The following full text is a publisher's version.

For additional information about this publication click this link.  
http://hdl.handle.net/2066/116654

Please be advised that this information was generated on 2017-08-10 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

Introduction

Our ability to learn associations between events with rewarding or punishing outcomes plays an important role in guiding our behaviour. 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 tendencies 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. Response reversal contributes to flexibility in behaviour, both under social and nonsocial circumstances. The amygdala and orbitofrontal cortex (OFC) have been shown to play a crucial role in response reversal. The amygdala has been linked to the establishment of stimulus–outcome associations, whereas the OFC has been

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.
Implicated in behavioural adaptation based on contingency changes. They 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 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 and Budhani and colleagues, 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 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. 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\textsuperscript{23} and the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)\textsuperscript{24} 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.\textsuperscript{12} 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,\textsuperscript{25,26} and psychopathy was scored using the Hare Psychopathy Checklist — Revised (PCL-R),\textsuperscript{27} 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.\textsuperscript{12,28} In each group, the event 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.\textsuperscript{29} 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\textsuperscript{29}). 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.](Image)
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 resembled those used in the studies conducted by von Borries and colleagues18 and Budhani and colleagues.7 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 RTs between 100 ms and 2 standard deviations were excluded from the analyses. The RTs between 100 ms and 2 standard deviations (psychopathy, control) as a between-subject factor. 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 were excluded to prevent floor effects.

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.17 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>
<thead>
<tr>
<th>Characteristic</th>
<th>Psychopathy</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>n = 18</td>
<td>n = 19</td>
</tr>
<tr>
<td>Age, yr</td>
<td>39 (7.9)</td>
<td>36 (8.8)</td>
</tr>
<tr>
<td>IQ</td>
<td>98 (9.1)</td>
<td>102 (5.6)</td>
</tr>
<tr>
<td>PCL-R score</td>
<td>31 (3.5)</td>
<td>—</td>
</tr>
<tr>
<td>No. comorbid disorders*</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Borderline</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Antisocial disorder</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Narcissism</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Self-reported drug use, %†</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Alcohol use‡</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cannabis</td>
<td>44</td>
<td>42</td>
</tr>
<tr>
<td>Cocaine</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>Amphetamine</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>Methamphetamine</td>
<td>17</td>
<td>24</td>
</tr>
<tr>
<td>Opiate/morphine/heroin</td>
<td>6</td>
<td>14</td>
</tr>
</tbody>
</table>

PCL-R = Hare Psychopathy Checklist — Revised; SD = standard deviation. *Only current disorders identified in the populations are reported. †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. ‡Alcohol use was defined as consumption of more than 2 glasses of alcohol a day on average. §The exact PCL-R scores of 2 participants scoring above threshold were not accessible. ‡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 ($F_{1,34} = 0.793, p = 0.46, \eta^2 = 0.007$), 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
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 adaptation of behaviour when the active use of information is required. The latter corroborates previous results showing delays in learning in individuals with psychopathy. These findings have important implications for current accounts of disturbed learning in these individuals.

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 and in children with high levels of psychopathic traits. 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, and psychopathy has been linked to higher rates of (poly)drug use. 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 VIDI (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**


