Archival Report

Biological Psychiatry: CNNI

# Representational Uncertainty in the Brain During Threat Conditioning and the Link With Psychopathic Traits

Inti A. Brazil, Christoph D. Mathys, Arne Popma, Sylco S. Hoppenbrouwers, and Moran D. Cohn

## ABSTRACT

**BACKGROUND:** Psychopathy has repeatedly been linked to disturbed associative learning from aversive events (i.e., threat conditioning). Optimal threat conditioning requires the generation of internal representations of stimulus–outcome contingencies and the rate with which these may change. Because mental representations are imperfect, there will always be uncertainty about the accuracy of representations in the brain (i.e., representational uncertainty). However, it remains unclear 1) to what extent threat conditioning is susceptible to different types of uncertainty in representations about contingencies during the acquisition phase and 2) how representational uncertainty relates to psychopathic features.

**METHODS:** A computational model was applied to functional neuroimaging data to estimate uncertainty in representations of contingencies (CoUn) and the rate of change of contingencies (RUn), respectively, from brain activation during the acquisition phase of threat conditioning in 132 adolescents at risk of developing antisocial personality profiles. Next, the associations between these two types of representational uncertainty and psychopathy-related dimensions were examined.

**RESULTS:** The left and right amygdala activations were associated with CoUn, while the bilateral insula and the right amygdala were associated with RUn. Different patterns of relationships were found between psychopathic features and each type of uncertainty. Callous-unemotional traits and impulsive-irresponsible traits uniquely predicted increased CoUn, while only impulsive-irresponsible traits predicted increased RUn.

**CONCLUSIONS:** The findings suggest that 1) the insula and amygdala differ in how these regions are affected by different types of representational uncertainty during threat conditioning and 2) CoUn and RUn have different patterns of relationships with psychopathy-related dimensions.

*Keywords:* Amygdala, Computational modeling, Fear conditioning, Insula, Psychopathy, Representations, Threat conditioning, Uncertainty

http://dx.doi.org/10.1016/j.bpsc.2017.04.005

In a dynamic world, success relies heavily on our ability to adapt our behavior to avoid aversive outcomes. Threats have a large impact on the modulation of our behavior (1). One developmental condition that has repeatedly been linked to diminished learning from aversive experiences is psychopathy (2,3), which encompasses callous-unemotional traits (e.g., blunted affect, lack of empathy or remorse), a grandiose-manipulative interpersonal style (e.g., dishonesty, superficial charm, lying, manipulation of others), and impulsive-irresponsible behavioral tendencies (e.g., thrill seeking, lack of impulse control) (4). Adolescents with high levels of psychopathic traits are at increased risk of engaging in antisocial behavior (5) and may be more difficult to treat because of their more severe antisocial behavior and diminished treatment responsivity (6). The maladaptive behavior seen in psychopathy is thought to be strongly influenced by disturbed learning from aversive events, such as threats (i.e., threat conditioning) (7–9), which is reflected in abnormal physiological and brain responses in both psychopathic adults (7–9) and youths with severe antisocial personality profiles (10,11).

Threat conditioning is multifaceted, and learning relies on interacting cognitive computations, similar to other forms of associative learning (12,13). Learning which stimuli are threatening requires accurate representations of threat contingencies. To maintain accurate representations of contingencies, we need to update the representations continually after each observation, also taking into account that the learned contingencies may change. However, our observations are imperfect (14). Therefore, there will always be some uncertainty regarding the accuracy of the cognitive

© 2017 Society of Biological Psychiatry. Published by Elsevier Inc. All rights reserved. 

ISSN: 2451-9022 

representations (i.e., representational uncertainty)[1] that are generated based on these imperfect observations (15). Learning about contingencies involves generating representations based on our estimates of the likelihood that contingencies will change and of the rate at which these changes occur (13). Because these representations are imperfect, there will be uncertainty about the accuracy of the representation of contingency changes (i.e., contingency uncertainty [CoUn]) and of the rate at which changes occur (i.e., change rate uncertainty [RUn]) (16). Empirical evidence indicates that these two types of representational uncertainties are hierarchically related because the rate at which changes in contingencies are perceived to occur will influence our belief about the overall likelihood that the contingencies will change (13,15). Furthermore, a recent study showed that beliefs about each type of uncertainty play key roles in modulating physiological responses to threats (17). Importantly, effective learning requires cognitive uncertainty to be minimized (18). Therefore, it is likely that uncertainty in the representation of contingencies may also play a role in understanding the impairments in threat conditioning seen in psychopathy.

Threat conditioning in psychopathy has often been investigated in case-control studies in which groups of individuals scoring high on psychopathic features are compared with a low- scoring group [e.g., (7,11)]. In one of the few threat conditioning studies employing a dimensional approach to psychopathy, Cohn *et al.* (19) reported a positive relationship between blood oxygen level–dependent (BOLD) activation in the amygdala and the insula (during acquisition learning) and impulsive-irresponsible traits in at-risk adolescents but reported a negative correlation between callous-unemotional traits and BOLD activation in these regions. Other studies using physiological measures have found a similar negative relationship between threat conditioning and fearless dominance (a construct that overlaps with callous-unemotional traits) but no correlation with impulsive-antisociality in undergraduates (20) and reduced threat conditioning in adult psychopathic individuals scoring high on interpersonal-affective deficits (21). Taken together, the evidence points toward decreased threat conditioning in individuals high on callous-unemotional traits, while the findings are mixed for the impulsive-antisocial features. However, these prior studies have approached threat conditioning as a unitary form of learning without taking the multifaceted nature of associative learning into account. As a consequence, how psychopathy may be related to any of the various interacting cognitive computations that subserve acquisition learning during threat conditioning has been overlooked. Systematically studying the integrity of these computations is essential for pinpointing the deficiencies in the threat conditioning mechanism in psychopathy.

In this study, we examined whether the threat conditioning impairments seen during the acquisition phase in relation to psychopathy may be partly attributed to increased uncertainty in representations of contingencies and their rate of change. Importantly, because the uncertainty computations are latent, it is impossible to directly quantify them without employing computational modeling approaches. A computational model was applied to the large functional magnetic resonance imaging (fMRI) dataset collected by Cohn *et al.* (19) to quantify uncertainty in the representations of contingency change and the rate of change in target brain areas during threat conditioning. The right and left amygdala and insula, respectively, were chosen as regions of interest (ROIs) because 1) these areas consistently show activation during threat conditioning across studies (22), 2) these areas show relatively large responses to uncertainty about threat contingencies (23,24), 3) these areas are linked to learning impairments in psychopathy (25), and 4) we aimed to maintain comparability with the very few previous studies on psychopathy dimensions and threat conditioning. Importantly, because two types of representational uncertainties were quantified, it was unlikely that all ROIs showed an equal amount of CoUn and RUn. To take account of this, we subsequently used Bayesian structural equation modeling (BSEM) to examine which of the ROIs was more affected by each type of representational uncertainty and determined how psychopathy-related personality dimensions uniquely predicted CoUn and RUn.

Extant findings suggest that threat conditioning is reduced in individuals scoring high on callous-unemotional features, and we hypothesized that this learning impairment should be linked to increased representational uncertainty in these individuals. This is based on the premise that disturbed learning should be related to a failure in reducing uncertainty in the information processed in the brain [cf. (25)]. But given that we sought to obtain a higher level of precision by parsing representational uncertainty into CoUn and RUn for the first time, it is difficult to predict which type of uncertainty is related to the impairments seen in those with elevated levels of callous-unemotional traits. The findings for the impulsive-irresponsible dimension are mixed, but the hyperconditioning found by Cohn *et al.* (19) suggests that elevated impulsive-irresponsible traits should be linked to reduced representational uncertainty during threat conditioning.

## METHODS AND MATERIALS

### Participants and Assessment

Participants were recruited from a Dutch cohort of 364 adolescents who were first arrested by the Dutch police before 12 years of age, all of whom had participated in three previous waves of a longitudinal study (26). The mean age at study entrance was 10.9 years (SD = 1.4). All participants were assessed using the Dutch Youth Psychopathic Traits Inventory (YPI) (27), a valid and reliable 50-item self-report instrument developed to assess psychopathic traits in juvenile community samples (28). In the current study, internal consistency of the total score and its constituting dimensions was good to excellent; Cronbach's $\alpha$ was .93 for the total score and was

[1] Note that the operationalization of uncertainty used in this study differs from other definitions of the term. Neuroimaging studies have primarily focused on how uncertainty due to unsure outcomes is processed in the brain [e.g., (48,49)], whereas we examined the uncertainty about the accuracy of cognitive representations of change. In other words, our definition of uncertainty in this study refers to the inaccuracy in representations that are formed, which is driven by imperfect observations in addition to uncertain outcomes.

.86, .89, and .85 for the callous-unemotional, grandiose-manipulative, and impulsive-irresponsible scales, respectively.

For the current study, data were reanalyzed from a subsample characterized by a wide range of externalizing risk ($N$ = 150) with an even distribution of participants at low risk, medium risk, and high risk of antisocial behavior. Of this sample, 18 participants were excluded from analyses due to invalid MRI data ($n$ = 9), invalid task performance ($n$ = 6), drug use prior to scanning ($n$ = 2), or missing questionnaire data ($n$ = 1). Analyses were performed on the remaining 132 participants (mean age = 17.7 years, SD = 1.6) (see Supplement).

## Procedure

This study was approved by the Institutional Review Board of the VU University Medical Center Amsterdam. All participants and their parents/custodians (if participants' age was under 18 years) signed for informed consent. Participants underwent a neuroimaging protocol in a Philips 3T Intera MRI scanner (Philips Healthcare, Best, The Netherlands) at the VU University Medical Center Amsterdam. All participants were instructed to refrain from using alcohol, cannabis, or psychostimulant medication for at least 24 hours before the MRI scan.

## Threat Conditioning Task

A classical differential delay threat conditioning task was employed (7). Pictures of two neutral male faces served as conditioned stimuli (CS), one of which (chosen at random during each experiment) was consistently paired with an aversive electric unconditioned stimulus (US) at the end of a 10-second viewing period (CS+; 100% reinforcement) during the acquisition period, while the other picture (CS−) was never followed by a US. The acquisition period, which consisted of eight trials of each CS, was preceded by a habituation phase during which CSs were presented four times each for 3.5 seconds without a US and was followed by an extinction phase during which CSs were presented four times each for 7 seconds and were not followed by a US either.

## fMRI Protocol

T1-weighted anatomical scans (180 slices, 1 $mm^3$ voxels, field of view = 256 × 256 mm, repetition time = 9.0 ms, echo time = 3.5 ms) were acquired using an 8-channel SENSE head-coil (Philips Healthcare). Furthermore, 400 T2*-weighted axial echo-planar images were acquired during threat conditioning (38 slices, 3-mm thickness, 2.29 × 2.29 in-plane resolution, field of view = 220 × 220 mm, repetition time = 2300 ms, echo time = 30 ms).

## Statistical Analyses

fMRI data were processed using SPM8 (Wellcome Trust Centre for Neuroimaging, University College London, London, UK), including realignment, unwarping, slice-time correction, normalization to Montreal Neurological Institute space based on the segmented anatomical scan, and 8-mm full width at half maximum smoothing. First-level models included separate regressors for CS+/− and CS− acquisition, US, and rating blocks. During acquisition, the first 5 seconds of each trial was modeled separately from the remainder of the trial (5 seconds) to account for fast within-trial habituation of threat neurocircuitry, focusing analyses on the first epoch only during acquisition (19). Realignment parameters were also included in first-level models to account for movement effects. Next, average neural response estimates for each ROI were extracted using the MarsBaR toolbox for SPM (29). Following previous work (23,30), we focused analyses on the amygdala and the insula. The amygdala and insula were anatomically defined using the Automated Anatomic Labeling atlas (31).

## Computational Modeling

To model how activity in the ROIs responded to CS in the experiment, the activation trajectory of the principal eigenvariate was extracted for each ROI. This trajectory was then analyzed using the hierarchical Gaussian filter (HGF) (15). The HGF is a generic hierarchical Bayesian time series model that explains a time series as a sequence of noisy observations of a hidden state. Crucially, the HGF also models the dynamics of change (i.e., volatility) of the hidden state explicitly and gives one-step update equations for the evolving estimate of both the hidden state and its volatility. This evolving estimate is a probability distribution referred to by us as a belief. In the HGF, beliefs follow a normal distribution and are fully specified by their mean and variance. The mean of the belief represents the most probable value of the hidden state, and the variance represents the uncertainty.

More formally, the HGF consists of a hierarchy of Gaussian random walks where the variance of each walk is a function of the value at the next higher level. Because the variance of its walk determines a quantity's likelihood of changing, each value but the one at the base of the hierarchy represents the volatility of the next lower level. In the current study, we used a two-level HGF where the first level represents the BOLD activation in a given brain region and the second level is the volatility of that activation (Supplemental Figure S1). We applied the HGF to the measured BOLD activation (represented as $u$ in the HGF model graph of Supplemental Figure S1) to infer the latent true activation at the time of measurement and its volatility along with the posterior uncertainty about these quantities. This resulted in two estimated belief trajectories for each subject and ROI: that of the BOLD activation ($x_1$ in the HGF hierarchy; see Supplemental Figure S1) and that of its volatility ($x_2$ in the HGF hierarchy; see Supplemental Figure S1). Each of the two belief trajectories consisted of a mean trajectory ($\mu_1$ regarding $x_1$ and $\mu_2$ regarding $x_2$) and a variance (i.e., uncertainty) trajectory ($\sigma_1$ and $\sigma_2$). Crucially, this means that every observed update implies a certain level of uncertainty, and by observing the evolution of the BOLD signal in the ROIs during threat conditioning, we can infer the evolution of the uncertainty the brain has about the quantity encoded by the activity in each ROI. Estimation was performed using the HGF Toolbox version 3.0 (https://www.tnu.ethz.ch/en/software/tapas.html), and fits could be obtained for all subjects and ROIs. All HGF modeling was performed before the BSEM described below, and all hyperparameters of the HGF analysis (number of levels, noise level, and regularizing priors) were chosen only in terms of this analysis.

## Bayesian Structural Equation Modeling

To assess which brain areas were substantially affected by representational uncertainty during threat conditioning in a

data-driven fashion, we employed BSEM to determine 1) which ROIs were involved in coding CoUn and RUn and 2) which psychopathy-related dimensions predicted each type of representational uncertainty. First, mean uncertainty estimates for learning were created by averaging and subtracting the uncertainty trajectories of CS− trials from those of the CS+ trials in each ROI, yielding four average scores representing the additional uncertainty found in the CS+ condition relative to the CS− condition for CoUn and RUn, respectively. In particular, the CoUn score was the mean of the uncertainty (i.e., variance $\sigma_1$) at the bottom level of the HGF hierarchy during the CS+ condition minus the mean of the same uncertainty $\sigma_1$ during the CS− condition. Correspondingly, the RUn score was the mean of the uncertainty (i.e., variance $\sigma_2$) at the second level of the HGF hierarchy during the CS+ condition minus the mean of the same uncertainty $\sigma_2$ during the CS− condition. Note that this subtraction method is the common approach used to obtain BOLD responses reflecting conditioning but that we used only the uncertainty estimate derived from the BOLD signal instead of the raw BOLD signal. Next, a BSEM was built where the four average CoUn estimates were loaded on a latent factor, which was regressed on the scores of the callous-unemotional, grandiose-manipulative, and impulsive-irresponsible scales of the YPI. The same procedure was followed for the estimates of RUn.

The analyses were conducted in Mplus version 7.4 (32) using a Bayesian estimator (PX1) with five Markov chain Monte Carlo chains and 100,000 iterations. The first half of the iterations was discarded (i.e., burn-in trials), and model fit was determined using different indexes for Bayesian testing: 1) a chi-square test for posterior predictive checking and 2) the posterior predictive *p* value (PPP value). Convergence of the Markov chain Monte Carlo chains was established with Gelman–Rubin's potential scale reduction factor (33). In general, a good fit is indicated by a 95% credibility interval (CI) for the chi-square posterior predictive check that includes the value 0; the PPP value should approach the value 0.50, and convergence is achieved when the potential scale reduction is below 1.05 (32). Significance of the regression weights was determined based on the 95% CIs of the Bayesian posterior distributions, and variables with 95% CIs not containing the value 0 were considered as significant.

## RESULTS

After creating the mean difference scores between the CS+ and CS− conditions for CoUn and RUn, respectively, the mean difference scores were used as outcome variables in the BSEMs. The first BSEM included each subject's mean CoUn estimates in the right and left amygdala and insula, respectively, loaded on a latent factor that was in turn regressed on the YPI scales. This model had very good fit (95% CI −19.22 to 25.35, PPP = .38) and indicated that the left and right amygdala loaded significantly on the latent factor capturing CoUn and that the callous-unemotional (β = .45, 95% CI 0.21 to 0.72) and impulsive-irresponsible (β = .27, 95% CI 0.05 to 0.51) scales were significant positive predictors of the latent factor (Figure 1). The grandiose-manipulative scale was not a significant predictor (β = .10, 95% CI −0.15 to 0.35). This procedure was repeated for change rate uncertainty in the four ROIs loaded on a latent factor measuring RUn that was regressed on the three YPI scales. This model also had a very good fit (95% CI −27.52 to 17.17, PPP = .67). The results showed that the left and right insula and the right amygdala loaded significantly on the latent factor and that the impulsive-irresponsible scale predicted the latent variable for RUn (β = .29, 95% CI 0.07 to 0.49) (Figure 2). The callous-unemotional (β = −.02, 95% CI −0.25 to 0.21) and grandiose-manipulative (β = −.12, 95% CI −0.35 to 0.12) scales were not significant predictors.

## DISCUSSION

This study is the first to provide direct quantifications of different types of representational uncertainty in the brain during threat conditioning in a large sample of adolescents at risk of developing persistent antisocial behavior. The findings show that a significant level of CoUn can be found in the left and right amygdala during threat conditioning, while the right amygdala and the left and right insula are more responsive to RUn. The mapping of these brain responses onto different

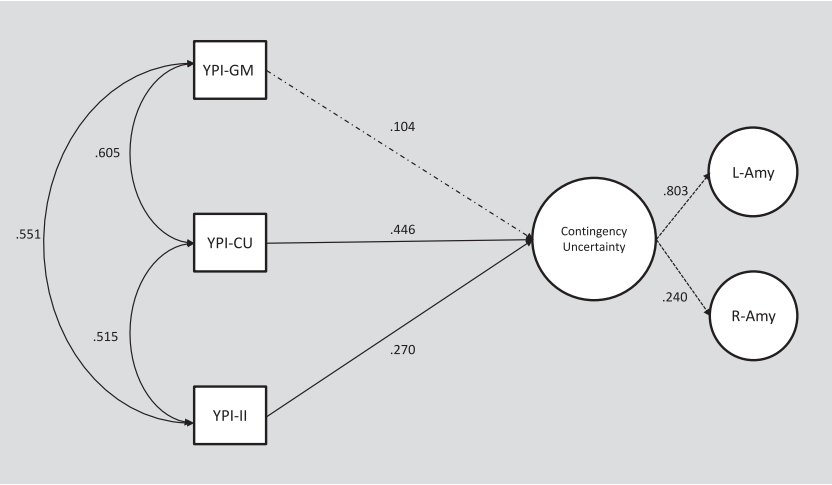


**Figure 1.** Structural equation model depicting the relationships among the Youth Psychopathic Traits Inventory (YPI) scales measuring callous-unemotional (YPI-CU), grandiose-manipulative (YPI-GM), and impulsive-irresponsible (YPI-II) traits and a latent factor representing the amount of perceptual uncertainty concerning contingencies, as estimated from the blood oxygen level–dependent signal trajectories. Only the left amygdala (L-Amy) and the right amygdala (R-Amy) loaded significantly on this latent factor. To increase readability, nonsignificant loadings are not depicted. Solid arrows represent significant unique correlations, dashed arrows represent nonsignificant effects, and dotted arrows represent factor loadings.

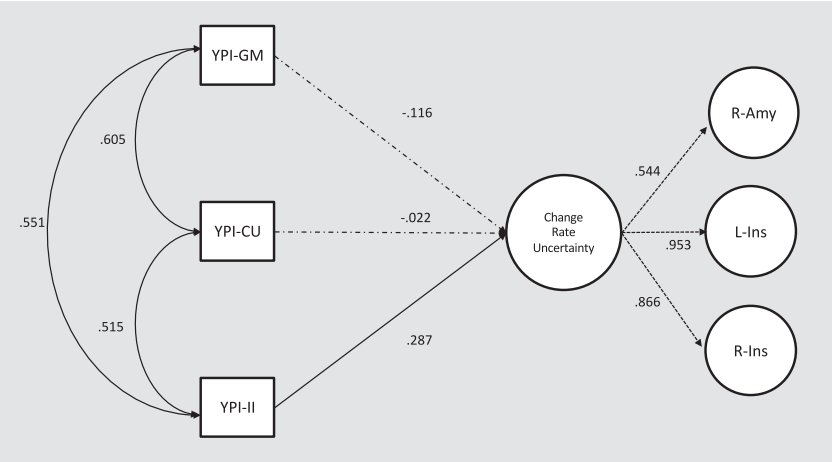


**Figure 2.** Structural equation model depicting the relationships among the Youth Psychopathic Traits Inventory (YPI) scales measuring callous-unemotional (YPI-CU), grandiose-manipulative (YPI-GM), and impulsive-irresponsible (YPI-II) traits and a latent factor representing perceptual uncertainty concerning the change rate of contingencies, as estimated from the blood oxygen level–dependent signal trajectories. The left insula (L-Ins), right insula (R-Ins), and right amygdala (R-Amy) loaded significantly on this latent factor. To increase readability, nonsignificant loadings are not depicted. Solid arrows represent significant unique correlations, dashed arrows represent nonsignificant effects, and dotted arrows represent factor loadings.

dimensions of psychopathy indicated that both callous-unemotional and impulsive-irresponsible traits were uniquely related to CoUn, while only impulsive-irresponsible features were positively linked to RUn (Figures 1 and 2).

Our results in an at-risk population are in line with those obtained in healthy samples, indicating that the amygdala (23,34) and the insula (24) are involved in processing uncertainty related to threat contingencies. Our results significantly advance this knowledge by specifying how activation in the insula and amygdala is related to uncertainty about the accuracy of different aspects of contingency representation. The amygdala seems to be particularly sensitive to CoUn during threat conditioning, but it also encodes a relatively high amount of RUn together with the insula. Given the hierarchical relationship between RUn and CoUn in our computational framework (13,15), these results converge with the suggestion that the insula passes information concerning aversive stimuli to the amygdala during threat conditioning (35). One tentative interpretation of the general pattern of results is that the interaction between the insula and the amygdala during threat conditioning might reflect a circuit in which the insula is primarily responsive to RUn and the information is then relayed to the right amygdala, which would then function as a point of entry, to inform computations taking place in the right and left amygdala that are related to the estimation of the likelihood that contingency changes occur. This proposal would also be consistent with the general notion that the insula relays somatosensory information to other areas to initiate adaptive responses (36).

Regarding the link with psychopathy, callous-unemotional and impulsive-irresponsible traits predicted increased CoUn in the amygdala. Because associative learning relies on successful reduction of uncertainty (15), heightened levels of uncertainty in the cognitive computations that are engaged ultimately should lead to reduced learning (18). In agreement with this prediction, Cohn *et al.* (19) reported a negative relationship between BOLD activation and callous-unemotional traits during threat conditioning in the dataset used in the current study, pointing toward mechanistic disturbances in threat conditioning (20,21). Our findings further specify one aspect of the mechanism that seems impaired in individuals with high levels of callous-unemotional traits, who seem to form more uncertain (i.e., less accurate) representations of contingency changes in the amygdala, while representations of change rate in the insula are relatively accurate.

Cohn *et al.* (19) also found a positive unique correlation between impulsive-irresponsible traits and BOLD signal during threat conditioning. This positive relationship, indicating enhanced learning from threats in those with increased levels of impulsive-irresponsible traits, was expected to be related to reduced representational uncertainty (i.e., more accurate representations) in the current study. Instead, we found that impulsive-irresponsible traits were positively linked to increased RUn and CoUn in contingency representations. Given these results, it seems plausible that there is a broader deficiency in forming representations concerning change in individuals scoring high on impulsive-irresponsible traits during threat conditioning, while the deficiency is limited to CoUn in those scoring high on callous-unemotional traits. However, representational uncertainty does not seem to be sufficient for explaining the enhanced threat conditioning seen with elevated impulsive-irresponsible traits (19). Therefore, it should be considered whether an additional mechanism could be interacting with the threat conditioning processes in this sample of at-risk adolescents. One possibility is that the insula and the amygdala are hypersensitive to aversive information in impulsive-irresponsible individuals. This interpretation builds on the previous finding that higher perceived uncertainty sensitizes the insula and amygdala to aversive information (23,24,34), presumably leading to exaggerated aversive responding in these regions. Thus, more representational uncertainty during aversive learning may be interacting with a bias toward exaggerated affective responding in the amygdala in individuals with high levels of irresponsible-impulsive features, while this bias might not be present in individuals with elevated callous-unemotional traits. One tentative prediction that follows is that impulsive-irresponsible individuals should also show exaggerated responses during extinction learning because the hypersensitization of the amygdala after threat conditioning combined with excessive representational uncertainty should interfere with the unlearning of contingencies during extinction. Obviously, this proposal is made with care

because we did not quantify amygdala bias in the current study and also because threat extinction has yet to be studied using a dimensional approach to psychopathy instead of group comparisons (11,37).

Interestingly, our findings also suggest that representational uncertainty could play an important role in explaining other learning impairments found in antisocial populations such as disturbed passive avoidance (38,39) and reversal learning (40,41). During reversal learning, for example, we learn that a stimulus previously associated with reward now may lead to negative outcomes, which requires us to adapt our beliefs and behavior to avoid negative consequences. Importantly, the change in stimulus–outcome associations introduces representational uncertainty about the contingencies and their rate of change, which affects how well and how fast we learn the new contingencies. From this perspective, the reversal learning impairment found in children (41) and adults (3,40,42) with psychopathic tendencies could reflect suboptimal management of representational uncertainty, similar to what we found in the current study. In line with this prediction, Budhani and Blair (41) found that the reversal learning impairment in boys with psychopathic tendencies got worse as the saliency of the contingency changes decreased. Such a reduction of saliency increases ambiguity and uncertainty about the contingencies, so their results suggest that increased CoUn may play a significant role in the reversal impairment. Future studies should try to confirm this expectation and determine the impact of representational uncertainty on reversal learning.

One critique of the current study could be that our task did not include a manipulation of uncertainty given that the aversive stimulus was always associated with the same neutral face during learning, which would make it clear when to expect the aversive stimulus. This argument stems from the issue that quantifying the impact of uncertainty on threat conditioning is not possible using traditional analytical approaches, thereby requiring the manipulation of uncertainty through the experimental task in blocked designs. However, the computational model used in the current study overcomes this limitation in that it directly quantifies the level of uncertainty about contingencies on a trial-by-trial basis. Therefore, the impact of uncertainty on the cognitive operations involved in learning can be measured without introducing drastic changes in contingencies through task design. Another potential issue is the possibility that the increased RUn in the insula found in those with elevated impulsive-irresponsible tendencies could be driven by a more general impairment in the integration of multiple sources of information used to generate contingency representations during threat conditioning (43,44). The insula can be seen as a processing hub implicated in various cognitive operations that include representations of pain, contextual appraisal, and general uncertainty (36). Thus, it is possible that increased RUn is a consequence of inaccurate lower level representations in the insula such as those pertaining to the pain stimulus. Such an account would also be in line with studies on reinforcement-based decision making in adolescents with conduct problems who show impairments in generating accurate representations of expected value in the insula (39,45). Together, these caveats and novel hypotheses highlight the need for further studies that focus on the interaction among personality dimensions, representational uncertainty, and their neural correlates during threat conditioning and other forms of associative learning.

In conclusion, the current study is the first to directly quantify different kinds of representational uncertainty during threat conditioning in an at-risk sample of adolescents. The results highlight the importance of examining how uncertainty in cognitive representations may be key to understanding some of the maladaptive characteristics often linked to psychopathy. A more precise understanding of the various interacting cognitive computations involved in maladaptive learning may lead to better (neuro)biology-oriented diagnostics and the development of targeted treatment approaches in various conditions showing disturbed associative learning (46,47).

## ACKNOWLEDGMENTS AND DISCLOSURES

MDC, AP, and IAB were supported by a Mosaic grant (Grant No. 017.007.022), a Brain and Cognition grant (Grant No. 056.23.010), and a Veni grant (Grant No. 451.15.014), respectively, from the Netherlands Organization for Scientific Research.

The authors thank Essi Viding for her valuable input.

All authors report no biomedical financial interests or potential conflicts of interest.

## ARTICLE INFORMATION

From the Donders Institute for Brain, Cognition and Behaviour (IAB), Radboud University, and Forensic Psychiatric Centre Pompestichting (IAB), Nijmegen, Department of Child and Adolescent Psychiatry (AP, MDC), VU University Medical Center Amsterdam, Amsterdam, Institute for Criminal Law & Criminology (AP), Leiden University, Leiden, and Department of Clinical Psychology (SSH), Erasmus University, Rotterdam, The Netherlands; Collaborative Antwerp Psychiatric Research Institute (IAB), University of Antwerp, Antwerp, Belgium; Centre for Psychology, Behaviour, & Achievement (IAB), Faculty of Health and Life Sciences, Coventry University, Coventry, Max Planck UCL Centre for Computational Psychiatry and Ageing Research (CDM) and Wellcome Trust Centre for Neuroimaging (CDM), Institute of Neurology, University College London, London, United Kingdom; and Scuola Internazionale Superiore di Studi Avanzati (CDM), Trieste, Italy.

Address correspondence to Inti A. Brazil, Ph.D., Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen 6525 HR, The Netherlands; E-mail: i.brazil@donders.ru.nl.

Received Feb 23, 2017; revised Apr 12, 2017; accepted Apr 12, 2017.

Supplementary material cited in this article is available online at http://dx.doi.org/10.1016/j.bpsc.2017.04.005.

## REFERENCES

1. LeDoux JE (2014): Coming to terms with fear. Proc Natl Acad Sci U S A 111:2871–2878.
2. Blair RJR, Mitchell DGV, Leonard A, Budhani S, Peschardt KS, Newman C (2004): Passive avoidance learning in individuals with psychopathy: Modulation by reward but not by punishment. Pers Individ Dif 37:1179–1192.
3. Gregory S, Blair RJR, Ffytche DH, Simmons A, Kumari V, Hodgins S, Blackwood N (2015): Punishment and the psychopath: An fMRI investigation of reinforcement learning in violent antisocial personality disordered men. Lancet Psychiatry 2:153–160.
4. Andershed HA, Kerr M, Stattin H, Levander S (2002): Psychopathic traits in non-referred youths: A new assessment tool. In: Blaauw E, Sheridan L, editors. Psychopaths: Current International Perspectives. The Hague, Netherlands: Elsevier, 131–158.
5. Hawes SW, Byrd AL, Waller R, Lynam DR, Pardini DA (2017): Late childhood interpersonal callousness and conduct problem trajectories

interact to predict adult psychopathy. J Child Psychol Psychiatry 58:55–63.
6. Wilkinson S, Waller R, Viding E (2016): Practitioner review: Involving young people with callous unemotional traits in treatment—does it work? A systematic review. J Child Psychol Psychiatry 57: 552–565.
7. Birbaumer N, Veit R, Lotze M, Erb M, Hermann C, Grodd W, Flor H (2005): Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. Arch Gen Psychiatry 62:799–805.
8. Flor H, Birbaumer N, Hermann C, Ziegler S, Patrick CJ (2002): Aversive Pavlovian conditioning in psychopaths: Peripheral and central correlates. Psychophysiology 39:505–518.
9. Hoppenbrouwers SS, Bulten BH, Brazil IA (2016): Parsing fear: A reassessment of the evidence for fear deficits in psychopathy. Psychol Bull 142:573–600.
10. Fairchild G, Stobbe Y, van Goozen SHM, Calder AJ, Goodyer IM (2010): Facial expression recognition, fear conditioning, and startle modulation in female subjects with conduct disorder. Biol Psychiatry 68:272–279.
11. Fairchild G, Van Goozen SH, Stollery SJ, Goodyer IM (2008): Fear conditioning and affective modulation of the startle reflex in male adolescents with early-onset or adolescence-onset conduct disorder and healthy control subjects. Biol Psychiatry 63:279–285.
12. Behrens TEJ, Hunt LT, Rushworth MFS (2009): The computation of social behavior. Science 324:1160–1164.
13. Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011): A Bayesian foundation for individual learning under uncertainty. Front Hum Neurosci 5:39.
14. Jazayeri M, Movshon JA (2006): Optimal representation of sensory information by neural populations. Nat Neurosci 9:690–696.
15. Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, Stephan KE (2014): Uncertainty in perception and the hierarchical Gaussian filter. Front Hum Neurosci 8:825.
16. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007): Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.
17. de Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, Bestmann S (2016): Computations of uncertainty mediate acute stress responses in humans. Nat Commun 7:10996.
18. Friston K (2010): The free-energy principle: A unified brain theory? Nat Rev Neurosci 11:127–138.
19. Cohn MD, Popma A, van den Brink W, Pape LE, Kindt M, van Domburgh L, *et al.* (2013): Fear conditioning, persistence of disruptive behavior and psychopathic traits: An fMRI study. Transl Psychiatry 3:e319.
20. Lopez R, Poy R, Patrick CJ, Molto J (2013): Deficient fear conditioning and self-reported psychopathy: The role of fearless dominance. Psychophysiology 50:210–218.
21. Veit R, Konicar L, Klinzing JG, Barth B, Yilmaz O, Birbaumer N (2013): Deficient fear conditioning in psychopathy as a function of interpersonal and affective disturbances. Front Hum Neurosci 7:706.
22. Sehlmeyer C, Schöning S, Zwitserlood P, Pfleiderer B, Kircher T, Arolt V, Konrad C (2009): Human fear conditioning and extinction in neuroimaging: A systematic review. PLoS One 4:e5865.
23. Herry C, Bach DR, Esposito F, Di Salle F, Perrig WJ, Scheffler K, *et al.* (2007): Processing of temporal unpredictability in human and animal amygdala. J Neurosci 27:5958–5966.
24. Sarinopoulos I, Grupe DW, Mackiewicz KL, Herrington JD, Lor M, Steege EE, Nitschke JB (2010): Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. Cereb Cortex 20:929–940.
25. Blair RJR (2013): The neurobiology of psychopathic traits in youths. Nat Rev Neurosci 14:786–799.
26. van Domburgh L, Vermeiren R, Blokland AAJ, Doreleijers TAH (2009): Delinquent development in Dutch childhood arrestees: Developmental trajectories, risk factors and co-morbidity with adverse outcomes during adolescence. J Abnorm Child Psychol 37:93–105.
27. van Baardewijk Y, Stegge H, Andershed H, Thomaes S, Scholte E, Vermeiren R (2008): Measuring psychopathic traits in children through self-report: The development of the Youth Psychopathic Traits Inventory–Child Version. Int J Law Psychiatry 31:199–209.
28. Andershed H, Hodgins S, Tengström A (2007): Convergent validity of the Youth Psychopathic Traits Inventory (YPI): Association with the Psychopathy Checklist: Youth Version (PCL: YV). Assessment 14: 144–154.
29. Brett M, Anton J, Valabregue R, Poline JB (2002): Region of interest analysis using the MarsBar toolbox for SPM 99. NeuroImage 16:S497.
30. Phelps EA, O'Connor KJ, Gatenby JC, Gore JC, Grillon C, Davis M (2001): Activation of the left amygdala to a cognitive representation of fear. Nat Neurosci 4:437–441.
31. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, *et al.* (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15:273–289.
32. Muthén LK, Muthén BO (1998–2012): Mplus: The Comprehensive Modeling Program for Applied Researchers: User's Guide. Los Angeles: Muthén & Muthén.
33. Gelman A, Rubin DB (1992): Inference from iterative simulation using multiple sequences. Statist Sci 7:457–472.
34. Whalen PJ (2007): The uncertainty of it all. Trends Cogn Sci 11:499–500.
35. Hartley CA, Fischl B, Phelps EA (2011): Brain structure correlates of individual differences in the acquisition and inhibition of conditioned fear. Cereb Cortex 21:1954–1962.
36. Singer T, Critchley HD, Preuschoff K (2009): A common role of insula in feelings, empathy and uncertainty. Trends Cogn Sci 13:334–340.
37. Cohn MD, van Lith K, Kindt M, Pape LE, Doreleijers TAH, van den Brink W, *et al.* (2016): Fear extinction, persistent disruptive behavior and psychopathic traits: fMRI in late adolescence. Soc Cogn Affect Neurosci 11:1027–1035.
38. Newman JP, Patterson CM, Howland EW, Nichols SL (1990): Passive avoidance in psychopaths: The effects of reward. Pers Individ Dif 11:1101–1114.
39. White SF, Pope K, Sinclair S, Fowler KA, Brislin SJ, Williams WC, *et al.* (2013): Disrupted expected value and prediction error signaling in youths with disruptive behavior disorders during a passive avoidance task. Am J Psychiatry 170:315–323.
40. Brazil IA, Maes JHR, Scheper I, Bulten BH, Kessels RPC, Verkes R, de Bruijn ERA (2013): Reversal deficits in individuals with psychopathy in explicit but not implicit learning conditions. J Psychiatry Neurosci 38:E13–E20.
41. Budhani S, Blair RJR (2004): Response reversal and children with psychopathic tendencies: Success is a function of salience of contingency change. J Child Psychol Psychiatry 46:972–981.
42. Budhani S, Richell RA, Blair RJR (2006): Impaired reversal but intact acquisition: Probabilistic response reversal deficits in adult individuals with psychopathy. J Abnorm Psychol 115:552–558.
43. Hoppenbrouwers SS, Van der Stigchel S, Sergiou CS, Theeuwes J (2016): Top-down attention and selection history in psychopathy: Evidence from a community sample. J Abnorm Psychol 125:435–441.
44. Hoppenbrouwers SS, Van der Stigchel S, Slotboom J, Dalmaijer ES, Theeuwes J (2015): Disentangling attentional deficits in psychopathy using visual search: Failures in the use of contextual information. Pers Individ Dif 86:132–138.
45. White SF, Tyler PM, Erway AK, Botkin ML, Kolli V, Meffert H, *et al.* (2016): Dysfunctional representation of expected value is associated with reinforcement-based decision-making deficits in adolescents with conduct problems. J Child Psychol Psychiatry 57:938–946.
46. Stephan KE, Iglesias S, Heinzle J, Diaconescu AO (2015): Translational perspectives for computational neuroimaging. Neuron 87:716–732.
47. Brazil IA, van Dongen JD, Maes JH, Mars RB, Baskin-Sommers AR (2016): Classification and treatment of antisocial individuals: From behavior to biocognition [published online ahead of print Oct 17]. Neurosci Biobehav Rev.
48. Kuhnen CM, Knutson B (2005): The neural basis of financial risk taking. Neuron 47:763–770.
49. Liu X, Powell DK, Wang H, Gold BT, Corbly CR, Joseph JE (2007): Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. J Neurosci 27:4587–4597.