

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/117595>

Please be advised that this information was generated on 2018-12-19 and may be subject to change.

# Reversal deficits in individuals with psychopathy in explicit but not implicit learning conditions

Inti A. Brazil, MSc; Joseph H.R. Maes, PhD; Inge Scheper, MSc; Berend H. Bulten, PhD; Roy P.C. Kessels, PhD; Robbert Jan Verkes, MD, PhD; Ellen R.A. de Bruijn, PhD

Brazil, Maes, Kessels — Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands; Brazil, Bulten, Verkes — Pompestichting, Nijmegen, The Netherlands; Scheper, Verkes — Radboud University Nijmegen Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Department of Psychiatry, Nijmegen, The Netherlands; Kessels — Radboud University Nijmegen Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Department of Medical Psychology and Geriatrics, Nijmegen, The Netherlands; de Bruijn — Leiden University, Department of Clinical, Health, and Neuropsychology, Leiden Institute for Brain and Cognition, Leiden, The Netherlands

**Background:** Psychopathy is a severe personality disorder that has been linked to impaired behavioural adaptation during reinforcement learning. Recent electrophysiological studies have suggested that psychopathy is related to impairments in intentionally using information relevant for adapting behaviour, whereas these impairments remain absent for behaviour relying on automatic use of information. We sought to investigate whether previously found impairments in response reversal in individuals with psychopathy also follow this dichotomy. We expected response reversal to be intact when the automatic use of information was facilitated. In contrast, we expected impaired response reversal when intentional use of information was required. **Methods:** We included offenders with psychopathy and matched healthy controls in 2 experiments with a probabilistic cued go/no-go reaction time task. The task implicated the learning and reversal of 2 predictive contingencies. In experiment 1, participants were not informed about the inclusion of a learning component, thus making cue-dependent learning automatic/incidental. In experiment 2, the instructions required participants to actively monitor and learn predictive relationships, giving learning a controlled/intentional nature. **Results:** While there were no significant group differences in acquisition learning in either experiment, the results revealed impaired response reversal in offenders with psychopathy when controlled learning was facilitated. Interestingly, this impairment was absent when automatic learning was predominant. **Limitations:** Possible limitations are the use of a nonforensic control group and of self-report measures for drug use. **Conclusion:** Response reversal deficits in individuals with psychopathy are modulated by the context provided by the instructions, according to the distinction between automatic and controlled processing in these individuals.

## Introduction

Our ability to learn associations between events with rewarding or punishing outcomes plays an important role in guiding our behaviour.<sup>1</sup> Of equal importance is our capacity to successfully alter established associations to maximize performance when the environment requires us to change our behaviour. When an event that previously led to reward (event A) currently leads to punishment and when a previously punished event (event B) now yields a reward, the optimal behaviour would be to reverse the initial response ten-

dencies by starting to respond to event B instead of event A. This principle is a driving force behind a form of behavioural adaptation known as response reversal. Response reversal can be defined as a change of behaviour following a reversal of the previously established relationships between events and their reinforcing value.<sup>2</sup> Response reversal contributes to flexibility in behaviour, both under social and nonsocial circumstances.<sup>3</sup> The amygdala and orbitofrontal cortex (OFC) have been shown to play a crucial role in response reversal.<sup>3,4</sup> The amygdala has been linked to the establishment of stimulus–outcome associations,<sup>5</sup> whereas the OFC has been

**Correspondence to:** I.A. Brazil, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands; I.Brazil@donders.ru.nl

*J Psychiatry Neurosci* 2013;38(4):E13-20.

Submitted Aug. 2, 2012; Revised Jan. 17, 28, 2013; Accepted Feb. 4, 2013.

DOI: 10.1503/jpn.120152

implicated in behavioural adaptation based on contingency changes.<sup>4</sup>

Patients with lesions of the OFC have been found to show intact acquisition of associative relationships during initial learning (hereafter termed "acquisition"), but to display disturbed response reversal.<sup>5</sup> Interestingly, disturbed response reversal has also been demonstrated in individuals with psychopathy.<sup>6,7</sup> Psychopathy is regarded as a severe personality disorder typified by emotional abnormalities in combination with severe antisociality.<sup>8</sup> There is growing evidence that cognitive deficiencies observed in psychopathy include abnormalities in emotional processing,<sup>9,10</sup> modulation of attention,<sup>11,12</sup> aspects of action monitoring<sup>13-15</sup> and, importantly, associative learning.<sup>6,7,16-18</sup> Behavioural results on associative learning in individuals with psychopathy show deficiencies in tracking stimulus–outcome contingencies and in subsequently altering behaviour in the face of changes in these contingencies.<sup>6,7</sup> In the study by Budhani and colleagues,<sup>7</sup> individuals with psychopathy showed normal acquisition but impaired response reversal in a task in which stimulus–outcome contingencies varied according to predetermined probabilities of gaining reward/punishment (i.e., probabilistic reinforcement learning). These results are in line with the predictions made by the Integrated Emotion Systems model of psychopathy.<sup>19</sup> This account was developed from a neurobiological perspective and proposes that cognitive, affective and behavioural abnormalities in individuals with psychopathy are due to deficiencies in a cortical network involving the amygdala and OFC. The model predicts that individuals with psychopathy should not display impaired acquisition of initial stimulus–response associations, as acquisition is not reliant on intact OFC and amygdala functioning. However, these individuals should show deficient response reversal, during which the integrity of OFC and amygdala functioning is crucial for modifying previously established stimulus–response relationships based on information conveyed by the outcomes.<sup>7</sup>

However, there are contradicting indications that learning deficiencies can also become evident during initial acquisition learning. Von Borries and colleagues<sup>18</sup> studied probabilistic reinforcement learning in psychopathy using event-related potentials (ERPs) and behavioural measures in a different paradigm. Participants were explicitly instructed to monitor and learn probabilistic associations through trial and error. This study did not include a reversal phase and revealed that individuals with psychopathy can also exhibit deficiencies during acquisition. Moreover, the electrophysiological results indicated that participants with psychopathy showed a specific deficiency in using information provided through negative feedback to learn and adapt their behaviour. This is consistent with other results suggesting impairments in the intentional use of available information to adapt behaviour in individuals with psychopathy. Our group<sup>14</sup> has shown that both behavioural and electrophysiological correlates of automatic (unconscious) processing of errors are intact in individuals with psychopathy, whereas later stages involved in controlled (conscious) processing of errors and behavioural adaptation are compromised.

Combined, the previous findings from our laboratory<sup>14,18</sup> suggest that in individuals with psychopathy, automatic adaptation of behaviour is unaffected, but impairments are present when adaptation relies on intentional use of available information. From this perspective, it can be hypothesized that response reversal is compromised in individuals with psychopathy specifically when instructions provide a context promoting controlled behavioural adaptation. This prediction is also supported by the observation that the participants with OFC lesions in the study by Rolls and colleagues<sup>20</sup> were aware of (and could verbalize) the fact that the contingencies had changed, but were still unable to execute response reversal. Also, in the studies by von Borries and colleagues<sup>18</sup> and Budhani and colleagues,<sup>7</sup> participants were aware that the goal of the task was to learn based on reinforcement. A second prediction offered by the distinction between automatic and controlled behavioural adaptation in individuals with psychopathy is that response reversal should be intact when automatic learning is predominant. It is important to note that our use of automatic learning refers to an implicit learning mechanism that does not rely on awareness of what is being learned<sup>21</sup> occurring incidentally. To our knowledge, there has been no previous exploration of response reversal in circumstances promoting automatic learning.

The aim of the present study was to explore the effect of learning context (automatic/incidental v. controlled/intentional) on response reversal in individuals with psychopathy by manipulating task awareness through the instructions given in 2 separate experiments. In experiment 1, the instructions facilitated automatic learning during a probabilistic cued go/no-go task.<sup>22</sup> Participants were instructed to react as quickly as possible whenever the go stimulus appeared. Importantly, they were not made aware that the task contained predictive relationships. We expected reversal learning to be intact in individuals with psychopathy compared with healthy controls under these conditions. In the second experiment, we used the same task, but the instructions were altered to make participants aware of the predictive relationships and the ability to learn from them, thereby facilitating intentional learning. Participants were instructed to actively monitor for predictive relationships and to respond appropriately to receive reward. We expected response reversal to be compromised in individuals with psychopathy under these circumstances.

## Methods

### *Participants and procedure*

Participants in the group with psychopathy were recruited from the in- and outpatient population of the Pompestichting Forensic Psychiatric Institute in Nijmegen, The Netherlands. This is a treatment facility for people who have committed offences partly owing to a DSM-IV Axis I and/or Axis II disorder. For inpatients, life in the clinic is designed to resemble everyday life outside of detention as much as possible. They are required to follow treatment, engage in educational activities, work, exercise and socialize.

Suitable candidates were initially selected based on available information about clinical status and history. Subsequently, we used the Dutch version of the MINI Psychiatric Interview<sup>23</sup> and the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)<sup>24</sup> to screen candidates who were willing to participate. We excluded individuals with Axis I disorders (i.e., bipolar disorder, depressive disorder, schizoaffective disorder, schizophrenia, schizophreniform disorder, delusional and other psychotic disorders, schizotypal disorder, schizoid personality disorder), as well as first-degree relatives with DSM-IV Axis I schizophrenia or schizophreniform disorder/somatic disorders. We also excluded those chronically using intoxicating substances and those using psychotropic medication at the time of testing.<sup>12</sup> Individuals with psychopathy and/or antisocial personality disorder were precluded from participating in the healthy control groups. All assessments were conducted by trained psychologists based on interviews and on available information from each patient's clinical files.

We estimated each participant's IQ using the Dutch version of the National Adult Reading Test,<sup>25,26</sup> and psychopathy was scored using the Hare Psychopathy Checklist — Revised (PCL-R),<sup>27</sup> an instrument yielding a psychopathy score based on file information and a semistructured interview. As is customary in Europe, the cut-off score was a PCL-R score of 26 or greater.<sup>12,28</sup> In each experiment, the group with psychopathy was matched for age and IQ with a community sample of healthy male volunteers with a similar level of intelligence and without a history of psychiatric disorders and criminal records whom we recruited through advertisements. As the controls did not have criminal records, no PCL-R scores were assessed in this population.

The experimental protocol was approved by the local ethics committee of the Radboud University Nijmegen. All participants received written information about the experiment, gave written informed consent and received financial compensation.

### Task and design

#### Experiment 1: automatic learning

This experiment consisted of an adapted version of a probabilistic cued go/no-go reaction time (RT) task developed by Fillmore and Rush.<sup>22</sup> Participants were seated in front of a 100 Hz computer screen on which events were presented against a white background. A trial started with the presentation of a fixation cross in the middle of the screen for 800 ms, followed by a blank screen for 500 ms, a cue and either a go or a no-go stimulus (Fig. 1). The cue consisted of a white rectangle (65 mm × 20 mm) with black borders presented in either a horizontal (flat cue) or a vertical orientation (tall cue) and was followed by a go or a no-go stimulus. Five different stimulus onset asynchronies (SOAs; 100, 200, 300, 400 and 500 ms) were used for the presentation of the cues to promote allocation of attention to the cues and to prevent anticipation effects for the onset of the imperative stimuli. The latter consisted of a green (go) and a blue (no-go) rectangle (65 mm × 25 mm) displayed in the centre of the screen for 1000 ms after

the cue signal. The orientation of the cue indicated whether a go or a no-go stimulus was more likely to appear.

We instructed participants to press a response button as quickly as possible when a go stimulus was presented and to suppress their response when a no-go stimulus was displayed. Positive or negative feedback was provided on all go trials but only on incorrect presses on no-go trials. More specifically, the RT relative to the onset of the go stimulus was displayed after a correct button press, which functioned as positive feedback. However, when a button press was made on a no-go trial, or when the participant did not press the response button on a go trial, the word "incorrect" appeared on the screen, representing negative feedback. To encourage fast responding, participants were told that they would receive 5 points if the RT was equal to or below 300 ms, receive no points if the RT was longer than 300 ms and lose 5 points if the word "incorrect" appeared on the screen. No feedback or reward was given for not responding on no-go trials.

The task consisted of 2 phases: an acquisition phase and a reversal phase, each consisting of 500 trials divided into 5 blocks of equal size. The whole experiment lasted about 50 minutes. During the acquisition phase, the go stimulus was preceded by the flat cue in 80% of the trials and by the tall cue in 20% of the trials. The no-go stimulus had a reversed cue mapping: 80% tall and 20% flat. Thus, the orientation of the white cue was linked to the likelihood of a go or a no-go stimulus being presented. Each SOA, stimulus, cue and cue-stimulus combination appeared an equal number of times during each block (for more details, see the study by Fillmore and Rush<sup>22</sup>). Participants were informed that a cue would appear to signal that a stimulus was coming and that they did not have to react to the cues. During acquisition, participants were expected to learn the probabilistic associations between the orientation of the cues and the type of stimulus that followed without explicit information about the true function of the cues. Thus, participants were not told that there was a predictive relationship between the cues and the stimuli, priming context-facilitated automatic learning of the predictive associations. After reversal occurred, the mappings between the cues and the stimuli were reversed without informing the participants. The acquisition learning effects were expected to become evident by a decrease in RTs

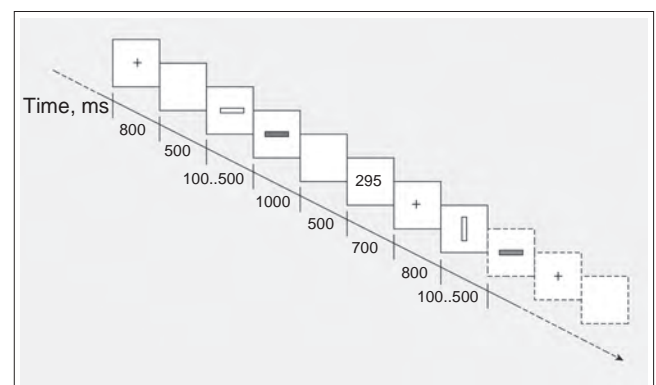


Fig. 1: Sequence of events and their timing during the experiments.

associated with the flat cue, predictive of the occurrence of the go stimulus on 80% of the trials during acquisition. In contrast, the RTs on go trials following the less predictive tall cue were expected to increase. After reversal, the opposite pattern was expected: RTs after the flat cue that was previously predictive of the go stimulus on 80% of the trials were expected to increase (as this cue no longer predicted the go stimulus), and RTs after the tall cue that previously signaled the no-go stimulus on 80% of the trials were expected to decrease. A short resting period was offered between blocks and participants were not informed about predictive relationships between the cues and the stimuli after the task.

### Experiment 2: controlled learning

We repeated the same experiment about 2.5–3 years later. However, in the second experiment, the colour scheme of the stimuli was adapted, and participants received different instructions. We told participants to react if a red rectangle (go stimulus) was presented and to withhold their response if a yellow rectangle (no-go stimulus) appeared after a cue. The key difference from experiment 1 was that the participants were explicitly instructed that there was a predictive relationship between the cues and the rectangles. They knew in advance that each of the 2 cues was more often followed by either the go or the no-go stimulus, but that these predictive relationships were probabilistic. These instructions increased awareness of the cue–stimulus contingencies and also resemble those used in the studies conducted by von Borries and colleagues<sup>18</sup> and Budhani and colleagues.<sup>7</sup> However, in our study, participants were not explicitly informed that the predictive relationships could change during the task to prevent the task from becoming too easy, thereby reducing the risk of floor effects.

### Analyses

We analyzed RTs on correct go trials using repeated-measures general linear models (GLMs) for each experiment separately. Acquisition learning was expected to become evident by a cue-dependent change in RTs between the start of the task (start acquisition) and the end of acquisition learning (end acquisition). Conversely, reversal learning was expected to become evident by a change in RTs between the end of acquisition and the end of the reversal phase (end reversal). Therefore, the GLMs included block (start acquisition, end acquisition, end reversal) and cue (tall, flat) as within-subject factors and group (psychopathy, control) as a between-subject factor. The  $\alpha$  level was set at 0.05. The levels of the block factor corresponded with the average RT of the first and last 100 trials of the acquisition phase and the last 100 trials of the reversal phase, respectively. The RTs between 100 ms and 2 standard deviations (SDs) above the mean overall RT were calculated for each participant, and RTs outside this range were excluded from the analyses to reduce the impact of outliers.<sup>29</sup> About 1.1% of the trials were excluded in experiments 1 and 2, respectively. We expected group differences in cue-dependent learning to become evident by significant group  $\times$  block  $\times$  cue interactions.<sup>22</sup> This interaction would indicate that the mean RTs differ signifi-

cantly between the groups as a function of the cue factor in 1 or more blocks. Post-hoc analyses were conducted with Bonferroni-corrected paired-samples *t* tests. Effect sizes ( $\eta^2$ ) were calculated for each of the within- and between-subject effects by dividing the corresponding sum of squares by the total sum of squares.<sup>30</sup> Greenhouse–Geisser-corrected *p* values are reported where appropriate.

## Results

### Participants

The group with psychopathy consisted of 18 participants in experiment 1 and 21 participants in experiment 2. These sample sizes are comparable to those used in related studies.<sup>6,7</sup> Nine individuals with psychopathy participated in both experiments. We initially included 20 participants in the healthy control group in experiment 1; however, 2 were then excluded because they did not completely understand or correctly follow the instructions. Therefore, 18 healthy controls participated in experiment 1 and 21 were included in experiment 2. One control participated in both experiments. Table 1 lists the demographic and clinical characteristics of participants in the psychopathy and control groups for each experiment. None of the participants reported being colour blind, and they all had normal or corrected-to-normal vision.

Accuracy data indicated that each group achieved near-perfect levels of accuracy in the experiments. The accuracy

**Table 1: Demographic and clinical characteristics of the psychopathy and control groups for each experiment**

Characteristic	Psychopathy		Control	
	Experiment 1 <i>n</i> = 18	Experiment 2 <i>n</i> = 21	Experiment 1 <i>n</i> = 18	Experiment 2 <i>n</i> = 19
General, mean (SD) [range]				
Age, yr	39 (7.9)	43 (7.0)	38 (8.4)	36 (8.8)
IQ	98 (9.1)	97 (10)	102 (6.8)	102 (5.6)
PCL-R score	31 (3.5) [26–38]	31 (3.4) <sup>§</sup> [26–38]	—	—
No. comorbid disorders*				
Borderline	0	1	0	0
Antisocial disorder	12	15	0	0
Narcissism	5	5	0	0
Self-reported drug use, % <sup>†</sup>				
Alcohol use <sup>‡</sup>	0	0	0	32 <sup>¶</sup>
Cannabis	44	43	17	21
Cocaine	28	33	0	5
Amphetamine	22	24	0	0
Methamphetamine	17	24	0	5
Opiate/morphine/ heroin	6	14	0	0

PCL-R = Hare Psychopathy Checklist — Revised;<sup>27</sup> SD = standard deviation.

\*Only current disorders identified in the populations are reported.

<sup>†</sup>Differences in percentage reported drug use were examined with 2-proportion *z* tests. For each group, the proportions for experiment 1 were tested against those of experiment 2.

<sup>‡</sup>Alcohol use was defined as consumption of more than 2 glasses of alcohol a day on average.

<sup>§</sup>The exact PCL-R scores of 2 participants scoring above threshold were not accessible.

<sup>¶</sup>Significant difference.

level in the psychopathy group was 97.7% in the acquisition phase and 97.2% in the reversal phase in experiment 1, and 99.8% and 99.5% in experiment 2. The control group had 97.3% and 96.9% accuracy in the acquisition and reversal phases, respectively, and 99.9% and 99.7% in experiment 2.

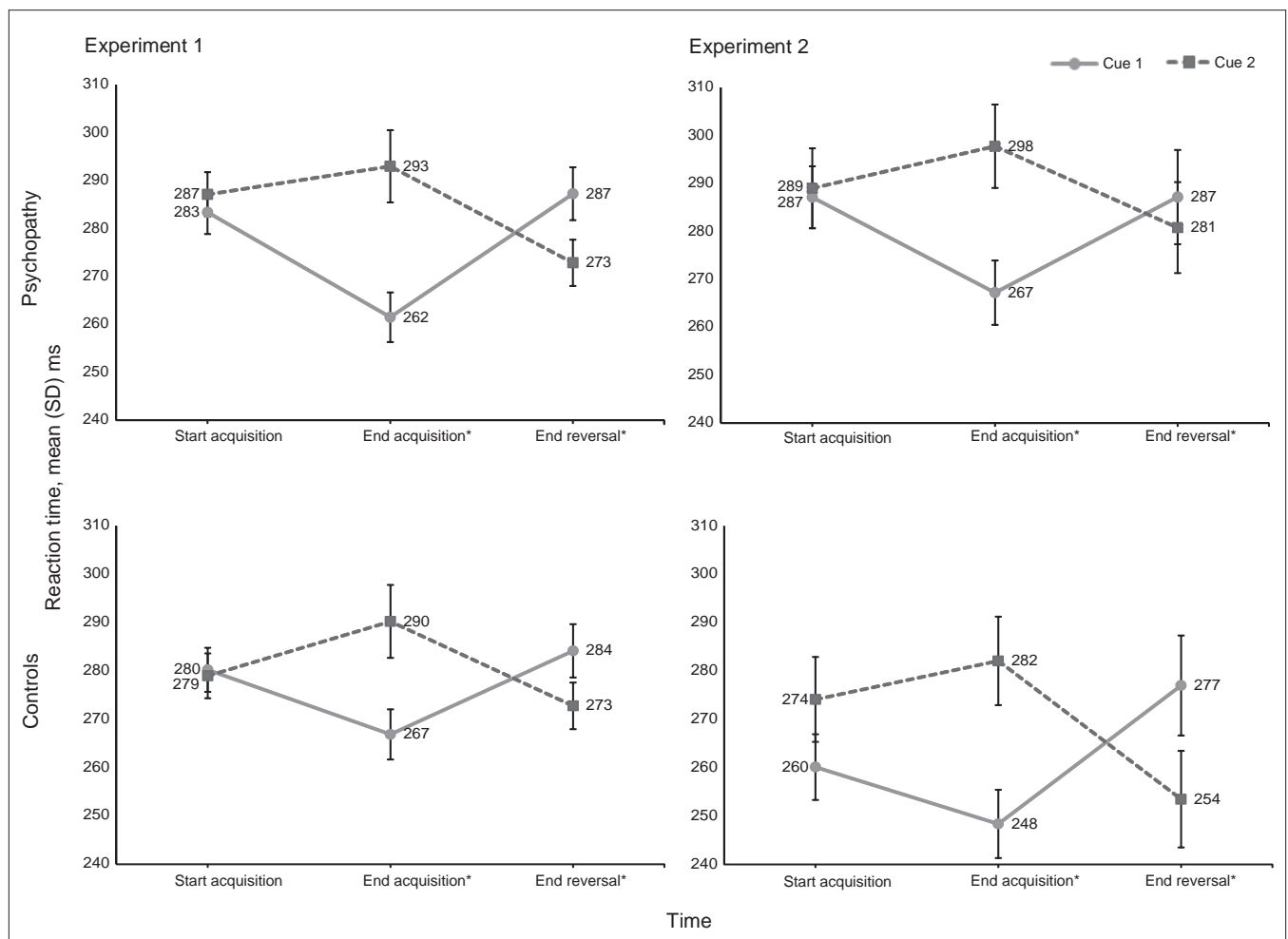
### Experiment 1: automatic learning

The left panel of Figure 2 displays the mean RTs for each of the groups, cues and trial blocks. As can be seen, there were no large differences between groups; both showed large RT differences between cues on trial blocks 2 and 3, but not block 1. Statistical analyses revealed no significant main effect for block ( $F_{2,68} = 1.40, p = 0.25, \eta^2 = 0.012$ ), cue ( $F_{1,34} = 3.69, p = 0.06, \eta^2 = 0.023$ ) or group ( $F_{1,34} = 0.116, p = 0.74, \eta^2 = 0.003$ ). A significant block  $\times$  cue interaction was indicative of successful learning in general ( $F_{2,68} = 34.6, p < 0.001, \eta^2 = 0.229$ ). This interaction reflected significantly faster responding after the more predictive cue than after the less predictive cue on the final block of the acquisition and reversal phases (264 v. 292 ms;  $t_{35} = -4.89, p < 0.001$ ) and (286 v. 273 ms;  $t_{35} = 5.03,$

$p < 0.001$ ), respectively, but not on the very first acquisition trial block (282 v. 283 ms;  $t_{35} = -0.436, p = 0.67$ ). The nonsignificant block  $\times$  group ( $F_{2,68} = 0.793, p = 0.46, \eta^2 = 0.007$ ), cue  $\times$  group ( $F_{1,34} = 0.377, p = 0.54, \eta^2 = 0.002$ ) and block  $\times$  cue  $\times$  group interactions ( $F_{2,68} = 0.681, p = 0.51, \eta^2 = 0.005$ ) indicated comparable performance between the groups.

### Experiment 2: controlled learning

The right side of Figure 2 shows the groups' mean RTs for each cue and trial block. The major difference between groups concerns responding during the reversal phase. On block 3, the control participants clearly displayed a difference in RTs between the 2 cues, whereas the individuals with psychopathy did not. There was a significant effect for cue ( $F_{1,38} = 11.1, p = 0.002, \eta^2 = 0.039$ ), reflecting higher RTs to the tall cue. The main effect for block ( $F_{2,76} = 0.565, p = 0.53, \eta^2 = 0.006$ ) and group ( $F_{1,38} = 3.20, p < 0.001, \eta^2 = 0.078$ ) did not reach significance. A significant block  $\times$  cue interaction was present ( $F_{2,76} = 39.2, p < 0.001, \eta^2 = 0.207$ ), whereas the block  $\times$  group ( $F_{2,76} = 0.124, p = 0.88, \eta^2 = 0.001$ ) and cue  $\times$  group



**Fig. 2:** Performance for each group in each experiment during the start of acquisition, end of acquisition and end of reversal phases. Mean group reaction times are reported with the error bars indicating their corresponding standard error.

( $F_{1,38} = 0.015$ ,  $p = 0.90$ ,  $\eta^2 < 0.001$ ) interactions were not significant. More importantly, the block  $\times$  cue  $\times$  group interaction was significant ( $F_{2,76} = 3.97$ ,  $p = 0.023$ ,  $\eta^2 = 0.021$ ). Bonferroni-corrected paired-samples  $t$  tests comparing RTs between the 2 cues in each block and for each group showed that in the psychopathy group, the difference in RT between the 2 types of cues was only significant at the end of the acquisition phase ( $t_{20} = -4.94$ ,  $p_{\text{bonf}} < 0.001$ ). However, in the control group, the difference between cues was significant both at the end of acquisition ( $t_{18} = -5.26$ ,  $p_{\text{bonf}} < 0.001$ ) and at the end of the reversal phase ( $t_{18} = 5.39$ ,  $p_{\text{bonf}} < 0.001$ ).

## Discussion

The aim of the present study was to investigate response reversal under automatic and controlled learning conditions in individuals with psychopathy. The findings point out that the presence of a response reversal deficit in individuals with psychopathy is modulated by the context in which learning occurs and are in line with the notion that some of the disturbances seen in individuals with psychopathy are related to a reduced capacity to intentionally use and manipulate information to adapt behaviour. More specifically, the results show that the deficiency is not present when response reversal occurs in a context in which automatic learning is predominant (experiment 1). Interestingly, however, abnormal response reversal was found in individuals with psychopathy when participants were instructed to actively monitor and manipulate associative relationships to perform successfully (experiment 2). The data also suggest that response reversal is not completely impaired in individuals with psychopathy, but that there is slower adaption of behaviour when the active use of information is required. The latter corroborates previous results showing delays in learning in individuals with psychopathy.<sup>18</sup> These findings have important implications for current accounts of disturbed learning in these individuals.

The Integrated Emotion Systems (IES) model cannot accommodate the results from experiment 1. That is, when the true nature of the task is made less salient by omitting any reference to the predictive relationship between stimuli, individuals with psychopathy are very capable of performing response reversals successfully. However, the current formulation of the IES model postulates general response reversal deficits in those individuals, irrespective of learning context and level of awareness. Therefore, one would not expect abnormal response reversal to be limited to explicit learning conditions.

The distinction between automatic and controlled cognitive processing in individuals with psychopathy offers novel predictions. On a neurocognitive level, our results can be explained by considering the role of the prefrontal cortex (PFC) in adapting behaviour. The PFC has been proposed to selectively bias cognitive processing to focus attention on relevant information while de-emphasizing competing information,<sup>31</sup> to use the relevant information to guide goal-directed behaviour.<sup>32</sup> Cohen and colleagues<sup>32</sup> originally classified this mechanism as a driving force behind cognitive control, a term now used to describe several types of cognitive functions. From this perspective, the impaired response reversal in experi-

ment 2 indicates that by making the cues more salient, we tapped into a deficiency in properly using information provided by both cues and imperative stimuli to guide behaviour in the psychopathy group, resulting in hampered response reversal. This interpretation is consistent with previous findings pointing out that inmates with psychopathy showed impaired performance under dual-task conditions when equal priority was given to both tasks<sup>33</sup> and with more recent results relating psychopathy to impairments in cognitive control.<sup>34</sup> Moreover, considering cognitive control deficits also offers an explanation for the unaffected response reversal found in experiment 1. Automatic processing is assumed not to rely on the integrity of the (prefrontal) brain system regulating cognitive control. During automatic learning in experiment 1, this system was bypassed, resulting in normal performance in the psychopathy group. Thus, the predictions offered by the dichotomy between automatic and controlled processing on a cognitive level converge with those made based on the biasing mechanism being referred to as cognitive control. However, there are also studies that used different behavioural measures and reported unaffected cognitive control in individuals with psychopathy.<sup>35,36</sup> These studies point out that not all aspects of cognitive control are compromised or that different behavioural indexes vary in sensitivity and suitability depending on the function being assessed. For instance, Blair and colleagues<sup>35</sup> used a series of neuropsychological tests, each known to be sensitive to different executive functions, in different cortical areas. Their results indicated that the psychopathy group did not show deficiencies on behavioural measures preferentially sensitive to functions of the anterior cingulate and dorsolateral prefrontal cortex, whereas measures quantifying functions of the OFC did show deficiencies.

## Limitations

One limitation of the present study is that it could be argued that nonpsychopathic offenders would have been a more valid comparison group. However, previous studies have shown that response reversal deficits are characteristic of psychopathy relative to general antisociality in adult offender samples<sup>6,7</sup> and in children with high levels of psychopathic traits.<sup>37</sup> Accordingly, it was not our primary intention to re-establish the link between response reversal deficits and psychopathy relative to generic antisociality. Furthermore, the absence of a difference in experiment 1 between the 2 groups is especially noteworthy given that the comparison involved more "contrasting" populations (psychopathic v. healthy individuals) than those in many other studies with similar sample sizes (psychopathic v. nonpsychopathic offenders). If anything, this should have increased the chance of detecting group differences in our experiment. Still, it will be beneficial to replicate these results in a study that includes a group of nonpsychopathic offenders. Reversal deficits have also been reported in (poly)drug users,<sup>22</sup> and psychopathy has been linked to higher rates of (poly)drug use.<sup>38</sup> One could argue that the reversal deficit seen in experiment 2 could be attributed to a history of drug use in the psychopathy group.

However, (self-reported) rates of drug use in the patient samples did not differ between the 2 experiments (Table 1), whereas the outcome did, implicating that drug use cannot be responsible for the difference. Finally, another argument is that the use of a different measure of response reversal (RTs v. amount of reversal errors) compared with those used in previous studies of adult psychopathy<sup>6,7</sup> reduces the comparability among studies. However, if disturbed response reversal is an essential aspect of psychopathy, it should also be present when a different method of assessing the same cognitive mechanism is used, thus providing additional support for the robustness of this cognitive deficiency. Further studies are needed to address the exact impact of manipulating the saliency of different pieces of information using these types of paradigms.

## Conclusion

The present study shows that deficient response reversal in individuals with psychopathy can be modulated by altering the nature of the learning context. The findings support the notion that some aspects of automatic processing of behaviour are intact in individuals with psychopathy, but that disturbances arise when information processing reaches controlled stages of processing and has to be used to guide goal-directed behaviour.<sup>14</sup> This view suggests that abnormal processing of information relevant for appropriate (re)adjustment of current behaviour becomes apparent when individuals with psychopathy have to actively monitor and manipulate information. These results also highlight the importance of considering the way information is offered to offenders with psychopathy during therapeutic interventions in forensic psychiatry settings. Using approaches that rely on automatic learning mechanisms might be an effective way to modify rigid and disruptive behaviour.

**Acknowledgements:** I.A. Brazil, E.R.A. de Bruijn and R.P.C. Kessels were supported by Mosaic (240-00-244), VENI (451-07-022) and VIDJ (452-08-005) grants, respectively, from the Netherlands Organization for Scientific Research (NWO).

**Competing interests:** As above for I.A. Brazil, E.R.A. de Bruijn and R.P.C. Kessels. Otherwise, none declared.

**Contributors:** I.A. Brazil, B.H. Bulten, R.J. Verkes and E.R.A. de Bruijn designed the study. I.A. Brazil and I. Scheper acquired and analyzed the data, which J.H.R. Maes, R.P.C. Kessels and E.R.A. de Bruijn also analyzed. I.A. Brazil, J.H.R. Maes, I. Scheper and E.R.A. de Bruijn wrote the article. I.A. Brazil, J.H.R. Maes, B.H. Bulten, R.P.C. Kessels, R.J. Verkes and E.R.A. de Bruijn reviewed the article. All authors approved its publication.

## References

1. Rushworth MFS, Behrens TEJ. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* 2008;11:389-97.
2. Cools R, Clark L, Owen AM, et al. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 2002;22:4563-7.
3. Kringelbach ML, Rolls ET. Neural correlates of rapid reversal learning in a simple model of human social interaction. *Neuroimage* 2003;20:1371-83.
4. Rolls ET. The orbitofrontal cortex and reward. *Cereb Cortex* 2000;10:284-94.
5. Davis M, Whalen PJ. The amygdala: vigilance and emotion. *Mol Psychiatry* 2001;6:13-34.
6. Mitchell DGV, Colledge E, Leonard A, et al. Risky decisions and response reversal: Is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? *Neuropsychologia* 2002;40:2013-22.
7. Budhani S, Richell RA, Blair RJR. Impaired reversal but intact acquisition: probabilistic response reversal deficits in adult individuals with psychopathy. *J Abnorm Psychol* 2006;115:552-8.
8. Hare RD, Hart SD, Harpur TJ. Psychopathy and the DSM-IV criteria for antisocial personality disorder. *J Abnorm Psychol* 1991;100:391-8.
9. Patrick CJ. Emotion and psychopathy: startling new insights. *Psychophysiology* 1994;31:319-30.
10. Müller JL, Sommer M, Wagner V, et al. Abnormalities in emotion processing within cortical and subcortical regions in criminal psychopaths: evidence from a functional magnetic resonance imaging study using pictures with emotional content. *Biol Psychiatry* 2003;54:152-62.
11. Newman JP, Curtin JJ, Bertsch JD, et al. Attention moderates the fearlessness of psychopathic offenders. *Biol Psychiatry* 2010;67:66-70.
12. Brazil IA, Verkes RJ, Brouns BHJ, et al. Differentiating psychopathy from general antisociality using the P3 as a psychophysiological correlate of attentional allocation. *PLoS ONE* 2012;7:e50339.
13. Munro GES, Dywan J, Harris GT, et al. ERN varies with degree of psychopathy in an emotion discrimination task. *Biol Psychol* 2007;76:31-42.
14. Brazil IA, De Bruijn ER, Bulten BH, et al. Early and late components of error monitoring in violent offenders with psychopathy. *Biol Psychiatry* 2009;65:137-43.
15. Brazil IA, Mars RB, Bulten BH, et al. A neurophysiological dissociation between monitoring one's own and others' actions in psychopathy. *Biol Psychiatry* 2011;69:693-9.
16. Newman JP, Kosson DS. Passive avoidance learning in psychopathic and nonpsychopathic offenders. *J Abnorm Psychol* 1986;95:252-6.
17. Blair RJR, Mitchell DGV, Leonard A, et al. Passive avoidance learning in individuals with psychopathy: modulation by reward but not by punishment. *Pers Individ Dif* 2004;37:1179-92.
18. von Borries AKL, Brazil IA, Bulten BH, et al. Neural correlates of error-related learning deficits in individuals with psychopathy. *Psychol Med* 2010;40:1559-68.
19. Blair RJR. Applying a cognitive neuroscience perspective to the disorder of psychopathy. *Dev Psychopathol* 2005;17:865-91.
20. Rolls ET, Hornak J, Wade D, et al. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J Neurol Neurosurg Psychiatry* 1994;57:1518-24.
21. DeKeyser R. Implicit and explicit learning. In: Doughty CJ, Long MH, editors. *The handbook of second language acquisition*. Malden (MA): Blackwell Publishing; 2003. p. 312-48.



22. Fillmore MT, Rush CR. Polydrug abusers display impaired discrimination-reversal learning in a model of behavioural control. *J Psychopharmacol* 2006;20:24-32.
23. Van Vliet IM, Leroy H, Van Megen HJM. De MINI-Internationaal neuropsychiatrisch interview: een kort gestructureerd diagnostisch interview voor DSM-IV en ICD-10 psychiatrische stoornissen. Leiden: LUMC; 2000.
24. Weertman A, Arntz A, Kerkhofs MLM. Gestructureerd klinisch interview voor DSM-IV persoonlijkheidsstoornissen (SCID-II). Amsterdam: Swets; 1996.
25. Schmand B, Bakker D, Saan R, et al. The Dutch Reading Test for Adults: a measure of premorbid intelligence level. *Tijdschr Gerontol Geriatr* 1991;22:15-9.
26. Evers A, Egberink IJL, Braak MSL, et al. *COTAN documentatie*. Amsterdam: Boom test uitgevers; 2009-2012.
27. Hare RD. *Manual for the Revised Psychopathy Checklist*. 2nd ed. Toronto (ON): Multi-Health Systems; 2003.
28. Rasmussen K, Storsæter O, Levander S. Personality disorders, psychopathy, and crime in a Norwegian prison population. *Int J Law Psychiatry* 1999;22:91-7.
29. Barnett VL, Lewis T. *Outliers in statistical data*. New York: John Wiley & Sons; 1994
30. Tabachnick BG, Fidell LS. *Using multivariate statistics*. 4th ed. Needham Heights (MA): Allyn & Bacon; 2001
31. Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 1995;18:193-222.
32. Cohen JD, Braver TS, O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philos Trans R Soc Lond B Biol Sci* 1996;351:1515-27.
33. Kosson DS, Newman JP. Psychopathy and the allocation of attentional capacity in a divided-attention situation. *J Abnorm Psychol* 1986;95:257-63.
34. Zeier JD, Baskin-Sommers AR, Hiatt Racer KD, et al. Cognitive control deficits associated with antisocial personality disorder and psychopathy. *Personal Disord* 2012;3:283-93.
35. Blair KS, Newman C, Mitchell DG, et al. Differentiating among prefrontal substrates in psychopathy: neuropsychological test findings. *Neuropsychology* 2006;20:153-65.
36. Dvorak-Bertsch JD, Sadeh N, Glass SJ, et al. Stroop tasks associated with differential activation of anterior cingulate do not differentiate psychopathic and non-psychopathic offenders. *Pers Individ Dif* 2007;42:585-95.
37. Budhani S, Blair RJR. Response reversal and children with psychopathic tendencies: success is a function of salience of contingency change. *J Child Psychol Psychiatry* 2005;46:972-81.
38. Smith SS, Newman JP. Alcohol and drug abuse-dependence disorders in psychopathic and nonpsychopathic criminal offenders. *J Abnorm Psychol* 1990;99:430-9.