Article

Oxytocin normalizes the implicit processing of fearful faces in psychopathy: a randomized crossover study using fMRI

John Tully [1,2,5] ✉, Arjun Sethi[1,5], Julia Griem[1], Yannis Paloyelis[3], Michael C. Craig[1], Steven C. R. Williams [3], Declan Murphy [1], Robert James Blair[4,6] & Nigel Blackwood[1,6]

Adults with antisocial personality disorder with (ASPD + P) and without (ASPD – P) psychopathy commit the majority of violent crimes. Empathic processing abnormalities are particularly prominent in psychopathy, but effective pharmacological interventions have yet to be identified. Oxytocin modulates neural responses to fearful expressions in healthy populations. The current study investigates its effects in violent antisocial men. In a placebo-controlled, randomized crossover design, 34 violent offenders (19 ASPD + P; 15 ASPD – P) and 24 healthy non-offenders received 40 IU intranasal oxytocin or placebo and then completed an fMRI morphed faces task examining the implicit processing of fearful facial expressions. Increasing intensity of fearful facial expressions failed to appropriately modulate activity in the bilateral mid-cingulate cortex in violent offenders with ASPD + P, compared with those with ASPD – P. Oxytocin abolished these group differences. This represents evidence of neurochemical modulation of the empathic processing of others' distress in psychopathy.

A small group of men engage in a life-course-persistent pattern of antisocial behavior[1]. These men are disproportionally responsible for violent crimes[2], resulting in considerable personal and societal costs[3]. They meet diagnostic criteria for conduct disorder in childhood and antisocial personality disorder (ASPD) in adulthood. However, there is considerable heterogeneity within this group. Approximately one-third of men with ASPD meet additional diagnostic criteria for psychopathy (ASPD + P)[4]. They exhibit callous-unemotional traits in childhood[5], begin offending at earlier ages and engage in a broader range and greater density of offending behaviors[6] than those without psychopathy (ASPD – P). Importantly, they also respond less well to psychosocial treatment programmes[7]. Clinical guidelines suggest that currently available evidence does not support the use of pharmacological interventions for the treatment of antisocial personality disorder or its associated behaviors of aggression, anger and impulsivity[8].

Abnormalities in reinforcement-based decision making and emotional (particularly empathic) responsiveness may help to explain the behaviors of these violent offenders. Decision-making abnormalities are observed in antisocial men with and without psychopathy[9] when they undertake tasks in which they must learn which responses to make to gain a reward or to avoid punishment. There are differences in the neural response to unanticipated punishment between those with and without psychopathy[10]. Such impairments may underpin reduced reinforcement sensitivity, resulting in impulsivity, frustration-induced reactive aggression and recidivism[11]. By contrast, deficits in different components of empathic processing appear to be relatively

[1]Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, UK. [2]Academic Unit of Mental Health and Clinical Neurosciences, School of Medicine, University of Nottingham, Jubilee Campus, Nottingham, UK. [3]Centre for Neuroimaging Sciences, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, UK. [4]Child and Adolescent Mental Health Centre, Mental Health Services, Capital Region of Denmark, Copenhagen, Denmark. [5]These authors contributed equally: John Tully, Arjun Sethi. [6]These authors contributed equally: Robert James Blair, Nigel Blackwood. ✉e-mail: john.tully@nottingham.ac.uk

specific to antisocial men with psychopathy. Behaviorally, individuals with psychopathy demonstrate impairments when explicitly asked to emotionally 'label' static two-dimensional images of a range of facial emotions[12], most markedly for fear and sadness[12,13]. Functional magnetic resonance imaging (fMRI) work indicates altered neural responses to empathy-inducing pictures of physically painful situations in individuals with psychopathy[14]. A reduced ability to recognize and respond to another's fear, pain and distress may be related to the use of goal-directed instrumental aggression that is particularly prevalent in individuals with psychopathy because the individual is less concerned by the distress of others and less fearful of punishment[15].

However, the exploration of a key aspect of empathic processing— implicit neural responses to others' facial expressions of fear—in violent antisocial men with ASPD, with and without psychopathy, has so far been limited. In healthy individuals, partially separable neural systems are involved in the processing of specific emotions, with prominent roles for the amygdala, insula, and anterior and mid-cingulate cortex in processing fearful expressions[13,16,17] (but see also ref. [18]). Studies in antisocial youth suggest reduced amygdala activity in response to fearful faces in children with high levels of callous-unemotional traits (the developmental precursor of ASPD + P)[19] but increased amygdala activity in those with low levels of callous-unemotional traits (the developmental precursor of ASPD − P)[20], compared with typically developing children. In adults, only two previous studies have been conducted to explore the implicit neural processing of fearful facial expressions in violent antisocial men categorized as ASPD +/− P. In a small pilot study[21], men with ASPD + P (compared with healthy non-offenders) showed significantly reduced activation in the core face-processing network[16] in response to fearful facial expressions at 'low' and 'prototypical' intensities. In a larger study[22], men with ASPD + P (compared with men with ASPD − P) showed reduced activation in the core face-processing network and associated emotional and motivational processing regions (orbitofrontal cortex and ventromedial prefrontal cortex), but increased dorsal insula responses, when passively viewing dynamic facial expressions of fear. However, this study lacked a non-offender control group. Two other studies conducted in samples of violent male offenders assessed the dimensional impact of psychopathic traits on implicit neural responses to emotional faces[23,24]. However, in one, not all violent offenders had ASPD[23], while in the other, participants were not assessed for presence of ASPD[24].

Furthermore, no study to our knowledge has investigated whether group differences in brain activation can be modified by pharmacological agents. One potential agent is oxytocin, a neuropeptide central to the regulation of complex social behaviors. Oxytocin has a key role in social functions such as emotion recognition[25], binding to receptors in social brain regions reported as functionally abnormal in ASPD (such as the amygdala and cingulate cortex[26]). In healthy individuals, oxytocin enhances the explicit emotional recognition of fearful faces[27] and significantly impacts activity within fearful facial-processing regions, including the amygdala, insula and anterior cingulate cortex[28]. In antisocial adults, a small behavioral study[29] has suggested that a single dose of 24 IU intranasal oxytocin can improve fearful expression recognition, at least in the short term. However, no previous work has examined the neural basis of this potential effect.

Hence, we carried out a double-blind, placebo-controlled, randomized crossover study in male violent offenders with ASPD + P and ASPD − P and healthy non-offenders to explore the impact of oxytocin on brain functional differences when implicitly processing others' distress in the form of fearful facial expressions at varying intensities. Considering previous studies investigating processing of facial expressions of fear and studies investigating regional effects of oxytocin, we hypothesized that (1) violent offenders with ASPD + P would show reduced modulation by fearful expression intensity within amygdala, anterior insula and bilateral anterior/mid-cingulate cortices compared with both violent offenders with ASPD − P and healthy non-offenders

(NO) under placebo and (2) intranasal oxytocin administration would reduce group differences in neural responding to fearful expression intensity within these regions.

Results

Behavioral data
Across the whole sample, participants successfully performed the gender rating task (mean accuracy 96.7% (s.d. = 1.0), mean response latency = 938.8 milliseconds (s.d. = 183.3)). For accuracy, there were no significant effects of group (NO, ASPD − P or ASPD + P; $\eta_P^2 = 0.071$, $F_{2,55} = 2.114$, $P = 0.13$), condition (placebo or oxytocin, $\eta_P^2 = 0.017$, $F_{1,42} = 0.553$, $P = 0.333$) or intensity of emotion (40%, 60%, 80% or 100%; $\eta_P^2 = 0.33$, $F_{1,42} = 1.899$, $P = 0.174$). For response latency, there were no significant effects of group ($\eta_P^2 = 0.41$, $F_{1,42} = 1.174$, $P = 0.317$), condition (placebo or oxytocin, $\eta_P^2 = <0.001$, $F_{1,42} = 0.001$, $P = 0.981$) or intensity of emotion ($\eta_P^2 = 0.38$, $F_{1,42} = 2.17$, $P = 0.146$). No group-by-intensity interactions for accuracy or response latency were observed.

fMRI results
Parametric modulation of neural responses by emotion intensity. The main 3dMVM (analysis of variance-style computations) analysis revealed significant activation of the right middle occipital gyrus, involving the primary visual cortex, extending into the right fusiform gyrus; the left middle occipital gyrus, extending into the left fusiform gyrus; and a separate region within the left fusiform gyrus, associated with modulation by fearful expressions (further details in Supplementary Information section 6).

Group differences in responses to modulated fearful expressions. In the three-group 3dMVM regions of interest (ROI) analyses for fearful expressions, there was an overall effect of group in the right and left mid-cingulate cortex (Table 1). Exploratory post hoc between-group analyses revealed four key findings:

- Violent offenders with ASPD + P showed reduced modulation of blood-oxygen-level-dependent (BOLD) responding by fearful expression intensity within bilateral mid-cingulate and right anterior insula compared with the group of violent offenders with ASPD − P under placebo conditions (Fig. 1a,b). The right anterior insula finding did not survive strict correction for multiple comparisons.
- Violent offenders with ASPD + P showed significant increases in modulation by fearful expression intensity in bilateral mid-cingulate cortex and left anterior insula under the oxytocin relative to the placebo condition (Fig. 2a,b).
- There were thus no group differences between ASPD + P and ASPD − P in the oxytocin condition; that is, differences under placebo were abolished under oxytocin. This was supported by a statistically significant group (ASPD − P, ASPD + P) by condition (placebo, oxytocin) interaction effect in left mid-cingulate cortex (Supplementary Figs. 3 and 4).
- There were no significant main effects or between- or within-group effects findings in amygdala in either placebo or oxytocin condition.

Discussion
We investigated the neural basis of implicit fearful facial emotion processing in violent offenders with antisocial personality disorder with and without psychopathy and the effect of intranasal oxytocin on brain functional differences. Offenders with ASPD + P displayed reduced modulation by fearful expression intensity in the anterior insula and mid-cingulate cortex (but not the amygdala) compared with offenders with ASPD − P. Oxytocin abolished differences in fear-associated activity within bilateral mid-cingulate cortex for the offenders with ASPD + P.

Table 1 | Significant BOLD responses to modulated fearful expressions (covaried for active substance misuse)

Region	BA	Voxels	X (left to right)	Y (posterior to anterior)	Z (inferior to superior)	Statistic[a]	P
Overall group effect							
Right mid-cingulate gyrus	24	29	+13.5	−4.5	−32.5	9.464 (F)	0.0003
Left mid-cingulate gyrus	23/24	28	−10.5	−10.5	−35.5	9.135 (F)	0.0003
ASPD+P<ASPD−P (placebo condition)							
Left mid-cingulate gyrus	23/24	109	−7.5	−13.5	−35.5	3.95	0.0002
Right mid-cingulate gyrus	24	92	+10.5	−4.5	−32.5	4.115	0.0001
Right anterior insula[b]	13	6	+40.5	−4.5	−17.5	3.045	0.003
Effect of oxytocin in ASPD+P (oxytocin>placebo)							
Right mid-cingulate gyrus	24	173	+10.5	+1.5	−26.5	4.033	0.0001
Left mid-cingulate gyrus	24	119	−10.5	+10.5	−29.5	4.074	0.0001
Left anterior insula	13	15	−34.5	+4.5	−14.5	3.498	0.0009
Group (ASPD+P versus ASPD−P)×condition (placebo versus oxytocin) interaction effect							
Left mid-cingulate gyrus	23/24	16	−7.5	−13.5	−32.5	3.21	0.0022

Hypothesized regions first, then ordered by cluster size. Significance threshold set at $P \leq 0.005$; all findings significant after cluster-wise correction for multiple comparisons unless noted otherwise. BA, Brodmann area. [a]Statistic refers to t-tests unless stated otherwise. [b]Findings not significant after cluster-wise correction for multiple comparisons.
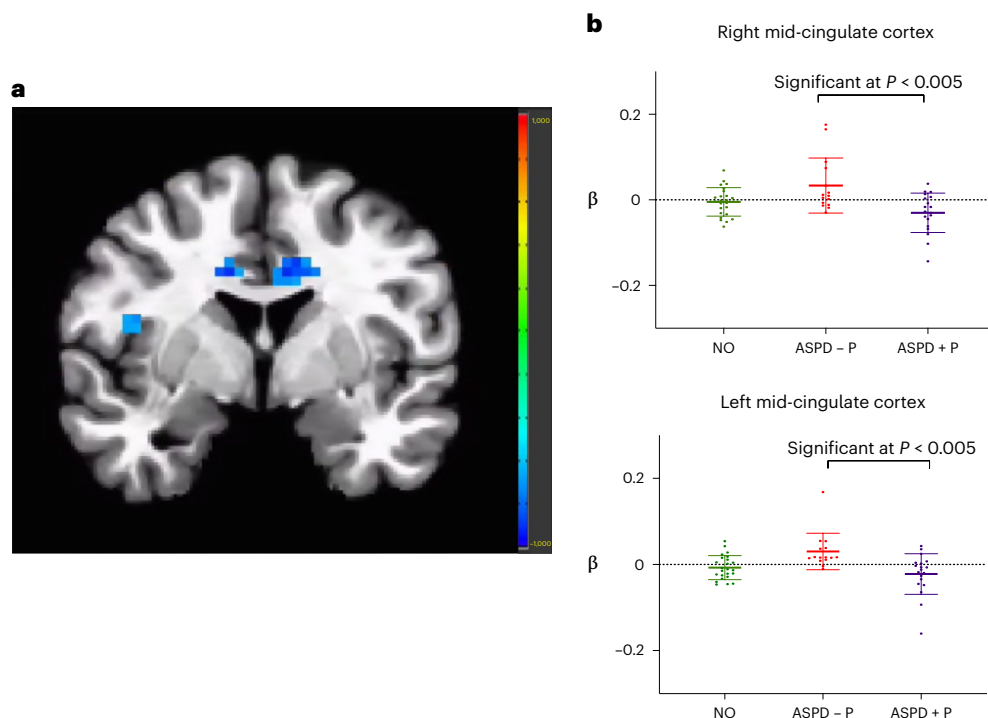

Fig. 1 | Group differences in placebo condition. a, Reduced modulation by fearful intensity in bilateral mid-cingulate cortex and right anterior insula in violent offenders with ASPD + P ($n = 19$) compared with violent offenders with ASPD − P ($n = 15$), placebo condition, $P = 0.005$. Color bar represents t statistic. **b**, Individual beta values for fear processing (modulated fear regressor) in bilateral mid-cingulate cortex, placebo condition, $P = 0.005$. Individual participants' data plotted as dots. Means are indicated by horizontal bars. Error bars represent standard deviations. Findings in insula did not survive multiple comparison corrections. For NO, $n = 24$.

The identification of reduced anterior insula and mid-cingulate cortex reactivity to implicit processing of facial expressions of fear in ASPD + P relative to ASPD − P is broadly consistent with the previous literature[19–22]. The anterior insula is critical in representing the salience of stimuli[30]. It generates an integrated awareness of one's cognitive, affective and physical state that becomes re-represented in the anterior cingulate cortex to facilitate homeostatic autonomic and behavioral responses[31]. The mid-cingulate cortex is a key part of reactive fear
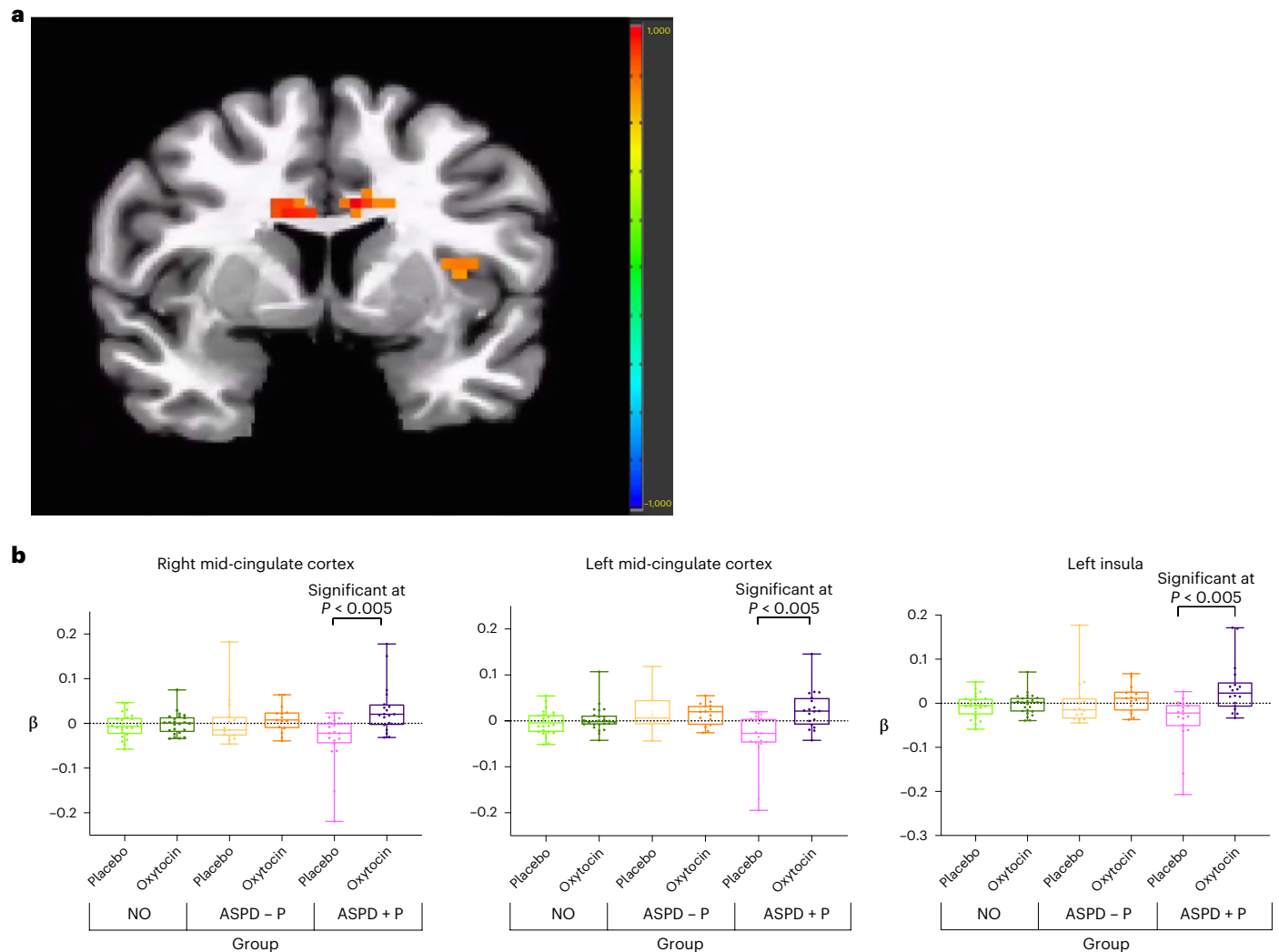
Fig. 2 | **Effect of oxytocin in ASPD + P. a**, Increased modulation by fearful intensity in bilateral mid-cingulate cortex and left insula in oxytocin relative to placebo condition in violent offenders with ASPD+P ($n = 19$), $P = 0.005$. Color bar represents *t* statistic. **b**, Individual beta values for fear processing (modulated fear regressor) for contrast oxytocin > placebo in bilateral mid-cingulate cortex and left insula in violent offenders with ASPD + P ($n = 19$). Individual participants' data plotted as dots. Bounds of upper whiskers represent maxima; bounds of lower whiskers represent minima; center box represents interquartile range, with mean represented by middle horizontal lines.

circuitry, helping to inform rapid decisions of escape from predators, which may be signaled by the fearful face of a conspecific[32]. It appears to coordinate emotional responses and motor actions according to learned values, particularly when a predatory threat is near[32]. Especially robust links have been demonstrated between activity in the anterior subdivision of the mid-cingulate cortex and the experience of more-intense states of negative affect[32], including fear[33]. The posterior mid-cingulate cortex may play a more specific role in threat appraisal and risk assessment by approaching the threat[34]. Our significant mid-cingulate cluster encompassed both anterior mid-cingulate cortex and posterior mid-cingulate cortex, suggesting that the impaired processing of fear in ASPD + P may be related to deficits in both responsivity to intensity and threat appraisal. Relative functional deficits in these key areas of the fear processing network in ASPD + P is in keeping with a model whereby impairment in the ability to recognize and integrate distress cues (such as fear in others) predisposes such individuals to especially pronounced aggressive behavior[15].

Our findings on the effect of oxytocin demonstrate for the first time, to our knowledge, that neural processing abnormalities in a robustly classified group of men with ASPD + P may be modified by neurochemical intervention. Oxytocin resulted in increased modulation

by fearful expression intensity in ASPD + P in left anterior insula and bilateral mid-cingulate cortex. These effects resulted in the baseline differences between ASPD + P and ASPD – P in the implicit processing of others' fear being abolished. The enhancement of fear-associated activity in these regions suggests that the fearful faces are accorded increased salience under the influence of oxytocin, with potential 'downstream' behavioral impacts[35,36]. The observed short-term normalization in the empathic processing of others' distress should prompt further investigation into the neurochemical modulation of the social cognitive abnormalities in this disorder that has such a profound personal and societal impact.

Some caveats should be noted in relation to our fMRI findings. First, the finding of reduced anterior insula reactivity under placebo in ASPD + P compared with ASPD – P in the right insula did not survive strict correction for multiple comparisons, and this finding therefore requires replication in a larger sample. Second, although our baseline finding was in the right insula, our finding of increased modulation by fearful expression intensity in ASPD + P following administration of oxytocin was observed in the left insula. Inconsistencies in findings regarding hemispheric laterality of effects of oxytocin in fMRI studies[28] have been discussed in depth elsewhere, and as a result, our finding of

a unilateral effect of oxytocin should be interpreted with some caution. Third, our findings in mid-cingulate cortex contained white matter. In fMRI analyses that have low spatial/anatomical resolution and a high degree of spatial smoothing, some degree of gray matter activation being localized within 'white matter' as represented by sharply defined high-resolution templates is relatively common and not necessarily a sign of noise. The shape of the cingulate cortex (which loops around the surrounding white matter) makes it particularly prone to this. Although we have taken several steps to remove noisy signal arising from motion that would result in spurious activations within white matter (Supplementary Information section 3), it is possible that our findings are affected by noise, for example, due to motion.

Furthermore, while there were significant differences between ASPD + P and ASPD – P at baseline (placebo) in insula and mid-cingulate cortex, there were not significant differences between either group and NO, although the pattern of parameter estimates in the mid-cingulate cortex and insula (ASPD – P activity > NO activity > ASPD + P activity) were in the expected direction. As such, we cannot be certain that either antisocial group showed significantly atypical face responsiveness compared with NO. However, it is important to note that the ASPD + P group showed a significantly greater response to oxytocin than did the ASPD – P group within mid-cingulate cortex and left insula; that is, oxytocin selectively increased the response within this these regions to increasing fearful facial expression intensity. Moreover, previous fMRI studies in youth with conduct disorder ± callous-unemotional traits[37] and adults with ASPD ± P[10] have demonstrated a similar pattern to that in our study: significant differences between the two antisocial groups but not between the antisocial groups and a normal control group. This likely reflects the use of underpowered samples in a difficult-to-recruit clinical population[38–41]. This also applies to the lack of significant findings in the amygdala in between- and within-group analyses, in both placebo and oxytocin conditions (see Supplementary Information for a further consideration of this negative finding). These factors suggest caution is required in interpretation of both the positive and negative findings in our study and further emphasize the importance of larger-scale collaborative projects to address such issues.

Some other limitations should also be noted. While both ASPD groups had similar lifetime histories of substance misuse, there were some differences in substance misuse measured on the day of scanning (for example, more of the ASPD + P group had recently consumed cocaine than had the ASPD – P group). However, such group differences were carefully controlled for in the imaging analyses, and the observed group differences at baseline and in response to oxytocin cannot be attributed simply to such differences. Finally, while our study benefited from use of a task measuring implicit responding to facial expressions of fear (see Supplementary Information for further discussion), we did not examine any potential impact on observable behaviors. The potential clinical significance of our findings therefore requires further specification.

This study also had important strengths. First, this study in a carefully characterized group of participants with ASPD established differences between antisocial groups in adulthood in the implicit processing of fearful facial expressions, a central aspect of empathic responding. Second, this study investigated the neural effects of oxytocin in this group, achieved using a randomized, placebo-controlled method. Further, diagnoses and Psychopathy Checklist–revised (PCL-R) ratings were made by trained clinicians, with use of official criminal records to help classify participants.

In conclusion, we have demonstrated that the implicit processing of fearful facial emotion expressions significantly differs between violent antisocial male offenders with and without psychopathy. Oxytocin abolished these group differences in the mid-cingulate cortex. This represents evidence of neurochemical modulation of the empathic processing of others' distress in psychopathy. Neurochemical modulation of core deficits in the condition could have profound implications for treatment of this complex disorder.

Methods

Participants and assessment

Between September 2017 and March 2020, we enrolled 58 men, aged 20 to 58 years, with an IQ greater than 70 as defined by the Wechsler Abbreviated Scale of Intelligence[41]. Offenders with convictions for violent crimes (murder, rape, attempted murder, grievous and actual bodily harm) who met *Diagnostic and Statistical Manual of Mental Disorders* 5th Edition (DSM-5) criteria for antisocial personality disorder (Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5 PD[42]) were recruited via the National Probation Service of England and Wales and local forensic personality disorder services. Healthy non-offenders were recruited from the general population using online advertisements and fliers in job centers and local recreational centers. Offenders and controls were recruited in parallel and interlaced in the study protocol sessions outlined in the following. All participants completed diagnostic (Structured Clinical Interview for DSM-5 Personality Disorders[42]) and PCL-R[43] interviews and authorized access to their criminal records. A cross-cultural validation study[44] of the PCL-R demonstrated that cut-off scores for psychopathy in men vary between North America (30 out of a possible 40 points) and Europe (25 out of a possible 40 points). In line with previous research in UK samples[9,10], we used a score of 25 as the threshold for psychopathy in this English population. We calculated total, factor 1 and factor 2 PCL-R scores for all participants. Factor 1 scores are a total of facet 1 (interpersonal traits, such as pathological lying) plus facet 2 traits (affective traits, such as lack of empathy), while factor 2 scores are a total of facet 3 (antisocial lifestyle traits, such as impulsivity) plus facet 4 traits (overt antisocial behaviors, such as criminal versatility). Exclusion criteria were history of major mental disorder (bipolar 1, bipolar 2, major depression or psychotic disorders) or self-reported neurological disorder, head injury resulting in loss of consciousness for 1 h or longer, severe visual or hearing impairment or contraindication to MRI.

After receiving a complete description of the study, all participants provided written consent. Ethical approval was obtained from the national UK research authority (National Health Service Health Research Authority Research and Ethics Committee, project number 15/LO/1083). All assessments were conducted by an experienced forensic psychiatrist (J.T.). Participants completed the Reactive–Proactive Aggression Questionnaire (RPQ)[45]. On the day of each MRI scan, participants provided a urine sample to assess for substance misuse. Following psychometric assessments, only participants who attended for two MRI sessions were included in the analyses.

The three groups did not differ significantly except for years of education and PCL-R total and facet scores (Table 2). The offenders with ASPD + P had significantly higher proactive, reactive and total aggression scores than offenders with ASPD – P. Cronbach's alpha for internal consistency of items was 0.79 for PCL-R and 0.91 for RPQ (see Supplementary Tables 1 and 2 for inter-item correlations). The offenders with ASPD + P also had a significantly higher rate of comorbid Cluster A personality disorder diagnosis compared with NO. Offender groups (with and without psychopathy) had a significantly higher rate of comorbid Cluster B personality disorder diagnosis compared with NO. Both offender groups also had a significantly higher rate of lifetime substance misuse disorders than did NO; however, the proportion of offenders with and without psychopathy with lifetime substance-use disorders did not differ. Urinary drug screening on the day of scanning revealed some significant differences in active illicit substance misuse (Supplementary Table 1); therefore, this was included as a covariate in fMRI analysis.

Study design and procedures

In a double-blind, placebo-controlled, randomized crossover design, participants self-administered, under instruction from the researcher,

Table 2 | Demographic and clinical characteristics

Demographic/clinical characteristic	Group			Group comparison		Post hoc tests (P values)		
	NO (n=24)	ASPD−P (n=15)	ASPD+P (n=19)	Statistic	P value	Control versus ASPD−P	Control versus ASPD+P	ASPD−P versus ASPD+P
Age (years)	37.6 (10.3)	40.9 (9.6)	38.7 (9.1)	0.52	0.59	0.92	1.0	1.0
Ethnicity (non-white)	7 (29.1%)	4 (26.6%)	5 (26.3%)	0.13	1.0	1.0	1.0	1.0
IQ	100.9 (12.6)	97.6 (16.7)	91.8 (11.3)	2.31	0.11	1.0	0.11	0.7
Duration of education (years)	13.9 (3.2)	10.8 (2.2)	10.0 (2.0)	13.07	<0.001**	0.002**	<0.001**	1.0
Age at first violent conviction	n/a	20.6 (6.2)	18.5 (5.5)	0.97	0.34	n/a		
Number of violent convictions	n/a	3.7 (3.1)	4.7 (2.9)	−0.87	0.39			
RPQ Reactive Aggression	6.0 (4.0)[a]	12.4 (4.8)[a]	17.2 (4.5)[a]	29.28	<0.001**	<0.001**	<0.001**	0.013*
RPQ Proactive Aggression	1.0 (1.3)[a]	7.8 (6.3)[a]	14.2 (6.2)[a]	33.71	<0.001**	0.003	<0.001**	0.016*
RPQ Total Aggression	6.7 (5.0)[a]	20.2 (10.8)[a]	31.4 (10.2)[a]	37.61	<0.001**	0.001	<0.001**	0.012*
PCL-R Facet 1 (Interpersonal)	0.75 (0.98)	2.02 (1.64)	4.42 (1.86)	32.25	<0.001**	0.038**	0.001**	0.001**
PCL-R Facet 2 (Affective)	0.75 (0.94)	3.13 (1.64)	5.42 (1.95)	50.40	<0.001**	0.001**	<0.001**	0.001**
PCL-R Facet 3 (Lifestyle)	1.16 (1.63)	5.46 (1.68)	7.52 (1.17)	99.15	<0.001**	<0.001**	<0.001**	0.001**
PCL-R Facet 4 (Antisocial)	0.91 (1.72)	6.33 (2.16)	8.63 (1.11)	117.96	<0.001**	<0.001**	<0.001**	<0.001**
PCL-R Total	2.70 (2.74)	18.41 (3.49)	28.84 (3.14)	391.56	<0.001**	<0.001**	<0.001**	<0.001**
Personality disorder other than ASPD (%)								
Cluster A	0	6.7	31.5	12.97	0.005**	0.38	0.0045**	0.10
Cluster B	0	20	36.8	10.19	<0.006**	0.016*	0.0016**	0.46
Cluster C	0	13.3	5.2	3.34	0.19	0.57	0.33	0.44
Lifetime substance-use disorder (%)	8.3	33.3	21	3.83	0.15	0.047*	0.23	0.42

Group data are mean (s.d.) unless otherwise stated. 'Statistic' refers to $F_{(2,55)}$ for continuous variables in three-group analyses; chi-squared/Fisher's exact test for categorical data. All statistical tests were two-sided. [a]Not all participants completed the RPQ: NO, $n=21$; ASPD−P, $n=12$; ASPD+P, $n=15$. *Statistically significant at $P<0.05$. **Statistically significant at $P<0.01$.

40 IU of IN-OT (Syntocinon; Novartis) or placebo (identical composition to Syntocinon except for the omission of oxytocin). Participants began the morphed faces task within 25–30 min of administration. The oxytocin dose used was the highest clinically applicable safe dose administered to human volunteers, in keeping with a protocol that demonstrated significant neural activation over a period of 25–78 min with this dose[46]. Further discussion about the dose and timing of intranasal oxytocin, and mechanism for delivery to the brain, is included in Supplementary Information section 1. At a second session (occurring between 3 and 28 days later), participants completed the fMRI task again under the alternative treatment condition. Participants were instructed to avoid food, drinks (except water) and nicotine two hours before starting the experiment. Participants completed the Morphed Faces task (Supplementary Information section 1). During the task, participants were asked to indicate the sex of each face with a left–right button press using the index and middle finger of their right hand during a single run, which lasted 9 min 56 s. Full description of the Morphed Faces task, together with information on data quality control and motion parameters, is available in Supplementary Information section 1. No participant reported adverse effects following the intranasal spray.

This trial was registered at ClinicalTrials.gov (ID NCT05383300). CONSORT guidelines were adhered to, and the CONSORT statement is available in Supplementary Information. Allocation of oxytocin or placebo was randomized in advance with randomization generator software that used permuted blocks of six (http://www.randomizer.org). Intranasal oxytocin spray (Syntocinon; Novartis) and placebo were provided and administered by Maudsley Pharmacy. The active intranasal oxytocin spray was an identical formulation to the oxytocin spray, except for no active agent. A kit was allocated to each participant. Bottles for oxytocin and placebo were identical, except randomly marked A or B. Each participant was first allocated spray 'A' and then spray 'B' on the subsequent visit. On the day of collection, only pharmacy staff knew which spray had been provided, and this was recorded in their records. Both the researcher and the participant were blinded to the formulation of the spray being used. At the end of the experiment, researchers were unblinded following instruction to Maudsley Pharmacy from the study principal investigator.

The target sample size was based on a previous computational study[47]. This showed that at thresholds that approach those used after correcting for multiple comparisons, about 24 participants are required to achieve 80% power at the single-voxel level for typical activations (an 80% true positive rate) in fMRI studies.

General linear model analysis of behavioral data

For the Morphed Faces task, means were first calculated across the whole sample for both accuracy and reaction time in rating the gender of the faces displayed. To investigate the effect of oxytocin and its interaction with other variables, for both accuracy and response latency data,

a three-group (NO, ASPD − P and ASPD + P) by two-condition (oxytocin and placebo) by four-intensity (40%, 60%, 80% and 100% of fearful facial expression) repeated-measures analysis of variance was conducted. Post hoc repeated-measures analysis of variance was performed for ASPD − P versus ASPD + P. Statistical Package for the Social Sciences version 25.0 was used. A threshold for significance of $P < 0.05$ was set for all tests.

Primary outcome measure and MRI processing
Whole-brain BOLD fMRI data were acquired using a 3.0 Tesla General Electric magnetic resonance scanner. The principle outcome measure was a regressor for modulation of neural activity (BOLD responsivity) by intensity of fearful expression. Specific MRI parameters, and full details of preprocessing and individual-level analyses and data quality control, are available in Supplementary Information sections 3 and 4.

MRI data-group analysis
Following preprocessing steps, modulated emotion data were entered into a three-group (NO, ASPD − P and ASPD + P) by two-condition (placebo and oxytocin) 3dMVM (analysis of variance-style computations) model. Within this framework, general linear tests were coded to assess differential effects of the drug between the groups. Post hoc t tests were conducted to decompose these interactions by examining between- and within-group effects. Correction for multiple comparisons was performed using a spatial clustering operation in the AFNI (Analysis of Functional NeuroImages) 3dClustSim utilizing the autocorrelation function (-acf) with 10,000 Monte Carlo simulations for the whole-brain analysis. Spatial autocorrelation was estimated from residuals from the individual-level general linear models. The initial threshold was set at $P = 0.005$. As outlined in the preceding, bilateral amygdala, anterior insula and mid-cingulate cortex were selected a priori for ROI analysis. Small-volume corrections, calculated using an anatomically defined mask (TT_Daemon, a Talaraich atlas from AFNI), yielded thresholds of $k = 13$ for anterior/mid-cingulate cortex, $k = 8$ for anterior insula and $k = 2$ for amygdala at an initial significance threshold of 0.005 (multiple comparison corrected $P < 0.05$).

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
All data are available from the authors upon reasonable request.

Code availability
All neuroimaging data were processed using AFNI code (available at https://afni.nimh.nih.gov/download).

References
1. Moffitt, T. E. Male antisocial behaviour in adolescence and beyond. *Nat. Hum. Behav.* **2**, 177–186 (2018).
2. Falk, Ö. et al. The 1% of the population accountable for 63% of all violent crime convictions. *Soc. Psychiatry Psychiatr. Epidemiol.* **49**, 559–571 (2014).
3. Wickramasekera, N., Wright, J., Elsey, H., Murray, J. & Tubeuf, S. Cost of crime: a systematic review. *J. Crim. Justice* **43**, 218–228 (2015).
4. Coid, J. & Ullrich, S. Antisocial personality disorder is on a continuum with psychopathy. *Compr. Psychiatry* **51**, 426–433 (2010).
5. Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R. & Stouthamer-Loeber, M. Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *J. Abnorm. Psychol.* **116**, 155–165 (2007).
6. Kosson, D. S., Lorenz, A. R. & Newman, J. P. Effects of comorbid psychopathy on criminal offending and emotion processing in male offenders with antisocial personality disorder. *J. Abnorm. Psychol.* **115**, 798–806 (2006).
7. Guy, L. S., Edens, J. F., Anthony, C. & Douglas, K. S. Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *J. Consult. Clin. Psychol.* **73**, 1056 (2005).
8. NCCFMH *Antisocial Behaviour and Conduct Disorders in Children and Young People: Recognition, Intervention and Management* (RCPsych Publications, 2013).
9. De Brito, S. A., Viding, E., Kumari, V., Blackwood, N. & Hodgins, S. Cool and hot executive function impairments in violent offenders with antisocial personality disorder with and without psychopathy. *PloS ONE* **8**, e65566 (2013).
10. Gregory, S. et al. Punishment and psychopathy: a case-control functional MRI investigation of reinforcement learning in violent antisocial personality disordered men. *Lancet Psychiatry* **2**, 153–160 (2015).
11. Bernasco, W., Elffers, H. & van Gelder, J.-L. *The Oxford Handbook of Offender Decision Making* (Oxford Univ. Press, 2017).
12. Dawel, A., O'Kearney, R., McKone, E. & Palermo, R. Not just fear and sadness: meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neurosci. Biobehav. Rev.* **36**, 2288–2304 (2012).
13. Marsh, A. A. & Blair, R. Deficits in facial affect recognition among antisocial populations: a meta-analysis. *Neurosci. Biobehav. Rev.* **32**, 454–465 (2008).
14. Decety, J., Chen, C., Harenski, C. & Kiehl, K. A. An fMRI study of affective perspective taking in individuals with psychopathy: imagining another in pain does not evoke empathy. *Front. Hum. Neurosci.* **7**, 489 (2013).
15. Blair, R. Emotion-based learning systems and the development of morality. *Cognition* **167**, 38–45 (2017).
16. Fusar-Poli, P. et al. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.* **34**, 418–432 (2009).
17. Tettamanti, M. et al. Distinct pathways of neural coupling for different basic emotions. *Neuroimage* **59**, 1804–1817 (2012).
18. Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E. & Barrett, L. F. The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* **35**, 121 (2012).
19. Marsh, A. A. et al. Reduced amygdala response to fearful expressions in children and adolescents with callous–unemotional traits and disruptive behavior disorders. *Am. J. Psychiatry* **165**, 712–720 (2008).
20. Sebastian, C. et al. Neural responses to fearful eyes in children with conduct problems and varying levels of callous–unemotional traits. *Psychol. Med.* **44**, 99–109 (2014).
21. Deeley, Q. et al. Facial emotion processing in criminal psychopathy. Preliminary functional magnetic resonance imaging study. *Br. J. Psychiatry* **189**, 533–539 (2006).
22. Decety, J., Skelly, L., Yoder, K. J. & Kiehl, K. A. Neural processing of dynamic emotional facial expressions in psychopaths. *Soc. Neurosci.* **9**, 36–49 (2014).
23. da Cunha-Bang, S., Fisher, P. M., Hjordt, L. V., Holst, K. & Knudsen, G. M. Amygdala reactivity to fearful faces correlates positively with impulsive aggression. *Soc. Neurosci.* **14**, 162–172 (2019).
24. Pardini, D. A. & Phillips, M. Neural responses to emotional and neutral facial expressions in chronically violent men. *J. Psychiatry Neurosci.* **35**, 390–398 (2010).
25. McCall, C. & Singer, T. The animal and human neuroendocrinology of social cognition, motivation and behavior. *Nat. Neurosci.* **15**, 681–688 (2012).
26. Boccia, M., Petrusz, P., Suzuki, K., Marson, L. & Pedersen, C. Immunohistochemical localization of oxytocin receptors in human brain. *Neuroscience* **253**, 155–164 (2013).

27. Shahrestani, S., Kemp, A. H. & Guastella, A. J. The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: a meta-analysis. *Neuropsychopharmacology* **38**, 1929–1936 (2013).

28. Tully, J., Gabay, A. S., Brown, D., Murphy, D. G. & Blackwood, N. The effect of intranasal oxytocin on neural response to facial emotions in healthy adults as measured by functional MRI: a systematic review. *Psychiatry Res. Neuroimaging.* **272**, 17–29 (2018).

29. Timmermann, M. et al. Oxytocin improves facial emotion recognition in young adults with antisocial personality disorder. *Psychoneuroendocrinology* **85**, 158–164 (2017).

30. Uddin, L. Q. Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* **16**, 55–61 (2015).

31. Fullana, M. et al. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol. Psychiatry* **21**, 500–508 (2016).

32. Shackman, A. J. et al. The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* **12**, 154–167 (2011).

33. Tolomeo, S. et al. A causal role for the anterior mid-cingulate cortex in negative affect and cognitive control. *Brain* **139**, 1844–1854 (2016).

34. Blanchard, T. C., Strait, C. E. & Hayden, B. Y. Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *J. Neurophysiol.* **114**, 2439–2449 (2015).

35. Menon, V. & Uddin, L. Q. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).

36. Newman, J. P., Curtin, J. J., Bertsch, J. D. & Baskin-Sommers, A. R. Attention moderates the fearlessness of psychopathic offenders. *Biol. Psychiatry* **67**, 66–70 (2010).

37. Poldrack, R. A. et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* **18**, 115–126 (2017).

38. Cremers, H. R., Wager, T. D. & Yarkoni, T. The relation between statistical power and inference in fMRI. *PloS ONE* **12**, e0184923 (2017).

39. Turner, B. O., Paul, E. J., Miller, M. B. & Barbey, A. K. Small sample sizes reduce the replicability of task-based fMRI studies. *Commun. Biol.* **1**, 62 (2018).

40. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).

41. Wechsler, D. & Hsiao-Pin, C. *WASI II: Wechsler Abbreviated Scale of Intelligence* 2nd edn (Psychological Corporation, 2011).

42. First, M. B., Williams, J. B., Benjamin, L. S. & Spitzer, R. L. *SCID-5-PD: Structured Clinical Interview for DSM-5 Personality Disorders* (APA, 2016).

43. Hare, R. D. *The Psychopathy Checklist–Revised* (Multi-Health Systems, 2003).

44. Cooke, D. J. & Michie, C. Psychopathy across cultures: North America and Scotland compared. *J. Abnorm. Psychol.* **108**, 58–68 (1999).

45. Raine, A. et al. The reactive–proactive aggression questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggress. Behav.* **32**, 159–171 (2006).

46. Paloyelis, Y. et al. A spatiotemporal profile of in vivo cerebral blood flow changes following intranasal oxytocin in humans. *Biol. Psychiatry* **79**, 693–705 (2016).

47. Desmond, J. E. & Glover, G. H. Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *J. Neurosci. Methods* **118**, 115–128 (2002).

Acknowledgments

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44220-023-00067-3.

Correspondence and requests for materials should be addressed to John Tully.

Peer review information *Nature Mental Health* thanks Ilse van de Groep and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature portfolio

Corresponding author(s):   John Tully

Last updated by author(s):   240223

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Psychometric data was entered manually into SPSS and Excel spreadsheets. Neuroimaging data was stored and accessed via centralised database (NaN), at Centre of Neuroimaging Sciences, Institute of Psychiatry, Psychology and Neuroscience, Kings College London.
Data analysis	Psychoemtric analysis was done with SPSS version 25. Neuroimaging analysis was done with AFNI Version 19.0.9. All neuroimaging data was processed using AFNI code (available at https://afni.nimh.nih.gov/download).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data is available from the authors upon reasonable request. All neuroimaging data was processed using AFNI code (available at https://afni.nimh.nih.gov/download).

Human research participants

Reporting on sex and gender	Sex is specified in tables and results
Population characteristics	Yes, all data available on basic demographics, violent offending, and clinical diagnoses reported
Recruitment	Detailed in 'Methods- Participants and assessment'
Ethics oversight	Detailed in 'Methods- Participants and assessment'

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Randomised, double-blinded, placebo-controlled, crossover study
Research sample	Adult male violent offenders with antisocial personality disorder +/- psychopathy, healthy non-offending adult male controls
Sampling strategy	Offenders recruited from probation and clinical services, controls recruited from the community by adverts
Data collection	Psychometric data collected manually by researchers and entered onto Excel and SPSS spreadsheets. MRI data automatically collected and stored on centalised database (NaN) at Centre of Neuroimaging Sciences, Institute of Psychiatry, Psychology and Neuroscience, Kings College London.
Timing	Samples recruited between September 2017 and March 2020
Data exclusions	A small number of subjects dropped out or did not generate sufficient neuroimaging data, as detailed in Supplementary Materials, Figures S1-S3.
Non-participation	n/a
Randomization	Detailed in 'Study Design and Procedures'- Allocation of oxytocin or placebo was randomised in advance with randomisation generator software that used permuted blocks of six (http://www.randomizer.org).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	This trial was registered at ClinicalTrials.gov (ID NCT05383300).
Study protocol	Available at clinicaltrials.gov: https://clinicaltrials.gov/ct2/show/NCT05383300?term=NCT05383300&draw=2&rank=1
Data collection	Detailed in METHODS. Clinical data based on interviews, neuroimaging data generated from MRI scanning sessions.
Outcomes	Detailed in METHODS. For key neuroimaging outcomes measures, general linear tests were coded to assess differential effects of drug between the groups. Post hoc t-tests were conducted to decompose these interactions by examining between- and within-group effects.