioned stimuli (pictures of Lego blocks); these were counterbalanced (see text). With respect to unconditioned stimulus presentations, " + " represents presentation of a positive IAPS (International Affective Picture System) picture and "−" presentation of a neutral IAPS picture.

reinforcement (AP, BX), but the critical comparison was between two additional test compounds, CP and CX. C was excitatory and was predicted to elicit high ratings indicating expectation of reinforcement. If P was a conditioned inhibitor, it should reduce this high rating to C, whereas the critical comparison stimulus, X, should not. CI would therefore be evident as lower ratings to CP than to CX. The identities of the stimuli used as P and X were counterbalanced across the participants, as were those of A and B (and C and V, see above).

The dependent variable was the mean rating given for each particular trial type, which was assessed at each stage of the experiment. Statistical analyses were by a mixed-design analysis of variance (ANOVA), with group (offender vs. control) as a between-subjects factor, and discrimination (e.g., A+ vs. U− and C+ vs. V−), reinforcement (reinforced or not), and trial block as within-subjects factors. Significant interactions were explored with simple main effects analysis using the pooled error term. Planned comparisons of the assessment score data were by t test, or Kruskal–Wallis where the data were not normally distributed (i.e., to compare the IPDE scores). The measure of effect size given for mixed-design ANOVAs was Pearson's correlation coefficient, r.

Correlational analyses (Pearson's or Spearman's, as appropriate) were used to compare overall learning and questionnaire scores. A summary measure of excitatory learning was provided by the difference in mean ratings on C and V trials during the initial training stage—that is, C − V. As C was the excitatory stimulus, the greater the C − V score, the higher the level of excitatory learning. A summary measure of CI was provided by the difference between the mean ratings on CP and CX trials given during the test stage—that is, CX − CP. P was the putative inhibitor and thus was supposed to suppress evaluation of C more than X; thus the higher this score, the greater the inhibitory learning.

An *a priori* exclusion criterion was applied to the elemental training performance: Five participants (two offenders and three controls) who failed to learn the simple discrimination between C+ and V− trials—that is, rating scores (C − V) ≤ 0—were excluded from the CI analysis (the C + /V− discrimination was fully counterbalanced for stimulus identity for this purpose).

Procedure
Participants were invited to take part in a research study on learning using a computer-based task. Before the task, each participant had to read the information sheet and sign a consent form.

The task instructions were that a cat "Mogwai" would bring participants either a positive picture or a neutral, boring picture, depending on what kind of Lego blocks she found in her basket. Participants were asked to guess or predict what kind of picture would follow presentation of the Lego blocks using a rating scale from 1 (neutral) to 9 (positive), with the rating 5 to reflect uncertainty. Reminder instructions were presented on screen at each stage of the procedure.

Before the start of the preexperimental phase, participants were shown some example CSs and USs, and further explanation was given as necessary. The samples of CS and US pictures were individually colour printed on a 4.5 × 6-cm card, and these pictures were representative of, but not subsequently used as, stimuli during the experiment. Participants were told that the whole computer-based experimental session would last about 20 min and comprise three stages. At the same time, they were shown an example of CS presentations with the rating scale and were told that during the experiment they would need to click the corresponding number to guess or predict the valence of the US (a positive or a neutral picture) according to the different Lego blocks that had been presented. Participants were encouraged to ask questions at this stage. The three stages of the computer-based experimental session then followed.

Preexperimental stage. During the first (preexperimental) stage of the experiment, participants were told they must guess what kind of picture the cat might bring based on the Lego blocks presented, although the instructions specified that no pictures would follow. After participants clicked on a number button to guess the US valence, the next presentation followed in a semirandom sequence. In this and all subsequent stages of the experiment, CS presentations were counterbalanced for right/left position on the screen across participants, and the various trial types were presented in a semirandom sequence (constrained by the total number of trials of a particular type scheduled in each stage). There was a total of 16 presentations, 2 of each stimulus or stimulus combination presented (these being A, C, AZ, AP, BX, CY, CP, and CX; see Table 2).

Training stages. On completion of the preexperimental ratings, the conditioning trials commenced, and US presentations were introduced. The instructions were as before, but with the exception that participants were advised that following their guess they would be shown the picture that the cat had brought. The first training stage used the CS elements and comprised six training blocks, each with two of each of the four kinds of trial (A + , U–, V–, and C+). After the participant had clicked a number button to predict the valence of the US to follow, a US, randomly selected from the pool of positive or neutral USs as appropriate, was shown on the screen for 1 s. This was followed by a 1-s gap before the next trial started, during which a picture of the cat Mogwai (around 6 × 6 cm) was presented in the middle of the screen on a white background (see Supplementary Materials Online). The second, compound training stage followed directly after this training with the CS elements and comprised four kinds of trial (AZ + , AP–, BX–, and CY+). There were 8 training trials with each of the reinforced compounds (AZ and CY) and 12 with each of the nonreinforced compounds (AP and BX). This proportion of reinforced to nonreinforced trial types has produced robust CI in previous studies; the different trial types were analysed in two blocks of 4 and 6, respectively.

Test stage. The test stage was exactly the same as the preexperimental stage, except that the number of presentations of the critical test compounds CP and CX was increased from 2 to 4. According to the CS compound on the screen, participants were asked to predict the valence of the US that would follow. However, no USs were presented in the test stage. As in the earlier stages of the experiment, there were on-screen reminders of the task instructions.

Throughout the experiment, whenever participants asked questions or made comments they were asked to try to focus on the task and to try to remember or guess which outcome (positive or neutral picture) was predicted by the Lego blocks.

Results

Preexperimental stage

The preexperimental stage established that there were no preexisting differences between the two critical stimulus compounds CP and CX. ANOVA with stimulus (CP vs. CX) and group (offender vs. control) as factors revealed no significant effects or interactions, largest $F(1, 46) = 2.22$, $p = .14$.

Training stage

Excitatory training. In the initial training stage, participants learned a simple discrimination between the reinforced A and C and the nonreinforced U and V; this provided a measure of simple excitatory learning in the two groups. It is evident from Figure 1 that both groups came to respond accurately on this task, although there was some indication that the offender group showed a slightly poorer discrimination, especially at the start of the training phase.

This impression was confirmed by the results of an ANOVA with group, discrimination (A+ vs. U–, C+ vs. V–), reinforcement, and training block. This showed a main effect of reinforcement, $F(1, 46) = 260.66$, $p < .001$, $r = .92$, which interacted with group, $F(1, 46) = 4.40$, $p = .04$, $r = .295$. Simple main effects analysis revealed that the groups differed on reinforced trials, $F(1, 92) = 4.14$, $p = .04$, but not on nonreinforced trials, $F(1, 92) = 3.08$, $p = .08$. However, the difference between reinforced and nonreinforced

trials was significant in both control and offender groups: $F(1, 46) = 166.38$, and $F(1, 46) = 98.68$, respectively, $p s < .001$. Thus, although both groups learned to respond differently on reinforced and nonreinforced trials, the offender group responded slightly less on reinforced trials than did the control participants, which could indicate a general learning impairment in this group.

Reinforcement also interacted significantly with training block, $F(5, 230) = 26.69$, $p < .001$, $r = .75$, confirming that the discriminations developed over the course of training. Simple main effects confirmed that responding on reinforced trials increased, and that on nonreinforced trials decreased—there were significant effects of block on both types of trial, $F(5, 460) = 16.24$, $p < .001$, and $F(5, 460) = 16.44$, $p < .001$, respectively. In addition, the difference between reinforced and nonreinforced trials was significant on all training blocks, smallest $F(1, 276) = 20.32$, $p < .001$. These results demonstrate that the discrimination developed over training—and there was no sign that this differed between the two groups: The interaction between reinforcement, training block, and group was not significant, $F < 1$.

Significant interactions were also found between training block and discrimination, and between training block, discrimination, and group, $F(5, 230) = 2.49$, $p = .03$, $r = .27$, and $F(5, 230) = 2.87$, $p = .02$, $r = .33$, respectively. However, as none of these interactions involved the factor reinforcement, they did not bear on the ease with which the discriminations were learned and so were not considered further. Nothing else was significant, the largest $F(1, 46) = 1.65$, $p = .21$.

Inhibitory training. During the second stage, participants were trained on the key discrimination between the reinforced AZ and the nonreinforced AP, which was designed to turn P into a conditioned inhibitor. They were also trained to discriminate between the reinforced CY and the nonreinforced BX. It is clear from Figure 2 that both groups learned this task, although the offender group showed a less clear discrimination between the two types of trial, especially at the start of training.
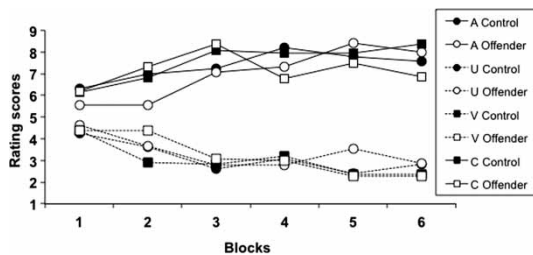

Figure 1. *Rating scores for A, U, V, and C in the control and offender groups at the elemental training stage. A rating of 9 reflected expectation of a positive image, 1 reflected expectation of a neutral image, and 5 indicated uncertainty.*
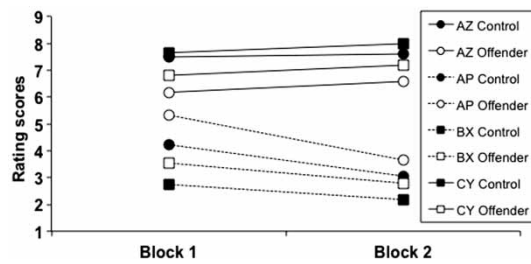
Figure 2. *Rating scores for AZ, AP, BX, and CY in the control and offender groups at the compound training stage. A rating of 9 reflected expectation of a positive image, 1 reflected expectation of a neutral image, and 5 indicated uncertainty.*

An ANOVA with group (offender vs. control), and discrimination (AZ+ vs. AP–, and CY+ vs. BX–), reinforcement (reinforced or not), and training block (1–2) as factors, revealed no main effect of either group, $F < 1$, or discrimination, $F(1, 46) = 2.25, p = .14$. However, the main effects of training block and reinforcement were significant, $F(1, 46) = 5.65, p = .02, r = .33$, and $F(1, 46) = 168.28, p < .001, r = .89$, respectively, as was the interaction between reinforcement and group, $F(1, 46) = 11.01, p = .002, r = .44$, suggesting that the offenders might have learned less effectively than the control participants. This was confirmed by simple main effects analyses, which revealed that the offenders responded less on reinforced trials and more on nonreinforced trial than did the controls, $F(1, 92) = 10.55, p = .002$, and $F(1, 92) = 5.35, p = .02$, respectively. Nonetheless, the difference between reinforced and nonreinforced trials was significant in both groups, $F(1, 46) = 46.60$ and 132.70 for offender and control groups, respectively, $ps < .001$, confirming that both groups had learned the discrimination.

There was a significant interaction between training block and reinforcement, $F(1, 46) = 32.85, p < .001, r = .65$, confirming that the discrimination developed over the two training blocks; simple main effects revealed that there was an effect of blocks for nonreinforced trials, $F(1, 92) = 29.07, p < .001$, but not for reinforced trials, $F(1, 92) = 2.68, p = .11$, and that the difference between reinforced and nonreinforced trials was significant on both training blocks,

$F(1, 92) = 92.89, p < .001$, and $F(1, 92) = 201.12, p < .001$, respectively.

A significant interaction was also found between discrimination and reinforcement, $F(1, 46) = 19.52, p = .001, r = .55$, suggesting there was a difference between the difficulty of the AZ+/ AP– and CY+/BX– discriminations. This might be expected: The nonreinforced AP compound contained element A that was also present on the reinforced AZ trials, whereas the CY+/BX– discrimination did not involve reinforcement of either element of the nonreinforced BX compound. The simple main effects accordingly revealed that ratings were higher on AP– trials than on BX– trials, $F(1, 92) = 15.70, p < .001$, but that there was no difference in ratings between CY+ and AZ+ trials, $F(1, 92) = 2.70, p = .10$. However, the difference between reinforced and nonreinforced trials was significant for both discrimination types, $F(1, 92) = 64.83$, and $F(1, 92) = 172.83$, for AZ/AP and CY/BX, respectively, $ps < .001$, suggesting that both discriminations had nonetheless been learned effectively. Nothing else was significant, largest $F(1, 46) = 1.92, p = .17$.

Test stage

The ratings for the training stimulus compounds were maintained in the test stage—an impression confirmed by statistical analysis—and so these data are not considered further. Figure 3 shows the rating scores for the two critical stimulus compounds, CP and CX, at the preexperimental and test stages. It is clear that during the preexperimental stage, the ratings of the two compounds were similar in both groups, but that at test, although the rating of CP was markedly lower than that of CX in control participants, no such difference was evident in the offender group. This description was confirmed by the results of an ANOVA with group, stage (preexperimental vs. test), and stimulus compound (CP vs. CX) as factors; this revealed a significant three-way interaction, $F(1, 46) = 10.03, p < .01$, suggesting that the groups differed in the way in which the CP/ CX ratings changed as a result of training. In order to explore this three-way interaction further, two further ANOVAs, with stage and
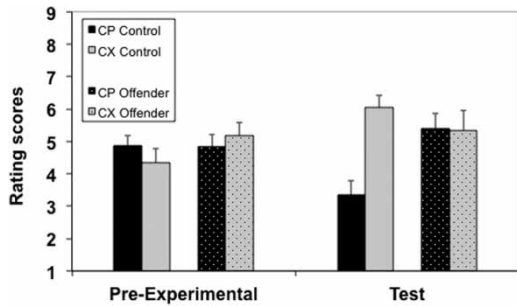
Figure 3. *Rating scores for the key comparison stimuli CP and CX at the preexperimental and the test stages in the control and offender groups. A rating of 9 reflected expectation of a positive image, 1 reflected expectation of a neutral image, and 5 indicated uncertainty.*

stimulus compound as factors, were performed on the data from each group. In the control group, this revealed no main effect of either stage, $F < 1$, or stimulus compound, $F(1, 23) = 8.00$, $p = .10$, but a significant interaction between these two factors, $F(1, 23) = 27.89$, $p < .001$, $r = .75$. Simple main effects revealed that there was no difference between CP and CX at the preexperimental stage, $F(1, 23) = 2.40$, $p = .14$, but that at test, participants gave significantly lower ratings to CP than to CX, $F(1, 23) = 30.80$, $p < .001$. The fact that P was able to suppress the positive ratings of C substantially more than was the control stimulus X is consistent with P being a conditioned inhibitor. No such pattern was evident in the offender group: Here the corresponding ANOVA revealed nothing significant, all Fs < 1. There was thus no evidence of CI in the offender group.

Although this apparent abolition of CI in the offenders was a striking effect, one could argue that the offenders were showing signs of impaired learning throughout the experiment, so that this test effect might reflect no more than a general learning deficit. Indeed, there was statistical evidence in the first stage of training that the offenders were worse at simple excitatory conditioning, and the associated effect size was moderate ($r = .295$). However, this suggestion is weakened by an examination of the individual differences that participants displayed in this task.

Individual differences

In the offender group, there was no significant correlation between any of the IPDE subscales and the summary measure of inhibitory or excitatory learning, largest $r(23) = .34$, $p = .10$. This was also true when the IPDE dimensional scores were subgrouped as three clusters according to *DSM–IV*, largest $r(23) = .17$, $p = .43$. Similarly, there was no correlation between PCL–R scores, either overall or by subscale, and the summary measure of excitatory or inhibitory learning, largest $r(24) = .29$, $p = .17$, for Factor 2.

However, there did appear to be differences between participants in the DSPD and PD units, despite the fact that levels of learning during the excitatory and inhibitory training phases were relatively similar in these two clinical subgroups (Figures 4 and 5). ANOVA of rating scores from the excitatory and inhibitory training phases, but comparing only the PD or DSPD subgroups, revealed no effects or interactions involving group, smallest $p = .14$. The test data for the two units are shown in Figure 6 (for comparison with Figure 3). Although the figure suggests that the DSPD showed relatively less inhibition at test than the PD participants, analysing these data in the same manner as those in Figure 3 did not support this conclusion, possibly in part because of the apparent difference between the two groups at the preexperimental stage (though this was not significant, Fs < 1).

However, the additional analyses of the summary scores of excitatory and inhibitory
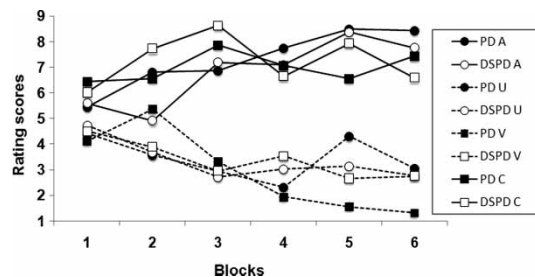

Figure 4. *Rating scores for A, U, V, and C in the PD (personality disorder) and DSPD (dangerous and severe personality disorder) groups at the elemental training stage. A rating of 9 reflected expectation of a positive image, 1 reflected expectation of a neutral image, and 5 indicated uncertainty.*
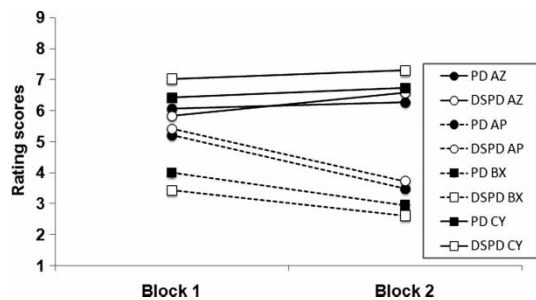
Figure 5. *Rating scores for AZ, AP, BX, and CY in the PD (personality disorder) and DSPD (dangerous and severe personality disorder) groups at the compound training stage. A rating of 9 reflected expectation of a positive image, 1 reflected expectation of a neutral image, and 5 indicated uncertainty.*

learning (as used for the other analyses in relation to individual differences) confirmed the impression that the level of CI was indeed further reduced in the DSPD participants. The summary excitatory learning scores (C – V) were 4.05, 3.89, and 5.12 for the PD, DSPD, and control participants, respectively; the corresponding scores for the CX – CP measure of CI were 1.91, −1.03, and 2.69. ANOVA with group (PD, DSPD, and control) and measure (excitatory vs. inhibitory) was performed on these data and revealed a significant interaction between these two factors, $F(2, 45) = 10.41$, $p < .001$, $r = .43$; simple main effects

Figure 6. *Rating scores for the key comparison stimuli CP and CX at the preexperimental and the test stages in the PD (personality disorder) and DSPD (dangerous and severe personality disorder) groups. A rating of 9 reflected expectation of a positive image, 1 reflected expectation of a neutral image, and 5 indicated uncertainty.*

showed that there was an effect of group on the CI measure, $F(2, 90) = 11.63$, $p < .0001$, but not on the excitation measure, $F(2, 90) = 1.43$, $p = .25$. Tukey's test showed that whereas DSPD differed from the control participants $p = .001$, PD participants did not, $p = .78$. A further analysis on these summary scores, but for PD and DSPD groups only, revealed an almost significant interaction, $F(1, 22) = 4.18$, $p = .053$, $r = .399$; simple main effects revealed that the groups differed on the inhibitory but not the excitatory learning measure, $F(1, 44) = 6.40$, $p = .015$, and $F < 1$, respectively.

We therefore conclude that the deficit in CI was greater in the DSPD participants. With respect to the source of the difference in performance between these two subgroups of offenders, the IPDE dimensional scores were not significantly different in the PD and DSPD units on any of the subscales, with one exception: There was a significant difference in avoidant PD between the two units, $\chi^2(1, N = 24) = 4.04$, $p = .05$ (PD unit mean scores = 7.13, $SD = 4.39$; DSPD unit mean scores = 2.75, $SD = 3.36$). Nothing else was significant, the largest $\chi^2(1, N = 24) = 2.20$, $p = .14$. Furthermore, when the IPDE dimensional scores were subgrouped as the three *DSM–IV* clusters, there was no significant difference between the two units, largest $t(22) = 1.84$, $p = .10$. However, consistent with the admissions criteria, participants in the DSPD unit had significantly higher PCL–R scores (by both the individual factors and the total scores) than those in the PD unit, smallest $t(22) = 2.53$, $p = .02$, $r = .47$. Thus, participants with sufficiently high PCL–R scores to warrant DSPD admission showed overall less CI than those with lower PCL–R scores. Moreover, PCL–R Factor 1 and Factor 2, as well as the total scores, were significantly correlated with the dimensional scores for ASPD, $r(24) = .53$, $p = .01$, $r(24) = .41$, $p = .04$, and $r(24) = .61$, $p = .001$, respectively, but not BPD, largest $r(24) = .24$, $p = .27$. In addition, although they did not differ in age, $t(22) = 0.28$, $p = .78$, the PD group had been incarcerated significantly longer than the DSPD group, $t = −2.43$, $p = .03$. This direction of effects is opposite to what we
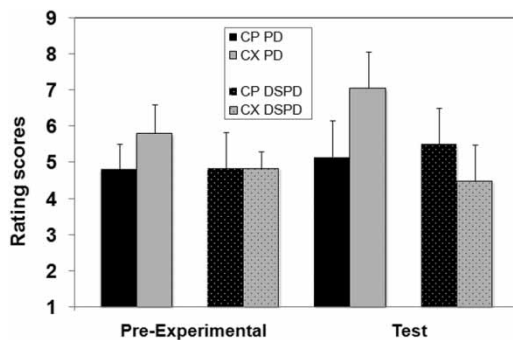
Downloaded by [ ] at 08:13 27 April 2012

would expect if incarceration were the cause of the CI deficit.

Thus, differences in ASPD could modulate the level of CI shown in these violent offenders. This suggestion is supported by further analyses using the SVRS. There was no significant difference by the Violence in Criminal Record subscale, $t(22) = 1.70$, $p = .10$, or by the Violence Total score, $t(22) = 1.32$, $p = .20$. There was a significant difference in participants recruited from the PD and the DSPD units in the level of Violence in Index Offence subscale, $t(22) = 2.55$, $p = .02$, but this took the direction that PD unit patients showed greater severity of violence in their index offence (PD unit mean scores = 3.00, $SD = 0.54$; DSPD unit mean scores = 2.12, $SD = 1.15$). It was the severity of Violence in the Institution that reflected the dangerousness of the DSPD compared to PD patients, $t(22) = 2.19$, $p = .05$ (PD unit mean scores = 1.62, $SD = 1.41$; DSPD unit mean scores = 2.94, $SD = 1.34$). With respect to mediation of the CI deficit, scores on the Violence in Criminal Record subscale were positively correlated with the measure of inhibitory learning, $r(24) = .41$, $p = .04$. Thus, the higher the violence ratings in relation to criminal record, the worse the CI learning performance.

Finally, medication status made no difference to performance on the task. There were no differences in excitatory learning scores (both measures), or in inhibitory learning scores (both measures), between offenders on ($n = 14$) and off ($n = 10$) medication: For antidepressants, largest $t(22) = 1.11$, $p = .28$; for anxiolytics, largest $t(22) = 0.78$, $p = .44$; for antidepressants and/or anxiolytics, largest $t(22) = 1.27$, $p = .22$; for antipsychotics, largest $t(22) = 1.54$, $p = .14$; for any form of psychotropic medication, largest $t(22) = 1.51$, $p = .15$.

Discussion

This experiment provided evidence that, while CI was clearly demonstrated in the community control participants, it was effectively absent in a group of offenders, particularly in the DSPD subgroup. The critical test comparison, ratings of CP

and CX, involved fully counterbalanced stimulus compounds and took into account preexperimental differences in their ratings. In addition, the training history of P and X was matched in every respect apart from the fact that P signalled the absence of a reinforcing picture, whereas X did not.

There was some evidence of a more general learning deficit in the offender group, so one must consider the possibility that our results do not represent a specific impairment in CI, but a more nonspecific effect on learning. For example, it may be that the positive pictures had less salience for the offender group, which could have lessened both excitatory and inhibitory learning in these participants. However, although we have no direct evidence against this possibility, the impairment in excitatory conditioning in the offender group was statistically less consistent, as well as being numerically more modest, than the impairment in CI, and the CI deficit was significantly greater than the effect on excitatory learning in offenders. This pattern became clearer when the PD and DSPD subgroups were considered separately: Here analysis of the summary measures of excitation and inhibition showed clearly that, while excitatory learning did not differ among the three groups, inhibitory learning did, and this difference stemmed from a selective deficit in inhibitory learning in the DSPD group. This pattern of results makes it unlikely that the inhibitory learning deficit was an indirect result of a more general learning deficit.

A curious feature of the results was that performance in the inhibitory learning phase appeared to be less impaired than performance on the summation test. This was especially evident in the comparison of the PD and DSPD groups, who showed essentially identical performance throughout training, but differed significantly at test. It might be that the critical difference between these two groups had less to do with inhibitory learning than with a differential tendency in the DSPD participants to process compound cues in a *configural* manner—such that they were unable to identify, or respond appropriately to, elements of the compound when they were presented in a different compound, as occurred in the summation test (cf., Pearce, 1987). Unfortunately at present we

do not have any data that would allow us to discriminate between these two possibilities.

Any firm conclusion that CI is impaired in this forensic population of course depends on the adequacy of the control condition. Although the offenders all had PD, while the community control participants did not, it must be acknowledged that the differences between forensic patients in a high-security hospital and healthy community controls are doubtlessly vast across many other social, economic, and psychological dimensions. Nonetheless, control participants were deliberately matched as closely as possible with offenders in terms of general factors, including educational level and socioeconomic status, and participants who did not learn in the first training stage were excluded. Moreover, the inevitable limitations of the control condition were mitigated by the fact that the inhibitory learning deficit was relatively confined to the DSPD subgroup of offenders. Still, differences in general intelligence or motivational factors among the groups cannot be ruled out; however, although such differences might be able to account for a generally different level of performance, it is less clear how they could explain the relatively specific deficit in CI.

It is also important to consider the possibility that medication might be sufficient explanation of the loss of CI in the offenders, as a relatively large proportion of this sample, especially of the DSPD participants, was on benzodiazepines or antipsychotics of some description, and such medications generally impair cognitive function. Moreover, we have previously found that CI was impaired under medication with clonidine in a subset of participants with Tourette syndrome, even though there was no overall impairment in the diagnosed group (Kantini et al., 2011). Nonetheless, a number of arguments may be made against the possibility that medication mediated the difference in CI observed in the present study. First, statistical analysis did not reveal any difference in performance on either the excitatory or inhibitory tasks according to medication, either in general or by specific type. Second, the diversity of medication regime makes it unlikely that this could provide a systematic account of the failure of offenders to

show CI. Nonspecific effects of medication—for example, on arousal, attention, or motivation to engage with the task—would be expected to depress performance throughout, and this was not the pattern of effects observed in the present study. For example, as noted above, by the end of training the offender group was performing as well as controls on the inhibitory discrimination, and a clear deficit only emerged in the transfer test.

Finally, there are other confounded factors that could contribute to the observed difference in CI. For example, the controls could not be matched for incarceration. Because of its admissions criteria, it was not possible—within the same institution—to recruit offenders without PD, psychiatric illness, or learning disability, as a further comparison group to allow us to examine the role of incarceration in a secure hospital. Nonetheless, given the significant difference in CI between the participants from the PD and DSPD units, we would argue that incarceration per se does not seem to be the critical issue—especially as the DSPD group, who showed worse CI than the PD group, had been incarcerated for less time, rather than more. We would suggest that a more natural conclusion is that inhibitory learning deficits may contribute to the cognitive profile of an individual whose behaviours result in incarceration—rather than suggest that incarceration per se has a selective cognitive effect. In summary, the pattern of results obtained is not obviously explicable in terms of the nature of our control condition.

It is not immediately apparent how such a selective effect on expression of CI in our offender group might be mediated. We have previously identified a role for individual differences in CI (using a different variant of the CI task) in a larger sample of normal participants (Migo et al., 2006), although the scales used in this earlier study measured variation in normal function, and so some discontinuity can be expected in the socially deviant sample who participated in the present study (Howard & Duggan, 2009). However, although a high proportion of the sample had confirmed ASPD and/or BPD according to the IPDE categorical diagnoses, there was no correlation between any of the IPDE dimensional measures and CI scores.

Similarly, there was no correlation between CI scores and psychopathy levels as measured by the PCL–R, either overall or by either of the subscales (Hare, 1991; Lykken, 1995). Certainly one should not place too much weight on this null result as, in common with other experimental studies of this type, the sample size was underpowered for such correlational analyses, and so the failure to find correlations here does not mean none could be found in a larger study.

Moreover, direct comparison of the PD and DSPD subgroups, using the same summary learning scores, showed that the DSPD group showed worse CI than the PD participants. Critically, these subgroups of offenders differed significantly in their PCL–R scores, an index of psychopathy—suggesting that the CI impairment could stem from this difference in the personality profile of these two participant groups. But it is possible that the true picture is more complex. DSPD unit patients are typically characterized by the co-occurrence of high PCL–R scores and an ASPD diagnosis and, frequently, a BPD diagnosis (Howard & Duggan, 2009). This pattern of comorbidity is associated with a significant degree of serious, in particular violent, offending (Coid & Ullrich, 2010; Kosson et al., 2006) and with high scores on a dimension of hostile impulsivity, characterized by aggression, resentment, deviance, and paranoid beliefs together with affective dyscontrol (Blackburn, 2009). This may provide a further explanation of the lack of correlation between the CI deficit and the PCL–R scores taken in isolation, since it is the co-occurrence of psychopathy with ASPD (and frequently BPD as well) that characterizes these deviant and disinhibited patients, rather than simply a high PCL–R score. Moreover, patients in the DSPD unit were found to have significantly lower ratings of avoidant personality disorder—that is, less social inhibition, feelings of inadequacy, hypersensitivity to negative evaluation, and avoidance of social interaction. To the extent this constellation of traits relates also to levels of cognitive inhibition, this profile is consistent with the finding of the present study, that their CI was relatively worse than that demonstrated in the PD unit.

We therefore conclude that the demonstration that the participants in the DSPD showed less CI than those in the PD unit is consistent with a relationship between CI and the level of deviancy and disinhibition within the offender group. In this sense it is significant that higher PCL–R scores, as was found in the DSPD patients, went together with elevated ASPD but not BPD scores, suggesting that higher ASPD could also be driving the poorer CI in the DSPD group. Conversely a diagnosis of psychopathy did not affect the level of CI in participants with comorbid BPD/ASPD, making it less likely that psychopathy in isolation was a critical factor. Importantly, correlational analyses showed that the higher the SVRS scores in relation to criminal record (Gunn & Robertson, 1976; Wong et al., 1993), the worse the CI learning performance. Although not a direct measure of personality, this finding points to the role of long-standing individual characteristics in relation to CI deficits.

Irrespective of how the effect was mediated, in terms of implications for our understanding of cognitive control in offenders, the loss of CI reflects a selective effect on an aspect of associative learning that could potentially be relevant to both the symptom profile and the ability to control unwanted behaviours that have identified triggers. Stimuli provided by environmental events are recognized to trigger associations that generate habitual thoughts and feelings (Ferguson & Cassaday, 1999; Lishman, 1987; Siegel, 1977; Stewart, de Wit, & Eikelboom, 1984; Watson, 1924), which can in turn generate unwanted behaviours. Similarly, other environmental events and thoughts provide contextual stimuli with the potential to become learned inhibitors of these unwanted associations. However, in offenders, precisely where the need to inhibit such associations is greater, impaired inhibition of responding based on S–S contingencies may leave some individuals less able to inhibit the unwanted thoughts that can lead to unwanted actions. Thus, therapeutic interventions to improve learning by inhibiting antecedent S–S associations could be an effective behavioural approach to symptom control (via an

action on the associative chain that generates triggers).

It should be noted that the S–S association being inhibited in the present study was between the Pavlovian CS and a motivationally significant outcome (a positive picture) rather than a specific image: The inhibited CR was elicited by a stimulus that had predicted a variety of different images that shared only their positive motivational value. We are also unable to specify exactly what the participants learned in order to be able to inhibit their behaviour. For example, the CI could have suppressed the ability of the conditioned excitor to elicit expectation of the outcome, but it might equally have suppressed performance of the conditioned response directly (Rescorla, 1993; Rescorla & Holland, 1977). The present study is not able to discriminate between these, or other, possibilities. But the critical point is that, by whatever mechanism the response was suppressed, it was the product of learning based on Pavlovian S–S contingencies, not an operant response that had been rewarded directly. This critical theoretical distinction sets the present task apart from other tests of impulsivity and broadens the operational and theoretical criteria by which disorders in impulsivity may be identified.

Theoretical accounts of ASPD, BPD, and psychopathy do not specifically address CI. However, the results of the present study are broadly consistent with predictions arising from a model of deficits in emotionally impulsive PD (Howard, Fenton, & Fenwick, 1982). Although we find CI deficits in association with offending behaviour and level of violence in participants' criminal record rather than psychopathy specifically, the results of the present study are similarly consistent with a recent reformulation of the response modulation deficit hypothesis of psychopathy (Gorenstein & Newman, 1980; Vitale & Newman, 2009). Newman and colleagues have suggested that the insensitivity to inhibitory cues that characterizes psychopathy is a result of anomalous attentional modulation of responding (Newman & Lorenz, 2003). This attentional anomaly has the consequence that contextual information has less effect on prepotent S–R responding (Newman & Wallace, 1993; Patterson & Newman, 1993;

Zeier, Maxwell, & Newman, 2009). We already know that bad actions are a defining feature of offenders; the present study provides a direct investigation of the cognitive mechanisms necessary to dampen certain kinds of trigger. Specifically, the results of the present study show that the contextual information provided by conditioned inhibitors has virtually no effect on the expression of prepotent associations based on Pavlovian S–S contingencies.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.

Ávila, C., & Parcet, M. A. (2001). Personality and inhibitory deficits in the stop-signal task: The mediating role of Gray's anxiety and impulsivity. *Personality and Individual Differences*, *29,* 975–986.

Barratt, E. S. (1985). Impulsiveness defined within a systems model of personality. In E. P. Speilburger & J. N. Butcher (Eds.), *Advances in personality assessment* (pp. 113–132). Hillsdale, NJ: Lawrence Erlbaum Associates.

Barratt, E. S. (1994). Impulsiveness and aggression. In J. Monahan & H. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 61–79). Chicago, IL: University of Chicago Press.

Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology*, *7,* 202–207.

Blackburn, R. (2009). Subtypes of psychopath. In M. McMurran & R. C. Howard (Eds.), *Personality, personality disorder and violence* (pp. 113–132). Chichester, UK: John Wiley & Sons.

Blackburn, R., & Coid, J. (1998). Psychopathy and the dimensions of personality disorders in violent offenders. *Personality and Individual Differences*, *25,* 129–145.

Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York, NY: Wiley-Interscience.

Cleckley, H. (1964). *The mask of sanity*. St. Louis, MO: Mosby.

Coid, J., & Ullrich, S. (2010). Antisocial personality disorder is on a continuum with psychopathy. *Comprehensive Psychiatry*, *51,* 426–433.

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology*, *108,* 58–68.

Dolana, M., & Doylea, M. (2007). Psychopathy: Diagnosis and implications for treatment. *Psychiatry*, *6,* 404–408.

Fazel, S., & Danesh, J. (2002). Serious mental disorder among 23,000 prisoners: Systematic review of 62 surveys. *Lancet*, *16,* 545–550.

Ferguson, E., & Cassaday, H. J. (1999). The Gulf War and illness by association. *British Journal of Psychology*, *90,* 459–475.

Flor, H., Birbaumer, N., Hermann, C., Ziegler, S., & Patrick, C. J. (2002). Aversive Pavlovian conditioning in psychopaths: Peripheral and central correlates. *Psychophysiology*, *39,* 505–518.

Gorenstein, E. E., & Newman, J. P. (1980). Disinhibitory psychopathology: A new perspective and a model for research. *Psychological Review*, *87,* 301–315.

Grann, M., Langstrom, N., Tengstrom, A., & Stalenheim, E. G. (1998). The reliability of file-based retrospective ratings of psychopathy with the PCL–R. *Journal of Personality Assessment*, *70,* 416–426.

Grings, W. W., Carey, C. A., & Schell, A. M. (1974). Comparison of two methods for producing response inhibition in electrodermal conditioning. *Journal of Experimental Psychology*, *103,* 658–662.

Grootens, K. P., van Luijtelaar, G., Buitelaar, J. K., van der Laan, A., Hummelen, J. W., & Verkes, R. J. (2008). Inhibition errors in borderline personality disorder with psychotic-like symptoms. *Progress in Neuro-psychopharmacology and Biological Psychiatry*, *32,* 267–273.

Gullo, M. J., Jackson, C. J., & Dawe, S. (2010). Impulsivity and reversal learning in hazardous alcohol use. *Personality and Individual Differences*, *48,* 123–127.

Gunn, J., & Robertson, G. (1976). Drawing a criminal profile. *British Journal of Criminology*, *16,* 156–160.

Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist–Revised*. Toronto, Canada: Multi-Health Systems.

Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the proposed DSM–IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, *100,* 391–398.

Harris, G. T., Rice, M. E., & Cormier, C. (1991). Psychopathy and violent recidivism. *Law and Human Behavior*, *15,* 625–637.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, *20,* 315–335.

Hart, S. D., & Hare, R. D. (1996). Psychopathy and antisocial personality disorder. *Current Opinion in Psychiatry*, *9,* 129–132.

Hiscoke, U. L., Langstrom, N., Ottosson, H., & Grann, M. (2003). Self-reported personality traits and disorders (DSM–IV) and risk of criminal recidivism: A prospective study. *Journal of Personality Disorders*, *17,* 293–305.

Howard, R., & Duggan, C. (2009). Mentally disordered offenders: Personality disorders. G. Towl & D. Crighton (Eds.), *Forensic psychology* (pp. 309–328). Oxford, UK: Blackwell Publishing.

Howard, R. C., Fenton, G. W. F., & Fenwick, P. B. C. (1982). *Event-related brain potentials in personality and psychopathology: A Pavlovian approach*. Letchworth, UK: Research Studies Press, J. Wiley & Sons.

Howells, K., Krishnan, G., & Daffern, M. (2007). Challenges in the treatment of dangerous and severe personality disorder. *Advances in Psychiatric Treatment*, *13,* 325–332.

Johansson, P., Kerr, M., & Andershed, H. (2005). Linking adult psychopathy with childhood hyperactivity–impulsivity–attention problems and conduct problems through retrospective self-reports. *Journal of Personality Disorders*, *19,* 94–101.

Kantini, E., Cassaday, H. J., Hollis, C. P., & Jackson, G. M. (2011). The normal inhibition of associations is impaired by clonidine in Tourette Syndrome. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *20,* 96–106.

Kiehl, K. A., Smith, A. M., Hare, R. D., & Liddle, P. F. (2000). An event-related potential investigation of response inhibition in schizophrenia and psychopathy. *Biological Psychiatry*, *48,* 210–221.

Kosson, D. S., Lorenz, A. R., & Newman, J. P. (2006). Effects of comorbid psychopathy on criminal offending and emotion processing in male offenders with antisocial personality disorder. *Journal of Abnormal Psychology*, *115,* 798–806.

Kosson, D. S., & Newman, J. P. (1986). Psychopathy and allocation of attentional capacity in a divided-attention situation. *Journal of Abnormal Psychology*, *95*, 257–263.

Kosson, D. S., Smith, S. S., & Newman, J. P. (1990). Evaluation of the construct validity of psychopathy in black and Caucasian male inmates: Three preliminary studies. *Journal of Abnormal Psychology*, *99*, 250–259.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Instruction manual and affective ratings* (Tech. Rep. No. A-6). Gainesville, FL: University of Florida.

Langstrom, N., Grann, M., Tengstrom, A., Lindholm, N., Woodhouse, A., & Kullgren, G. (1999). Extracting data in file-based forensic psychiatric research: Some methodological considerations. *Nordic Journal of Psychiatry*, *53*, 61–67.

Lesch, K. P., & Merschdorf, U. (2000). Impulsivity, aggression, and serotonin: A molecular psychobiological perspective. *Behavioral Sciences & the Law*, *18*, 581–604.

Lishman, W. A. (1987). *Organic psychiatry: The psychological consequences of cerebral disorder* (2nd ed., pp. 207–276). Oxford, UK: Blackwell Science.

Logue, A. W. (1988). Research on self-control: An integrating framework. *Behaviour and Brain Science*, *11*, 665–709.

Logue, A. W., Tobin, H., Chelonis, J. J., Wang, R. Y., Geary, N., & Schachter, S. (1992). Cocaine decreases self-control in rats: A preliminary report. *Psychopharmacology*, *109*, 245–247.

Loranger, A. W., Sartorius, N., Andreoli, A., Berger, P., Buchheim, P., Channabasavanna, S. M., et al. (1994). The International Personality Disorder Examination, IPDE. The WHO/ADAMHA International Pilot Study of Personality Disorders. *Archives of General Psychiatry*, *51*, 215–224.

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology*, *55*, 6–10.

Lykken, D. T. (1995). *The antisocial personalities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Migo, E. M., Corbett, K., Graham, J., Smith, S., Tate, S., Moran, M. P., et al. (2006). A novel test of conditioned inhibition correlates with personality measures of schizotypy and reward sensitivity. *Behavioural Brain Research*, *168*, 299–306.

Moeller, F. G., Barratt, E. S., Dougherty, D. M., Schmitz, J. M., & Swann, A. C. (2001). Psychiatric aspects of impulsivity. *The American Journal of Psychiatry*, *158*, 1783–1793.

Neumann, D. L., Lipp, O. V., & Siddle, D. A. T. (1997). Conditioned inhibition of autonomic Pavlovian conditioning in humans. *Biological Psychology*, *46*, 223–233.

Newman, J. P. (1987). Reaction to punishment in extra-verts and psychopaths: Implications for the impulsive behaviour of disinhibited individuals. *Journal of Research in Personality*, *21*, 464–480.

Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*, *96*, 257–263.

Newman, J. P., & Lorenz, A. R. (2003). Response modulation and emotion processing: Implications for psychopathy and other dysregulatory psychopathology. In R. J. Davidson, K. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 1043–1067). New York, NY: Oxford University Press.

Newman, J. P., Patterson, C. M., & Kosson, D. S. (1987). Response perseveration in psychopaths. *Journal of Abnormal Psychology*, *96*, 145–148.

Newman, J. P., & Schmitt, W. (1998). Passive avoidance in psychopathic offenders: A replication and extension. *Journal of Abnormal Psychology*, *107*, 527–532.

Newman, J. P., & Wallace, J. F. (1993). Diverse pathways to deficient self-regulation: Implications for disinhibitory psychopathology in children. *Clinical Psychology Review*, *13*, 690–720.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*, 220–246.

Nigg, J. T., Silk, K. R., & Stavor, G. (2005). Disinhibition and borderline personality disorder. *Development and Psychopathology*, *17*, 1129–1149.

Patterson, C. M., & Newman, J. P. (1993). Reflectivity and learning from aversive events: Toward a psychological mechanism for the syndromes of disinhibition. *Psychological Review*, *100*, 716–736.

Pavlov, I. P. (1927). *Conditioned reflexes*. London, UK: Oxford University Press.

Pearce, J. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review*, *94*, 61–73.

Prichard, J. C. (1837). *A treatise on insanity and other disorders affecting the mind*. Philadelphia, PA: Haswell, Barrington & Haswell.

Rentrop, M., Backenstrass, M., Jaentsch, B., Kaiser, S., Roth, A., Unger, J., et al. (2007). Response inhibition in borderline personality disorder: Performance in a go/no-go task. *Psychopathology*, *41*, 50–57.

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*, 77–94.

Rescorla, R. A. (1993). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*, *21*, 327–336.

Rescorla, R. A., & Holland, P. C. (1977). Associations in Pavlovian conditioned inhibition. *Learning and Motivation*, *8*, 429–447.

Rice, M. E., Harris, G. T., & Cormier, C. (1992). Evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior*, *16*, 399–412.

Rubio, G., Jiménez, M., Rodríguez-Jiménez, R., Martinez, I., Iribarren, M. M., Jiménez-Arriero, M. A., et al. (2007). Varieties of impulsivity in males with alcohol dependence: The role of Cluster-B personality disorder. *Alcoholism, Clinical and Experimental Research*, *31*, 1826–1832.

Ruchsow, M., Groen, G., Kiefer, M., Buchheim, A., Walter, H., Martius, P., et al. (2008). Response inhibition in borderline personality disorder: Event-related potentials in a go/nogo task. *Journal of Neural Transmission*, *115*, 127–133.

Schmauk, F. J. (1970). Punishment, arousal, and avoidance learning in sociopaths. *Journal of Abnormal Psychology*, *76*, 325–335.

Siegel, S. (1977). Morphine tolerance acquisition as an associative process. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*, 1–13.

Stein, D. J., Hollander, E., & Liebowit, M. R. (1993). Neurobiology of impulsivity and the impulse control disorders. *Journal of Neuropsychiatry and Clinical Neurosciences*, *5*, 9–17.

Stein, D. J., Towney, J., & Hollander, E. (1995). The neuropsychiatry of impulsive aggression. In E. Hollander & D. Stein (Eds.), *Impulsivity and aggression* (pp. 91–105). New York, NY: Wiley.

Stewart, J., de Wit, H., & Eikelboom, R. (1984). Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. *Psychological Review*, *91*, 251–268.

Thiébot, M. H., Le Bihan, C., Soubrié, P., & Simon, P. (1985). Benzodiazepines reduce the tolerance to reward delay in rats. *Psychopharmacology*, *86*, 147–152.

Thornquist, M. H., & Zuckerman, M. (1995). Psychopathy, passive-avoidance learning and basic dimensions of personality. *Personality and Individual Differences*, *19*, 525–534.

Vitale, J. E., & Newman, J. P. (2001). Response perseveration in psychopathic women. *Journal of Abnormal Psychology*, *110*, 644–647.

Vitale, J. E., & Newman, J. P. (2009). Psychopathic violence: A cognitive-attention perspective. In M. McMurran & R. C. Howard (Eds.), *Personality, personality disorder and violence* (pp. 247–263). Chichester, UK: John Wiley & Sons.

Watson, J. B. (1924). *Behaviorism*. New York, NY: Norton.

Wilkinson, G. M., Lovibond, P. F., Siddle, D. A. T., & Bond, N. (1989). Effects of fear-relevance on electrodermal safely signal learning. *Biological Psychology*, *28*, 89–104.

Wong, M., Lumsden, J., Fenton, G., & Fenwick, P. (1993). Violence ratings of special hospital patients. *Journal of Forensic Psychiatry*, *4*, 471–480.

Zeier, J. D., Maxwell, J. S., & Newman, J. P. (2009). Attention moderates the processing of inhibitory information in primary psychopathy. *Journal of Abnormal Psychology*, *118*, 554–563.