

Fear extinction, persistent disruptive behavior and psychopathic traits: fMRI in late adolescence

Moran D. Cohn,¹ Koen van Lith,¹ Merel Kindt,² Louise E. Pape,^{1,3}
Theo A. H. Doreleijers,¹ Wim van den Brink,⁴ Dick J. Veltman,⁵
Arne Popma,^{1,6}

¹Department of Child and Adolescent Psychiatry, VU University Medical Center and ²Department of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands, ³Department of Radiology, NYU Langone Medical Center, Center for Brain Imaging, New York, NY, USA, ⁴Academic Medical Center University of Amsterdam, Amsterdam Institute for Addiction Research, Amsterdam, The Netherlands, ⁵Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands, and ⁶Institute for Criminal Law & Criminology, Leiden University, Leiden, The Netherlands

Correspondence should be addressed to Moran D. Cohn, Department of Child and Adolescent Psychiatry, VU University Medical Center, De Bascule, P.O. Box 303, 1115 ZG Duivendrecht, The Netherlands. E-mail: m.cohn@debascule.com

Abstract

Children diagnosed with a Disruptive Behavior Disorder (DBD, i.e. Oppositional Defiant Disorder or Conduct Disorder), especially those with psychopathic traits, are at risk of developing persistent and severe antisocial behavior. Reduced fear conditioning has been proposed to underlie persistent antisocial development. However, we have recently shown that both DBD persisters and desisters are characterized by increased fear conditioning compared with healthy controls (HCs). In this study, we investigated whether brain function during fear extinction is associated with DBD subgroup-membership and psychopathic traits. Adolescents from a childhood arrestee cohort (mean age 17.6 years, s.d. 1.4) who met criteria for a DBD diagnosis during previous assessments were re-assessed and categorized as persistent DBD ($n = 25$) or desistent DBD ($n = 25$). Functional MRI during the extinction phase of a classical fear-conditioning task was used to compare regional brain function between these subgroups and 25 matched controls. Both DBD persisters and desisters showed hyperreactivity during fear extinction, when compared with HCs. Impulsive-irresponsible psychopathic traits were positively associated with responses in the fear neurocircuitry and mediated the association between neural activation and group membership. These results suggest that fear acquisition and fear extinction deficits may provide an endophenotype for an emotionally hyper-reactive subtype of antisocial development.

Key words: fear conditioning; fMRI; antisocial behavior; psychopathy; extinction

Introduction

Children who engage in antisocial behavior cause substantial economic losses (Cohen and Piquero, 2009) and are at risk of persistent antisocial development (Moffitt *et al.*, 2002), but very little is known about the mechanisms involved in the persistence of antisocial behavior. There is, however, increasing evidence that psychopathic traits [i.e., grandiose-manipulative

(GM) traits, callous-unemotional (CU) traits and impulsive-irresponsible traits (II)], and especially CU traits, increase the risk of persistent antisocial behavior and predict negative treatment outcomes (for review, see Frick *et al.*, 2014). In addition, psychopathic traits have been associated with specific neurobiological processes, e.g. CU traits are associated with reduced amygdala responsiveness to some—but not all—emotional

Received: 2 September 2014; Revised: 1 May 2015; Accepted: 2 June 2015

© The Author (2015). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

stimuli (for review, see Blair, 2013). As such, CU traits have been included in DSM-5 as a subtype specifier for conduct disorder (CD) as 'limited prosocial emotions'.

However, there are many unresolved issues regarding the interrelations between the different psychopathic traits, their specific relations with persistent antisocial behavior and the underlying neurobiological processes. For example, similar to the neurobiological and phenomenological heterogeneity observed among antisocial youths in general (Blair, 2013), not all persistent offenders are characterized by CU traits and low emotional reactivity: part of this population even report high anxiety and shows emotional hyperreactivity (Hodgins, 2007). Moreover, there is hardly any data on the neural processes that may distinguish persistent from desistent antisocial youths and antisocial youths from healthy controls (HCs).

We have recently shown (Cohn et al., 2013) that neural activation during fear conditioning does not differentiate adolescents with a persistent or a desistent Disruptive Behavior Disorder [DBD; Oppositional Defiant Disorder (ODD) or CD] compared with HCs: both DBD-persisters (DBD-p) and DBD-desisters (DBD-d) showed increased activation of the fear circuit in response to cues predicting an aversive electric unconditioned stimulus vs neutral cues. Importantly, we also found that these neural activation patterns were related to the presence of II psychopathic traits, and group differences in neural activation were in fact mediated by these II traits. In a multiple regression model simultaneously including all psychopathic traits, II and GM traits were positively associated, whereas CU traits were negatively associated with activation of the fear neurocircuitry. Although the main findings in this study were not explained by anxiety scores, the observed hyperreactivity during fear conditioning and the positive association with II psychopathic traits suggest that these youngsters—as a group—were likely to be emotionally hyperreactive. However, our previous study only reported on fear acquisition and not on fear extinction, i.e. the experimental phase in which fear cues are not followed by an aversive stimulus anymore, which should lead to extinction of fear responses. This is important, because fear hyperreactivity and its clinical symptoms may actually result from an inability to down-regulate previously acquired fear responses (Graham and Milad, 2011).

While the evidence for fear acquisition deficits in psychopathy is relatively robust (Hare, 1978; Raine, 1993), this is not the case for fear extinction deficits. For example, two neuroimaging studies reported deficient fear conditioning in small samples of adult psychopaths when compared with HCs, but no group differences were observed during fear extinction (Veit et al., 2002; Birbaumer et al., 2005). Similarly, studies have implicated fear acquisition deficits in subgroups of antisocial juveniles (Raine and Venables, 1981; Raine et al., 1996; Fairchild et al., 2008), but did not report any differences during fear extinction (Fairchild et al., 2008). However, emotional hyporeactivity characterizes only one part of persistent offenders (Hodgins, 2007). Hyperreactivity during fear acquisition has been found in antisocial juveniles from lower social class (Raine and Venables, 1981) and in our previous fear acquisition study (Cohn et al., 2013). In summary, there is some evidence for aberrant fear responsiveness in subgroups of antisocial individuals, but the current literature on fear extinction in antisocial individuals is limited by an exclusive focus on the emotionally hyporeactive subgroup. As such, it seems relevant to investigate whether persistence of DBD in our hyperreactive sample is associated with an inability to down-regulate previously acquired fear responses in the face of changing contingencies.

Importantly, our previous study (Cohn et al., 2013) also showed that group differences in neural responses during fear acquisition were largely explained by differences in psychopathic traits. In addition, various studies have shown that psychopathic traits are not a unitary construct: i.e. that its dimensions are distinctly associated with behavioral criterion variables (Patrick et al., 2005; Hicks and Patrick, 2006) and regional brain function (Carré et al., 2013; Cohn et al., 2013). Indeed, we found that neural responses during fear acquisition were positively associated with GM and II psychopathic traits and negatively with CU psychopathic traits (Cohn et al., 2013). However, nothing is known about the associations between the different psychopathic traits and brain activation during fear extinction.

In this report, we present extinction data from our fear conditioning experiment (Cohn et al., 2013), that is, fMRI data from a large cohort of childhood arrestees followed up in late adolescence, allowing the distinction between DBD-p, DBD-d and HCs. Because extinction studies in antisocial samples are scarce, we used a basic paradigm assessing extinction *learning*—relying on function of the basic fear neurocircuitry [amygdala, insula and anterior cingulate cortex (ACC)]—rather than its *consolidation* or *modulation*—relying on the ventromedial prefrontal cortex and hippocampus, respectively (Quirk and Mueller, 2008; Delamater and Westbrook, 2014). Consequently, we used amygdala, insula and ACC as a priori regions of interest (ROIs). In the absence of previous studies on this topic, and given the results of this study during the fear acquisition phase (Cohn et al., 2013), we hypothesize: (1) hyperreactivity of the fear neurocircuitry during fear extinction in the DBD-p and the DBD-d group as compared with HC; and (2) hyperreactivity of the fear neurocircuitry to be related to, and mediated by, GM and II psychopathic traits, and CU psychopathic traits to be associated with hyporeactivity of the fear neurocircuitry. In addition, as a secondary outcome measure, we expect similar associations of DBD subgroups and different psychopathic traits with skin conductance responses (SCRs).

Materials and methods

Participants

This study reports on analyses regarding the extinction phase of a fear conditioning experiment, the acquisition data of which have been reported in Cohn et al. (2013). The study sample was recruited from a cohort of childhood arrestees who had all committed acts below the age of 12 that would be prosecutable above the Dutch age of criminal responsibility (12 years), ranging in severity from very mild (e.g. petty theft) to severe (e.g. sexual abuse, robbery). Subjects had participated in three previous waves of this longitudinal study: mean age at study entrance 10.9 (s.d. 1.4) and mean age at wave three 13.1 (s.d. 1.5). Eighty participants of this cohort had met criteria for a DBD (i.e. ODD or CD) on the Diagnostic Interview Schedule for Children (DISC-IV, see below; Shaffer et al., 2000) in at least one of the previous waves; 56 of whom participated in the current neuroimaging study. Six of these participants were excluded from analyses because of invalid MRI data, i.e. poor coverage due to susceptibility artifacts. The remaining 50 subjects were subdivided in those still meeting criteria for DBD (i.e. DBD-persisters, $n = 25$) and those who did not meet criteria anymore (i.e. DBD-desisters, $n = 25$) at follow-up at age 17.6.

Twenty-five matched HCs were recruited from the same childhood arrestee cohort. This group had never met criteria for

DBD, scored low on current wave aggression (Reactive-Proactive Aggression; RPQ, Raine et al., 2006) and psychopathic traits (Youth Psychopathic Traits Inventory; YPI, Andershed et al., 2002) and had no history of other axis 1 or 2 disorders. This study sample is identical to that reported on in Cohn et al. (2013), except for one participant who had to be excluded because of premature termination of the experiment. Table 1 lists the clinical characteristics of these subgroups.

Procedure

This study was approved by the IRB of the VU University Medical Center Amsterdam. All participants and their parents/custodians (if the participant's age was below 18) signed for informed consent. Participants were visited at home for behavioral testing and underwent a neuroimaging protocol in a Philips 3T Intera MRI scanner at the Academic Medical Center (Amsterdam, The Netherlands). All participants were instructed to refrain from using alcohol, cannabis or psychostimulant medication for at least 24 h before the MRI-scan.

Assessment

Both the parent and youth version of the National Institute of Mental Health DISC-IV (Shaffer et al., 2000) were used to assess criteria for DSM-IV ODD and CD: diagnoses were made when participants met diagnostic criteria according to either the parent or youth report. Desistence was defined as not meeting DSM-IV criteria, i.e. endorsing less than four ODD symptoms in the last half year, and less than three CD symptoms in the last year.

The Youth Psychopathic traits Inventory is a valid and reliable 50-item self-report instrument, developed in order to assess psychopathic traits in juvenile community samples

(Andershed et al., 2002). In this study, internal consistency (Cronbach's alpha) of the total score and its constituting dimensions were good to excellent: CU alpha 0.87; GM alpha 0.93; II alpha .88. Dimensions were correlated but far from collinear ($r = 0.62-0.67$).

The RPQ (Raine et al., 2006), Child Behavior Checklist (Achenbach, 1991), Youth Self Report (Achenbach, 1991) and Wechsler Intelligence Scale for Children-III (Wechsler, 1974) were used for descriptive purposes.

Fear conditioning task

A classical differential delay fear-conditioning task was employed, adapted from Birbaumer et al. (2005). Pictures of two neutral male faces served as conditioned stimuli (CS), one of which was consistently paired with an aversive electric unconditioned stimulus (US) at the end of a 10-s viewing period (CS+; i.e. 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, in which CSs were presented four times each for 3.5 s without a US and was followed by an extinction phase in which CSs were presented four times each for 7 s and were not followed by a US either. During the extinction phase the previously reinforced CS (CS+) is referred to as CS+/-.

Skin conductance data were collected to assess whether fear extinction was successful. Skin conductance was measured with MRI-compatible Ag/AgCl-electrodes and BIOPAC recording hardware and software (AcqKnowledge 4.1). Data were extracted from the raw signal with the Versatile Stimulus Response Registration Program (Molenkamp, 1998). SCRs (μ S) were calculated by subtracting the baseline (i.e. the mean skin conductance level of the 2 s before CS-onset) from the maximum skin conductance level during the CS.

Table 1. Characteristics of DBD subgroups and controls

	HC (n = 25)	DBD-d (n = 25)	DBD-p (n = 25)	Group difference
Male gender, no. (%)	22 (88%)	20 (80%)	18 (72%)	Fisher's exact $P = 0.43$
Low SES neighborhood, no. (%)	18 (72%)	15 (60%)	13 (54%)	Fisher's exact $P = 0.44$
Non-Western ethnicity, no. (%)	10 (40%)	4 (16%)	9 (36%)	Fisher's exact $P = 0.37$
Age, mean (s.d.), y	17.8 (1.2)	17.6 (1.7)	17.3 (1.4)	$F_{2,68} 0.7, P = 0.50$
DBD age of onset, mean (s.d.), y	-	6.6 (3.5)	6.5 (3.0)	$T_{47} 0.15, P = 0.87$
IQ, mean (s.d.)	92.5 (12.1)	91.2 (15.6)	86.7 (13.2)	$F_{2,65} 1.1, P = 0.34$
RPQ Aggression, mean (s.d.)	4.4 (2.6)	12.2 (7.0)	18.0 (8.8)	Welch $_{2,36.5} 36.8, P < 0.001^{a,b,c}$
CBCL Internalizing, mean (s.d.)	47.8 (10.8)	53.4 (12.0)	60.9 (6.3)	Welch $_{2,41.9} 14.0, P < 0.001^{b,c}$
YSR Internalizing, mean (s.d.)	41.8 (8.2)	48.4 (10.1)	53.2 (10.5)	$F_{2,71} 8.7, P < 0.001^{a,b}$
CBCL Externalizing, mean (s.d.)	44.1 (6.6)	57.0 (8.9)	66.6 (6.1)	Welch $_{2,43.6} 72.2, P < 0.001^{a,b,c}$
YSR Externalizing, mean (s.d.)	43.4 (5.5)	55.3 (8.5)	60.1 (11.4)	Welch $_{2,43.0} 30.9, P < 0.001^{a,b}$
CD (%)	0	0	13 (52%)	Fisher's exact $P < 0.001^{b,c}$
ADHD (%)	5 (20%)	12 (48%)	16 (64%)	Fisher's exact $P = 0.006^b$
PTSD (%)	0 (0%)	0 (0%)	2 (9%)	Fisher's exact $P = 0.10$
YPI Callous-Unemotional, mean (s.d.)	20.2 (4.0)	24.2 (7.5)	26.9 (9.0)	Welch $_{2,41.4} 7.5, P = 0.002^{a,b}$
YPI GM, mean (s.d.)	23.2 (3.4)	32.2 (10.1)	33.3 (10.9)	Welch $_{2,37.0} 17.1, P < 0.001^{a,b}$
YPI II, mean (s.d.)	24.5 (5.2)	33.0 (9.6)	36.9 (7.8)	Welch $_{2,44.2} 26.0, P < 0.001^{a,b}$
YPI total Psychopathic traits, mean (s.d.)	67.9 (8.4)	89.3 (23.8)	97.2 (23.8)	Welch $_{2,37.7} 23.7, P < 0.001^{a,b}$
Aversive stimulus level, mean (s.d.), mA	36.4 (12.7)	33.6 (13.6)	39.3 (10.5)	$F_{2,72} 1.3, P = 0.27$
Mean translation, mean (s.d.), mm	0.08 (0.02)	0.11 (0.06)	0.13 (0.07)	Welch $_{2,39.1} 7.3, P = 0.002^b$
Mean rotation, mean (s.d.), °	0.07 (0.02)	0.11 (0.06)	0.13 (0.09)	Welch $_{2,38.2} 7.5, P = 0.002^{a,b}$

SES, Socio-economic status; RPQ, Reactive Proactive aggression Questionnaire; CD, Conduct Disorder; ADHD, Attention Deficit/Hyperactivity Disorder; PTSD, Post-Traumatic Stress Disorder; CBCL, Child Behavior Checklist; YSR, Youth Self Report; YPI, Youth Psychopathic traits Inventory; HC, healthy control; DBD-d, desistent DBD subgroup; DBD-p, persistent DBD subgroup

^aSignificant difference between HC and desisters

^bHC vs persisters

^cdesisters vs persisters

fMRI protocol

T1-weighted anatomical scans (180 slices, 1 mm³ voxels, FOV 256 × 256 mm², TR 9.0 ms, TE 3.5 ms) were acquired using an 8-channel SENSE head-coil. Furthermore, 400 T2* weighted axial echo-planar images were acquired during fear conditioning (38 slices, 3-mm thickness, 2.29 × 2.29 in-plane resolution, FOV 220 × 220 mm², TR 2300, TE 30 ms).

Statistical analyses

Repeated-measures General Linear Models (rmGLMs) were used to assess average differential SCRs to extinction trials (2, 3 and 4) as compared with the final acquisition trials (6, 7 and 8). CS-type (CS+/- vs CS-) and time (late acquisition vs extinction) were used as within-subject measures and diagnostic group (DBD-p vs DBD-d vs HC) as between-subject measure. To account for violations of compound symmetry and sphericity, Wilks' Lambda multivariate tests of significance were used. Similar rmGLMs were performed with psychopathic traits (CU, GM, II) as dimensional between-subject measures.

fMRI data were processed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>), including realignment, unwarping, slice-time correction, normalization to MNI space based on the segmented anatomical scan, and 8-mm FWHM smoothing. First-level models included separate regressors for CS+/- and CS- during habituation, acquisition and extinction, US and rating blocks. As participants were only able to detect the change in contingency during extinction (i.e. the CS+ not being followed by a US anymore and becoming a CS+/-) after the first unreinforced CS+/-, the first trial was modeled as a separate regressor. Similar to the fear acquisition study, we separated the fear extinction trials in two 3.5 s epochs to account for fast within-trial habituation of fear neurocircuitry—focusing analyses on the first epoch only—and modeled first-order time-modulation regressors to account for between-trial habituation effects. Realignment parameters were also included in first-level models to account for movement effects. Next, contrast images were calculated, subtracting the CS- from the CS+/- for the second, third and fourth extinction trial, which were entered into a one-way ANOVA to assess between-group differences (specifically testing our hypothesis that DBD-p and DBD-d would show higher differential responses than HC, i.e. [(DBD-p & DBD-d) vs HC]). Post hoc pairwise analyses were performed to assess whether DBD-p and DBD-d differed from HC, and DBD-p from DBD-d. In addition, we performed three separate regression analyses to evaluate the relations between dimensional measures of psychopathic traits and [CS+/- > CS- responses. Finally, we performed a multiple regression analysis incorporating all three psychopathic traits dimensions to evaluate the unique association for each dimensional measure of psychopathic traits while controlling for any suppressor effects from the remaining predictor dimensions.

Similar to our previous fear acquisition study, analyses were conducted at a whole-brain level, as well as in a priori ROIs (see 'Introduction' section) involved in fear acquisition and extinction, i.e. the amygdala, insula and ACC (Mechias et al., 2010). Amygdala and insula were anatomically defined using the Automated Anatomic Labeling atlas (Tzourio-Mazoyer et al., 2002). Similar to previous neuroimaging fear conditioning studies (Well et al., 2012; Cohn et al., 2013), we assessed ACC effects using a 16-mm radius sphere centered on the peak coordinates ($x = -2$, $y = 14$, $z = 40$) from a recent fear conditioning meta-analysis (Mechias et al., 2010). Results are reported at a multiple-comparison corrected threshold of Family Wise Error $P < 0.05$,

both at the whole-brain level and—using small volume correction—in the ROIs.

Mediation analyses were performed using bootstrapping procedures with 5000 resamples as implemented in the SPSS macro INDIRECT provided by Andrew Hayes (<http://www.afhayes.com>; Preacher and Hayes, 2008), thresholded at two-sided $\alpha 0.05$ (i.e. lower boundary of the 95% confidence interval > 0). Mean neuronal activation of the ACC ROI was extracted as an unbiased estimate using the MarsBaR toolbox for SPM (<http://marsbar.sourceforge.net>; Brett et al., 2002).

Results

Conditioning indices

Skin conductance analyses showed significant fear extinction effects (Table 2). Post hoc analysis showed that significant fear extinction had occurred, with higher differential [CS+/- > CS-] responses during the late acquisition phase ($M = 0.39$, $s.d. = 1.00$) than during the extinction phase ($M = 0.04$, $s.d. = 0.56$, paired $t_{74} = 2.9$, $P = 0.005$). No significant interactions with DBD group membership (all $P > 0.35$) or psychopathic traits (all $P > 0.39$) were observed. However, there were significant between-subject effects, indicating that DBD group membership and II traits ([marginally] $F_{1,71} = 4.0$, $P = 0.050$, partial $\eta^2 = 0.05$) were related to the skin conductance rmGLM intercept. Post hoc analysis revealed that these effects were driven by significantly higher mean SCRs in the DBD-p group than in the HC group, and by a significant positive correlation between II traits and mean SCR ($r_{73} = 0.25$, $P = 0.03$). Mean SCR in DBD-d, however, did not differ significantly from mean SCR in either DBD-p or HC ($P > 0.05$). Notably, while there were no significant time*group or time*traits interactions, significant associations with SCR were found during extinction (DBD-p > HC; II traits: $r_{73} = 0.28$, $P = 0.016$) but not during acquisition (DBD-p > HC; II traits: $r_{73} = 0.14$, $P = 0.25$), suggesting that extinction findings could not be entirely explained by acquisition differences.

Similarly, when all three psychopathic traits dimensions were simultaneously entered into a single rmGLM, none of the psychopathic traits showed a significant interaction with CS, time or CS*time (all $P > 0.22$) but significant main effects of II traits ($F_{1,69} = 12.3$, $P = 0.001$, partial $\eta^2 = 0.15$) and, additionally, CU traits ([marginally], $F_{1,69} = 3.1$, $P = 0.08$, partial $\eta^2 = 0.04$) were found. Post hoc analysis revealed that, again, II traits were related to a higher mean SCR ($\beta = 0.58$, $t_{69} = 3.8$, $P < 0.001$), CU traits were related to lower mean SCR ([marginally], $\beta = -.30$, $t_{69} = -1.9$, $P = 0.06$), and GM traits were unrelated to SCR.

Functional MRI

One-sample t -tests for the contrast [CS+/- > CS-] during extinction revealed that significant fear extinction had taken place, with only marginally significant remaining differential activation in the left ACC ($t_{72} = 3.5$, $z = 3.4$, $p_{FWE-svc} = 0.086$ at $[-8 22 28]$).

Between-group analyses (Figure 1 and Table 3) revealed significantly higher responses in the ACC, left insula and right insula in the combined DBD-p and DBD-d groups, compared with the HC group, during the extinction of previously acquired conditioned fear responses. Pairwise comparisons revealed that DBD-p showed higher differential responses in the ACC compared with HC. Furthermore, compared with HC, DBD-d showed higher responses in the right insula and ACC (marginally).

Table 2. Differential SCRs during extinction and acquisition

Repeated-measures GLM				
Multivariate tests				
CS	$F_{1,72}^{\dagger} = 7.9$	$P = 0.006^*$	partial $\eta^2 = .10$	
CS*DBD	$F_{2,72} = 0.5$	$P = 0.61$	partial $\eta^2 = 0.01$	
Time	$F_{1,72} = 6.4$	$P = 0.014^*$	partial $\eta^2 = 0.08$	
Time*DBD	$F_{2,72} = 1.1$	$P = 0.35$	partial $\eta^2 = 0.03$	
CS*time	$F_{1,72} = 8.0$	$P = 0.006^*$	partial $\eta^2 = 0.10$	
CS*time*DBD	$F_{2,72} = 0.8$	$P = 0.43$	partial $\eta^2 = 0.02$	
Between-subject-effect				
DBD	$F_{2,72} = 3.3$	$P = 0.043^*$	partial $\eta^2 = 0.08$	
Differential SCRs				
	HC (n = 25) M (s.d.)	DBD-d (n = 25) M (s.d.)	DBD-p (n = 25) M (s.d.)	Group difference
Acquisition	0.25 (0.50)	0.50 (0.65)	0.53 (0.63)	$F_{2,72} = 1.6, P = 0.21$
Extinction	0.10 (0.20)	0.24 (0.34)	0.49 (0.66)	$F_{2,72} = 4.9, P = 0.01^a$
Average SCR (across acquisition and extinction)	0.18 (0.30)	0.37 (0.42)	0.51 (0.59)	$F_{2,72} = 3.3, P = 0.043^a$

[†]All multivariate F-values refer to Wilks' Lambda F-values.

CS, Conditioned stimulus type [CS+ vs CS-] effects; DBD, Disruptive behavior disorders subgroup [HC vs DBD-d vs DBD-p] effects; HC, healthy control; DBD-d, desistent DBD subgroup; DBD-p, persistent DBD subgroup.

^aSignificant difference between HC persisters.

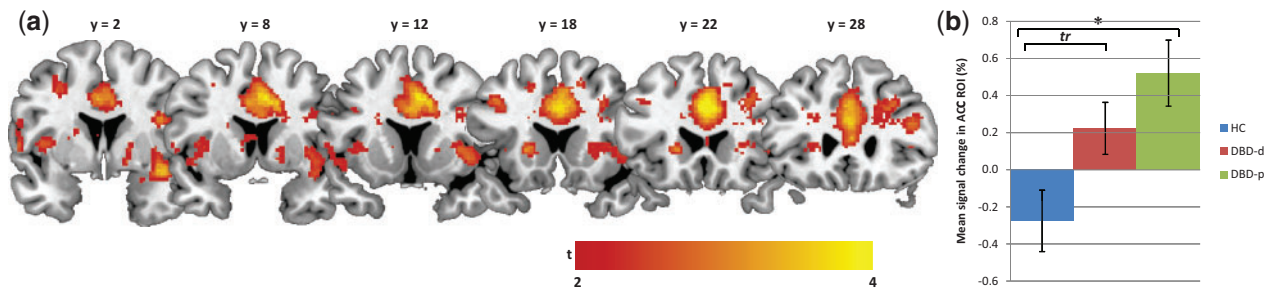


Fig. 1. ACC responsiveness during fear extinction in DBD subgroups. (A) Statistical parametric map for the contrast [(DBD-p & DBD-d) > HC] overlaid on an anatomical template (thresholded at $t > 2$ for display purposes, ranging from $t = 2$ [red] to $t = 4$ [yellow]). (B) Bar graph, showing mean % signal change in the ACC ROI per DBD subgroup, and indicating which group differences reached Family Wise Error corrected significance. DBD-p, DBD-persisters; DBD-d, DBD-desisters; HC, healthy controls;

* $p_{FWE-SVC} < 0.05$; $tr = 0.05 < p_{FWE-SVC} < 0.1$.

Table 3. Significant between-group differences in activity for conditioned responses to CS+ vs CS- during extinction (n = 75)

Group comparison	Brain region	$P_{FWE-SVC}$	Z-score	x	y	z
HC < (DBD-p and DBD-d)	ACC	0.014	3.9	2	22	32
	Left insula	0.026	3.8	-40	-8	10
	Right insula	0.048	3.6	40	-2	-10
HC < DBD-p	ACC	0.018	3.9	2	20	28
HC < DBD-d	Right insula	0.026	3.8	40	-2	-10
	ACC	0.058	3.5	6	14	40

CS+, conditioned stimulus followed by unconditioned stimulus; CS-, conditioned stimulus never followed by unconditioned stimulus; $P_{FWE-SVC}$, family wise error small volume correction for multiple comparison; DBD-d, desisters; DBD-p, persisters; HC, controls, ACC, anterior cingulate cortex.

Simple regression analyses (Figure 2 and Table 4) showed significant positive associations between II traits and neural responses during fear extinction in the right fusiform gyrus at the whole-brain level, as well as in the right amygdala and insula in ROI analyses. GM and CU traits were not significantly associated with neural responses during fear extinction in any of the ROIs or at the whole brain level. When all dimensions were simultaneously entered into a multiple regression model,

II traits were positively associated with neural responses during fear extinction in the right amygdala and left parahippocampal gyrus at the whole-brain level. ROI analyses revealed similar positive associations with responses in the ACC, left amygdala and right amygdala.

Since DBD-p and DBD-d membership as well as II traits were positively associated with ACC responsiveness during fear extinction, we performed a post hoc analysis to test whether II traits mediated the relation between neural responses during fear extinction and DBD-p and DBD-d membership (tested dichotomously vs HC, i.e. [(DBD-p+DBD-d) vs HC]). Bootstrapping tests indicated that II traits significantly mediated the path from ACC to pooled (DBD vs HC CI 95% 0.19–1.2; Figure 3) as well as individual DBD group membership (DBD-d vs HC: CI 95% 0.06–1.21; DBD-p vs HC: CI 95% 0.16–3.2).

We performed post hoc analyses to assess whether differential neural response patterns during extinction did not only reflect the residual effects of higher initial fear acquisition. No significant phase (acquisition vs extinction) by group or trait interactions were found, but associations were strongest—and only significant—during extinction (see Supplementary data), indicating that extinction findings are unlikely to solely reflect initial fear acquisition differences. Moreover, none of the associations between neural response patterns during extinction and

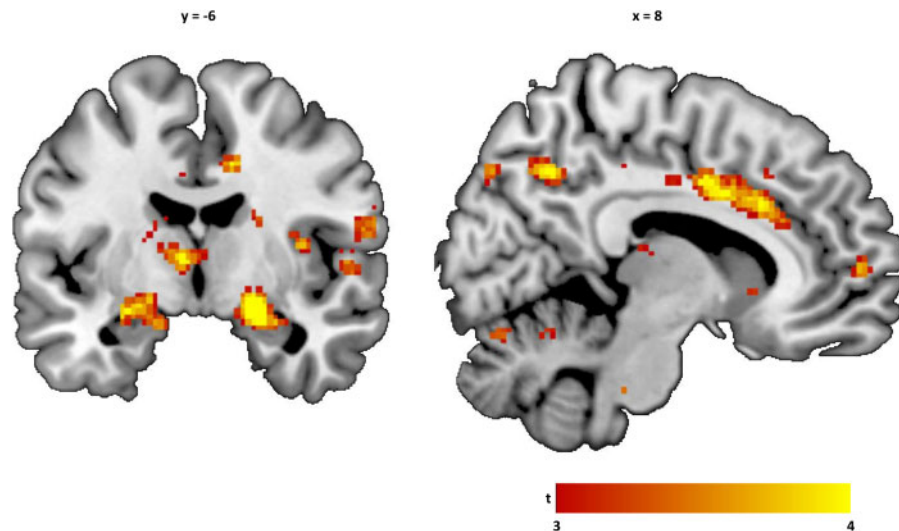


Fig. 2. ACC and amygdala responses during fear extinction and YPI II traits. Statistical parametric map showing unique positive associations with YPI Impulsive Irresponsible traits, i.e. controlling for the other dimensions, overlaid on an anatomical template (thresholded at $t > 2$ for display purposes, ranging from $t = 2$ [red] to $t = 4$ [yellow]). YPI, Youth Psychopathic Traits Inventory.

Table 4. Psychopathic trait dimensions' relation with regional BOLD response for CS+ vs CS- ($n = 73$)^a

	Brain region	P_{FWE}	T_{71} -score (Z-score)	x	y	z
Simple regression						
	YPI GM traits	ns				
	YPI Callous-unemotional traits	ns				
	YPI II traits					
	Fusiform gyrus	0.034 [†]	5.1 (4.7)	28	-40	-14
	Amygdala	0.025 ^{svc}	3.4 (3.3)	22	-6	-12
	Insula	0.018 ^{svc}	4.1 (3.9)	32	-14	22
Multiple regression, controlling for the effects of both other dimensions						
	YPI GM traits	ns				
	YPI Callous-unemotional traits	ns				
	YPI II traits					
	Amygdala	0.016 [†]	5.4 (4.9)	22	-10	-16
		0.002 ^{svc}	4.2 (4.0)	22	-6	-12
		0.004 ^{svc}	4.0 (3.8)	-26	-6	-14
	Parahippocampal gyrus	0.025 [†]	5.2 (4.8)	-22	-26	-14
	ACC	0.013 ^{svc}	4.2 (3.9)	8	20	30

^aYouth Psychopathic traits Inventories were missing for two participants.

CS+, conditioned stimulus followed by unconditioned stimulus; CS-, conditioned stimulus never followed by unconditioned stimulus, P_{FWE} , family wise error corrected p ; ACC, anterior cingulate cortex; svc, small volume corrected family wise error P -value.

[†]Whole-brain family wise error P -value.

both DBD-group membership and II traits scores were mediated by neural response patterns during acquisition (see Supplementary data). Finally, neither attention deficit/hyperactivity disorder diagnostic status nor anxiety scores confounded the reported findings (see Supplementary data).

Discussion

This study shows that, in childhood arrestees followed up in late adolescence, neural responses during fear extinction do not differ significantly between DBD persisters and DBD desisters. As expected, both DBD persisters and desisters, when compared with HCs, are characterized by hyperreactivity of the fear neurocircuitry during fear extinction. Furthermore, II traits are positively associated with neural responses during extinction, and partly mediate the association between ACC activation during extinction and DBD group membership, while no significant

association is observed between GM or CU traits and fear responses during extinction. Imaging findings are corroborated by similar associations of DBD subgroups and II traits with SCRs, although DBD desisters did not differ significantly from either persisters or controls in this respect. These observations extend our previous findings of hyperreactivity of the fear neurocircuitry in DBD persisters and desisters during fear acquisition, showing that these exaggerated anticipatory fear response patterns persist when cues are no longer followed by aversive reinforcement. Finally, the current findings show that certain psychopathic traits constitute relevant phenotypes capturing the neurobiological processes underlying the development of antisocial behaviors.

To the best of our knowledge, this is the first neuroimaging study showing fear extinction deficits in antisocial juveniles. Although part of this association could be driven by the hyperreactivity detected during fear acquisition (as phase-interaction

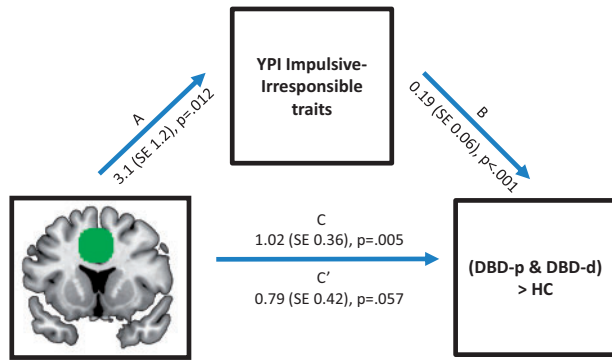


Fig. 3. Mediation of the association between ACC responsiveness and DBD subgroup membership by YPI II traits. Significant mediation of the association between neural responses in the ACC during fear condition and DBD subgroup membership, i.e. [(DBD-p and DBD-d) vs HC], by YPI II traits was ascertained by bootstrapping. Paths represent associations between (a) IV and M, (b) M and DV, (c) total effect of IV on DV, (c') direct effect of IV on DV, i.e. correcting for M. ACC, anterior cingulate cortex; YPI, Youth Psychopathic Traits Inventory; IV, independent variable; M, mediator; DV, dependent variable.

terms did not reach significance), both brain and SCR analyses showed that the associations of fear responses with DBD group membership and II traits were most pronounced during extinction and were not mediated by acquisition response patterns, suggesting that findings do not merely reflect the residual effects of initial hyperreactivity during fear acquisition. Several neuroimaging fear conditioning studies in psychopathic adults found no differences with HCs during fear extinction (Veit et al., 2002; Birbaumer et al., 2005), but their participants hardly showed fear acquisition, and these studies were underpowered to detect small differences. However, in antisocial juveniles, the combination of reduced fear acquisition and normal fear extinction has been reported (Fairchild et al., 2008). The latter finding is in contrast with our findings. Possible explanations for this discrepancy include neurobiological heterogeneity among antisocial youths and between-sample differences. More specifically, considering the theoretical frameworks on neurobiological heterogeneity in antisocial and psychopathic development laid out by Blair (2013) and Hodgins (2007), one could argue that both DBD persisters and desisters in this study were close to the hyperreactive end of a spectrum of emotional reactivity. However, contrary to these theoretical frameworks, the results in this study were not explained by anxiety scores, warranting their replication. Supportive evidence for extinction deficits in antisocial juveniles comes from a number of studies assessing extinction learning using operant reversal learning paradigms, showing both behavioral (Budhani and Blair, 2005; as previously reported in adults: Newman et al., 1987; Mitchell et al., 2002; Budhani et al., 2006) and neural abnormalities (Finger et al., 2008). This study suggests that such deficits may also become manifest in the context of fear extinction, and may differentiate DBD persisters and desisters from controls. However, fear extinction did not differ between DBD persisters and desisters, suggesting that other biological, psychological or social factors may explain the behavioral differences. We have recently shown that reward processing may provide one such explanation, because only DBD persisters in this sample were characterized by aberrant neural response patterns during reward outcomes (Cohn et al., 2014).

Notwithstanding the need for further studies on the processes underlying deficient fear extinction in DBD and II traits, their association may allow for a mechanistic interpretation of the reactive aggression commonly seen in relation to these

traits (Falkenbach et al., 2008). Specifically, we speculate that an inability to down-regulate fear responses under safe conditions may manifest as the 'hostile attribution bias' (Dodge et al., 1990) that has been observed in severely aggressive individuals. This suggestion has two important implications. First, it underscores the relevance of treating co-morbid axis I and axis II conditions characterized by high levels of anxiety in antisocial juveniles, as anxiety has been convincingly related to fear extinction deficits (Graham and Milad, 2011). Although *post hoc* analyses showed that anxiety symptom scores did not explain the current findings, there may still be synergetic mechanisms in this respect. Second, while the range of psychopharmacological treatments approved for antisocial disorders is limited, there is a wide range of experimental interventions developed specifically for the fear extinction deficits in anxiety disorders (e.g. d-cycloserine, Norberg et al., 2008; propranolol, Kindt et al., 2009; yohimbine, Powers et al., 2009). The current findings suggest that—even when antisocial juveniles do not meet criteria for these disorders, such therapies may provide experimental alternative interventions for hostile attribution tendencies or reactive aggression, the efficacy of which should be the topic of future studies.

Although we hypothesized that GM traits would be positively and CU traits would be negatively associated with differential fear responses when controlling for the other psychopathy dimensions, we did not find these associations in this study. However, a trend-level negative association between CU traits and SCRs was found when controlling for the other psychopathy dimensions. In addition, inspection of neuroimaging data at a lower statistical threshold revealed that CU traits were related to a lower differential fear response in the right amygdala during extinction ($t_{69} = 2.9$, $z = 2.8$, $p_{FWE-svc} = .08$ at [28 2–14]). Given the number of comparisons in this study and the trend-level significance of these findings, their replication is warranted.

The main strengths of this study were the large sample of at-risk adolescents and the focus on clinically relevant and prospectively ascertained developmental subgroups. However, the current findings should be regarded in light of several limitations. First, we used a very basic and rather short extinction scheme: due to the 100% reinforcement strategy (which was employed to enhance robust fear acquisition), subjective awareness of the contingency change is likely to have occurred immediately after trial 1, limiting the extinction phase. In addition, the lack of detailed subjective report of US expectancy did not allow inference on two competing hypotheses, i.e. whether participants were unaware of the contingency change or showed persistent fear responses despite adequate cognitive awareness. Furthermore, the current experimental procedure does not allow inference on the *cause* of fear extinction deficits as the task employed was not optimized to investigate the actual contingency information updating process, but rather its consequences (i.e. the relative decrease in anticipatory responses). Thus, future studies should employ paradigms with higher ecological validity, i.e. with <100% reinforcement schemes and longer extinction periods, as well as detailed reports of US expectancy. Finally, although the psychopathic traits measure employed in this study is valid and reliable (Andershed et al., 2002; Skeem and Cauffman, 2003), its self-report nature may not capture all the variance of the psychopathy construct and we advise future studies to additionally employ reports by parents or teachers.

In conclusion, this study suggests that the previously reported hyperreactivity of fear neurocircuitry in DBD persisters

and desisters extends to the extinction phase, which may result in hostile attribution bias and aggressive defense tendencies, with potential therapeutic implications. In addition, this study underscores the relevance of dimensional phenotypical measures of psychopathic traits when interpreting neurobiological findings in antisocial youth and, ultimately, in personalized interventions.

Acknowledgements

The authors thank P. F. C. Groot for technical assistance.

Funding

This study was funded by a Netherlands Organisation for Scientific Research (NWO) Mosaic grant (017.007.022) and by a NWO Brain & Cognition grant (056-23-010). Funders were not involved in any phase of the study.

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

- Achenbach, T.M. (1991). Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, VT, University of Vermont.
- Andershed, H., Kerr, M., Stattin, H., Levander, S. (2002). Psychopathic traits in non-referred youths: initial test of a new assessment tool. In: Blaauw, E., Sheridan, L., editors. *Psychopaths: Current International Perspectives*. The Hague, The Netherlands, Elsevier, 131–58.
- Birbaumer, N., Veit, R., Lotze, M., et al. (2005). Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study. *Archives of General Psychiatry*, *62*(7), 799–805.
- Blair, R.J. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews. Neuroscience*, *14*(11), 786–99.
- Brett, M., Anton, J.-L., Valabregue, R., et al. (2002). Region of interest analysis using an SPM toolbox. In: Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan. Available on CD-ROM in *NeuroImage*, *16*(2), abstract 497.
- Budhani, S., Blair, R.J. (2005). Response reversal and children with psychopathic tendencies: success is a function of salience of contingency change. *Journal of Child Psychology and Psychiatry*, *46*(9), 972–81.
- Budhani, S., Richell, R.A., Blair, R.J. (2006). Impaired reversal but intact acquisition: probabilistic response reversal deficits in adult individuals with psychopathy. *Journal of Abnormal Psychology*, *115*(3), 552–8.
- Carré, J.M., Hyde, L.W., Neumann, C.S., et al. (2013). The neural signatures of distinct psychopathic traits. *Social Neuroscience*, *8*(2), 122–35.
- Cohen, M.A., Piquero, A.R. (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, *25*(1), 25–49.
- Cohn, M.D., Popma, A., van den, B.W., et al. (2013). Fear conditioning, persistence of disruptive behavior and psychopathic traits: an fMRI study. *Translational Psychiatry*, *3*, e319.
- Cohn, M.D., Veltman, D.J., Pape L.E., et al. (2014). Incentive processing in persistent disruptive behavior and psychopathic traits: a functional Magnetic Resonance Imaging study in adolescents. *Biological Psychiatry*. doi: 10.1016/j.biopsych.2014.08.017.
- Delamater, A.R., Westbrook, R.F. (2014). Psychological and neural mechanisms of experimental extinction: a selective review. *Neurobiology of Learning and Memory*, *108*, 38–51.
- Dodge, K.A., Price, J.M., Bachorowski, J.A., et al. (1990). Hostile attributional biases in severely aggressive adolescents. *Journal of Abnormal Psychology*, *99*(4), 385–92.
- Fairchild, G., van Goozen, S.H., Stollery, S.J., et al (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. *Biological Psychiatry*, *63*(3), 279–85.
- Falkenbach, D., Poythress, N., Creevy, C. (2008). The exploration of subclinical psychopathic subtypes and the relationship with types of aggression. *Personality and Individual Differences*, *44*(4), 821–32.
- Finger, E.C., Marsh, A.A., Mitchell, D.G., et al. (2008). Abnormal ventromedial prefrontal cortex function in children with psychopathic traits during reversal learning. *Archives of General Psychiatry*, *65*(5), 586–94.
- Frick, P.J., Ray, J.V., Thornton, L.C., et al. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin*, *140*(1), 1–57.
- Graham, B.M., Milad, M.R. (2011). The study of fear extinction: implications for anxiety disorders. *The American Journal of Psychiatry*, *168*(12), 1255–65.
- Hare, R.D. (1978). Electrodermal and cardiovascular correlates of psychopathy. In: Hare, R.D., Schalling, D., editors. *Psychopathic Behavior: Approaches to Research*. New York, Wiley, 107–44.
- Hicks, B.M., Patrick, C. J. (2006). Psychopathy and negative emotionality: analyses of suppressor effects reveal distinct relations with emotional distress, fearfulness, and anger-hostility. *Journal of Abnormal Psychology*, *115*(2), 276–87.
- Hodgins, S. (2007). Persistent violent offending: what do we know? *The British Journal of Psychiatry Supplement*, *49*, s12–s14.
- Kindt, M., Soeter, M. (2014). Fear inhibition in high trait anxiety. *PLoS One*, *9*(1), e86462.
- Kindt, M., Soeter, M., Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature Neuroscience*, *12*, 256–8.
- Mechias, M.L., Etkin, A., Kalisch, R. (2010). A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *NeuroImage*, *49*(2), 1760–68.
- Mitchell, D.G., Colledge, E., Leonard, A., et al. (2002). Risky decisions and response reversal: is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? *Neuropsychologia*, *40*(12), 2013–22.
- Moffitt, T.E., Caspi, A., Harrington, H., et al. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: follow-up at age 26 years. *Development and Psychopathology*, *14*(1), 179–207.
- Molenkamp, E. (1998). Versatile Stimulus Response Registration Program (VSRPP98). www.test.uva.nl/ozip_psychology/index.php?Page0Software. Amsterdam: University of Amsterdam.
- Newman, J.P., Patterson, C.M., Kosson, D.S. (1987). Response perseveration in psychopaths. *Journal of Abnormal Psychology*, *96*(2), 145–8.
- Norberg, M.M., Krystal, J.H., Tolin, D.F. (2008). A meta-analysis of D-cycloserine and the facilitation of fear extinction and exposure therapy. *Biological Psychiatry*, *63*(12), 1118–26.

- Patrick, C.J., Hicks, B.M., Krueger, R.F., et al. (2005). Relations between psychopathy facets and externalizing in a criminal offender sample. *Journal of Personality Disorder*, *19*(4), 339–56.
- Powers, M.B., Smits, J.A., Otto, M.W., et al. (2009). Facilitation of fear extinction in phobic participants with a novel cognitive enhancer: a randomized placebo controlled trial of Yohimbine augmentation. *Journal of Anxiety Disorder*, *23*(3), 350–6.
- Preacher, K.J., Hayes, A.F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–91.
- Quirk, G.J., Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*, *33*, 56–72.
- Raine, A. (1993). *The Psychopathology of Crime: Criminal Behaviour as a Clinical Disorder*. San Diego, CA: Academic Press.
- Raine, A., Dodge, K., Loeber, R., et al. (2006). The reactive-proactive aggression questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior*, *32*(2), 159–71.
- Raine, A., Venables, P.H. (1981). Socialization and classical conditioning: a biosocial interaction. *Personality and Individual Differences*, *2*, 273–83.
- Raine, A., Venables, P.H., Williams, M. (1996). Better autonomic conditioning and faster electrodermal half-recovery time at age 15 years as possible protective factors against crime at age 29 years. *Developmental Psychology*, *32*(4), 624–30.
- Shaffer, D., Fisher, P., Lucas, C.P., et al. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry* *39*(1), 28–38.
- Skeem, J.L., Cauffman, E. (2003). Views of the downward extension: comparing the Youth Version of the Psychopathy Checklist with the Youth Psychopathic traits Inventory. *Behavioral Sciences and the Law*, *21*(6), 737–70.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–89.
- Veit, R., Flor, H., Erb, M., et al. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neuroscience Letters*, *328*(3), 233–36.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.
- Well, S.V., Visser, R.M., Scholte, H.S., et al. (2012). Neural substrates of individual differences in human fear learning: evidence from concurrent fMRI, fear-potentiated startle, and US-expectancy data. *Cognitive Affective and Behavioral Neuroscience*, *12*(3), 499–512.