BJPsych The British Journal of Psychiatry (2026)
228, 6, 530–538. doi: 10.1192/bjp.2026.10605

Original Article

Effects of opioid antagonism on functional MRI correlates of explicit and implicit threat processing in healthy volunteers

Nathan T. M. Huneke, Henk van Steenbergen, Harry A. Fagan, Laura Molteni, Algirdas Midveris, Naomi Phillips, Angela Darekar, Nic J. A. van der Wee, Matthew Garner and David S. Baldwin

Background

We need to identify novel, tractable therapeutic targets for anxiety disorders. Converging evidence suggests the endogenous opioid system plays a role in modulating affective processing, but its contribution to regulation of threat processing in humans remains unclear.

Aims

We investigated the neural correlates of non-specific opioid antagonism on explicit and implicit regulation of threatening stimuli in healthy volunteers, using functional magnetic resonance imaging (fMRI).

Method

In a randomised, double-blind, placebo-controlled, crossover design, 38 healthy participants received the opioid antagonist naltrexone (50 mg) or placebo before completing two tasks during fMRI: (a) a cognitive emotional reappraisal task probing explicit regulation and (b) a face-viewing task probing implicit processing.

Results

Contrary to our hypothesis, we found naltrexone reduced distress ratings during the reappraisal task ($p = 0.044$) without impairing regulation success. Explicit regulation in the reappraisal task engaged lateral prefrontal regions similarly across drug conditions. However, naltrexone attenuated ventromedial prefrontal cortex, thalamus and caudate activation when viewing negative images. Naltrexone additionally altered ventromedial prefrontal cortex activity and in task-positive regions including right premotor area and frontal pole compared with placebo when viewing emotional faces. In particular, naltrexone increased left middle frontal gyrus activity when viewing fearful faces.

Conclusions

Our results support a role for opioid signalling in automatic emotional regulation, but not in explicit regulation. Furthermore, naltrexone appeared to diminish activity in task-positive regions in response to emotional faces. These findings are consistent with a model where endogenous opioids 'fine-tune' affective responses to both negative and positive stimuli. Future research should explore dose–response effects, kappa-opioid contributions and whether similar results are seen in clinical populations.

Keywords

Antianxiety drugs; anxiety or fear-related disorders; neuroimaging; psychopharmacology; neuroscience.

Anxiety disorders are the most common form of mental disorder, with an estimated 12-month prevalence of 8.4–21.3%.[1] They confer a substantial economic burden: estimated costs to the UK are £20 billion annually,[2] and they are the sixth leading cause of non-fatal health loss globally.[3] Current treatments for anxiety disorders are not 'ideal'. Psychotherapeutic interventions are limited in availability, medications cause unwanted side-effects, and first-line interventions often fail, with remission or response rates at only 40–50%.[4,5] There is a need, therefore, to identify new tractable treatment targets for anxiety disorders.

Converging evidence suggests the endogenous opioid system plays a role in emotional and threat processing. The endogenous opioid system encompasses the endogenous peptides beta-endorphin, the enkephalins and dynorphins, and their G-protein-coupled receptors: the mu-, delta- and kappa-opioid receptors; and the non-opioid nociceptin/orphanin FQ receptor.[6] Work in animal models implicates the mu- and delta-opioid systems in stress-aversion and anxiolysis.[7,8] There is also some evidence for a similar role in humans. For example, infusion of 0.4 mg of naloxone (an opioid antagonist) has been shown to significantly reduce the anxiolytic effects of diazepam in patients anticipating minor orthopaedic surgery.[9] In a 2017 study, 43 healthy volunteers underwent an observational fear conditioning procedure after being given the opioid antagonist naltrexone 50 mg or placebo. Amygdala responses during observational learning of threat cues were increased in the naltrexone group, and this correlated positively with long-term fear conditioned memory.[10] Further, subclinical depressive and anxiety symptoms are associated with reduced mu-opioid receptor availability in healthy volunteers.[11] Thus preliminary evidence suggests a role for endogenous opioids in modulating distress and anxiety-related behaviour.

Patients with anxiety disorders exhibit reduced activity in prefrontal emotion regulation circuits and increased reactivity within limbic regions important in emotion generation, including the amygdala, insula and anterior cingulate cortex.[12] The mechanisms that cause this apparent dysregulation of emotional reactivity remain incompletely understood. Given the data above, it is possible that dysregulation of the endogenous opioid system might be involved in dysregulation of threat reactivity in patients with anxiety disorders. Intriguingly, a recent positron emission tomography study in healthy volunteers demonstrated that acute unconditioned fear (exposure to a live snake) reduced endogenous opioid release, and neural correlates of habituation to this stimulus depended on baseline mu-opioid receptor availability.[13] These results imply that endogenous opioids are important in homeostatic regulation of anxiety responses to threatening stimuli. However, the mechanisms mediating this remain unknown.

In this study, we aimed to understand whether endogenous opioids are important in the regulation of anxiety in response to

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

threatening stimuli in healthy volunteers. Emotion regulation can be achieved through conscious, active mechanisms, or through unconscious, automatic mechanisms. To explore endogenous opioids in both processes, we investigated whether naltrexone altered neural activity during an explicit cognitive emotional reappraisal task and during an implicit emotional face-viewing task. We hypothesised that through naltrexone's mu-opioid antagonism, subjective distress would increase compared with placebo, and that this would be associated with increased activity within limbic regions and decreased activity in prefrontal emotion regulation circuits. As far as we are aware, this is the first study to explore the effects of endogenous opioid antagonism on explicit cognitive emotional reappraisal and on implicit emotional processing.

Method

This study was approved by the Ethics and Research Governance Office at the University of Southampton (reference: 53540). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2013. All participants provided written informed consent. Participants and procedures overlap with our previously published manuscript, in which we investigated naltrexone effects on attention following reward receipt, outside the scanner.[14] No analyses or outcomes are duplicated here.

Participants

We recruited 40 healthy volunteers (mean age 24.38±4.88 years, 24 females) through adverts placed on the University of Southampton campus and in the local community. Participants who responded to advert were invited to attend a face-to-face screening interview to determine eligibility. We excluded participants if they had current or lifetime history of psychiatric illness (as assessed by the Mini International Neuropsychiatric Interview for DSM-5 (MINI)[15]), chronic physical illness, alcohol or drug dependence, if body mass index was <18 or $>28\,kg/m^2$, use of medication in the past 8 weeks, regular use of illicit substances, left or mixed handedness (as assessed by the Edinburgh Handedness Inventory – Short Form[16]), and any contraindication to naltrexone or magnetic resonance imaging (MRI) (see Supplementary Material for full list of exclusion criteria, available at: https://doi.org/10.1192/bjp.2026.10605).

Study procedures

We used a randomised, within-individual crossover design with two sessions. Eligible participants were invited to attend University Hospital Southampton, Southampton, UK, for an MRI scan on two occasions. Procedures for each session were identical and took place at the same time of day. On each occasion, participants randomly received either naltrexone 50 mg or a matched placebo capsule in a counterbalanced order under blinded conditions. Participants then waited 1 h for naltrexone to reach peak plasma concentration[17] before entering the scanner. Following the scan (approximately 45 min in duration), participants completed a modified monetary incentive delay task in a quiet room (results published separately).[14] Following each session, participants were debriefed and had their blood pressure and heart rate checked to ensure they were safe to leave. Sessions were scheduled at least 7 days apart to reduce the impact of carry-over effects, although it should be noted more recent work suggests a minimum interval of 15 days ensures complete elimination.[18]

Functional MRI tasks

Participants completed two tasks in the scanner: an emotional reappraisal task and a face-viewing task. Participants interacted with the tasks via button presses on response grips. Tasks are described briefly below; for full details see the Supplementary Material.

Emotional reappraisal task

We used the emotional reappraisal task included in the Netherlands Study of Depression and Anxiety study, published previously.[19] The task comprised five conditions: fearful/happy (regulate/attend), and neutral (attend). Each condition comprised three blocks of four images, presented in pseudo-randomised order such that the task always began and finished with a neutral block, and that two identical blocks were not consecutively presented. Each block began with the instruction presented in the middle of the screen, followed by four successive images shown for eight seconds each, separated by a jittered fixation cross. After each block, participants reported how they were feeling on a visual analogue scale ranging from 'very distressed' to 'very happy' (scored from −50 to +50). After responding, the block ended with a jittered fixation cross. In total, 15 blocks were presented resulting in a total task duration of approximately 16 min.

Face-viewing task

In this task, participants were shown images of actors portraying either fearful, happy or neutral facial expressions from the NimStim set.[20] To probe implicit emotional processing, we asked participants to identify each actor's biological sex as quickly and accurately as possible. Images were presented in a random order within blocks of the same facial expression. During each block, 16 faces were shown for 400 ms each (inter-trial interval of 850 ms). Each block was separated by an 18 s fixation cross followed by a 2 s reminder of the correct button trigger for each biological sex. Each of the three emotional blocks was repeated five times for a total task duration of approximately 10 min.

Image preprocessing

See Supplementary Material for image acquisition methods. Before preprocessing, image quality was inspected with MRIQC.[21] This resulted in exclusion of scans with multiple framewise displacements >3 mm in any axis and other outlying image quality metrics, including number of outlier volumes, Analysis of Functional NeuroImages (AFNI) quality index, or temporal signal/noise ratio ($n = 2$ for emotional reappraisal task, and $n = 3$ for faces task). Preprocessing was performed using fMRIPrep version 20.2.4 (https://www.biorxiv.org/content/10.1101/2025.05.14.654069v1),[22] which is based on Nipype 1.6.1 (https://zenodo.org/records/15054184).[23] In brief, preprocessing of the functional images included slice timing correction, susceptibility distortion correction, motion correction, co-registration to the anatomical image and subsequent normalisation to standard space. We also applied automatic removal of motion artifacts by using independent component analysis (ICA-AROMA)[24] on the standard space blood oxygen level dependent time series after spatial smoothing with a 6 mm full-width half maximum Gaussian kernel. See the Supplementary Material for full detail regarding preprocessing steps.

Statistical analyses

Behavioural analyses

We conducted behavioural analyses through linear mixed-effects modelling (estimated using restricted maximum likelihood), using

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

the afex package in R (R Core Team, Vienna, Austria; https://CRAN.R-project.org/package=afex).[25] For all models, degrees of freedom were estimated via the Kenwood–Roger method. Where effects were significant, we conducted *post hoc* pairwise comparisons (*t*-tests) to assess for significant differences between factors. All mixed-effect models used sum-to-zero contrast coding. For model specification details, see the Supplementary Material.

Functional MRI analyses

Functional MRI (fMRI) analyses were conducted with FEAT version 6.00, part of FSL (FMRIB's Software Library, Oxford, UK; https://fsl.fmrib.ox.ac.uk/fsl/docs/index.html).[26] At the individual level, we applied a general linear model to the fMRI time series for each task condition convolved with a single-gamma canonical haemodynamic response function.[27] Separate general linear models were set up for each task. We modelled both tasks as block designs. For the emotion regulation task, the model consisted of five task regressors: neutral, attend negative, attend positive, downregulate negative and upregulate positive. The model for the faces task consisted of three task regressors: neutral faces, fearful faces and happy faces. Contrasts between these regressors were computed at the individual level and passed to the higher level. We additionally applied a high-pass filter cut-off of 105 s for the emotional reappraisal task and 90 s for the faces task. Higher-level analysis was carried out through mixed-effects modelling (FLAME stage 1 with outlier deweighting).[28] *Z*-statistic images were thresholded using clusters determined by $Z > 2.3$ and a corrected cluster significance threshold of $p = 0.05$.[29] In significant clusters, parameter estimates versus baseline were extracted to explore drivers of drug × task interactions.

Results

Of the 40 participants recruited, one withdrew after randomisation but before testing. A second participant was excluded before analysis, because of a significant ghosting artifact on all MRI scans resulting from equipment failure ($n = 1$), leaving data from 38 participants.

Emotional reappraisal task

Behavioural results

Behavioural results for the emotional reappraisal task are summarised in the Supplementary Material. For neutral images, we did not find a significant effect of drug condition or of order on mixed-effects modelling (F's < 0.96, p's > 0.33), showing that affective ratings were similar across drug conditions ($\text{mean}_{\text{placebo}} = 4.85$ *v.* $\text{mean}_{\text{naltrexone}} = 6.06$).

For negative images, mixed effects modelling showed significant effects of task instruction ($F_{37} = 17.63$, $p < 0.001$) and of drug condition ($F_{37} = 4.33$, $p = 0.044$) on subjective ratings of distress, with no significant interaction. As expected, estimated marginal means indicated that attend trials were associated with significantly increased distress compared with regulate trials ($\text{mean}_{\text{attend}} = -12.66$, 95% CI −17.00 to −8.27] *v.* $\text{mean}_{\text{regulate}} = -7.67$, 95% CI −12.00 to −3.37, $t_{(37)} = 4.20$, $p < 0.001$, $d = 1.31$). In addition, naltrexone was associated with reduced distress compared with placebo across both attend and regulate blocks ($\text{mean}_{\text{placebo}} = -11.24$, 95% CI −15.70 to −6.76 *v.* $\text{mean}_{\text{naltrexone}} = -9.08$, 95% CI −13.20 to −4.98, $t_{(37)} = 2.08$, $p = 0.044$, $d = 0.57$). There was no significant effect of order.

For analysis of positive image blocks, we excluded an additional participant who misunderstood the task instructions and believed they were required to downregulate their emotional response during 'regulate' trials. In the remaining 37 participants, mixed-effects modelling showed a significant effect of task instruction on positive affect ($F = 41.47$, $p < 0.001$). Estimated marginal means indicated that, as expected, positive affect was significantly higher in regulate trials compared with attend trials ($\text{mean}_{\text{attend}} = 15.20$, 95% CI 12.40−18.10 *v.* $\text{mean}_{\text{regulate}} = 21.30$, 95% CI 18.40−24.20, $t_{(36)} = 6.44$, $p < 0.001$, $d = 1.13$). There were no significant effects of drug condition or order, or significant drug condition × task instruction interaction.

fMRI results

Two participants were excluded from fMRI analyses in this task because of excessive head motion and poor image quality. Results from contrasts within each drug condition are summarised in Supplementary Table 1. Across both conditions, there was relatively increased activity in bilateral lateral occipital cortex extending to fusiform gyrus when attending to emotional compared with neutral images. Confirming earlier work, we additionally saw relatively increased activity in left dorso- and ventrolateral prefrontal cortex, left temporal cortex, and right cerebellum in regulate > attend contrasts, across both conditions.

When comparing conditions, the placebo > naltrexone comparison for the attend negative > neutral contrast revealed that there was increased activity in the bilateral medial thalamus extending to left caudate, and in left ventromedial prefrontal cortex (Table 1). This effect appeared driven by relatively lower thalamus/caudate activity to neutral images under placebo, and to negative images instead under naltrexone. Similarly, ventromedial prefrontal cortex activity was significantly reduced when viewing neutral images under placebo, and at a trend level when viewing negative images under naltrexone (Fig. 1(a) and Supplementary Table 2). There were no significantly different clusters of activity in the naltrexone > placebo comparison for this contrast. In the regulate > attend negative contrast, we found a single significant cluster in the naltrexone > placebo comparison. Activity was increased in left lateral occipital cortex under naltrexone during regulate blocks (Fig. 1(b) and Supplementary Table S2). There were no differences between drug conditions in the regulate > attend positive contrast.

Face-viewing task

Behavioural results

Results for this task are summarised in the Supplementary Material. With reaction times for correct responses as the outcome variable, mixed-effects modelling revealed a non-significant trend toward an effect of drug condition ($F_{37} = 3.62$, estimate $= 0.011$, $p = 0.065$), where reaction times were slower under naltrexone. Order and valence effects were not significant.

Regarding accuracy, we found a significant effect of valence on percentage of correct responses ($F_{36} = 3.53$, $p = 0.040$), but there was no effect of order or drug condition, and no significant interaction with drug condition. *Post hoc* contrasts of estimated marginal means revealed a significantly increased accuracy for both fearful and happy faces compared with neutral faces, across conditions (ts > 2.00, ps < 0.05). On inspection of model assumptions, we noted that residuals were non-normally distributed. We therefore conducted sensitivity analyses with square root transformation of error rate and with arcsine-square-root transformation of accuracy as outcome variables. The effect of valence remained significant, driven by increased accuracy for fearful compared with neutral faces across conditions (ts > 3.00, ps < 0.01), with no significant effects of order or condition.

fMRI results

Three participants were excluded because of excessive motion and poor image quality, leaving 35 participants for analysis.

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

Table 1 Clusters of significantly different activation between the naltrexone and placebo conditions in the emotional reappraisal task

Contrast	Anatomical region	Peak voxel (MNI) *x*	*y*	*z*	*k*	Peak *Z*	*p*-value	Brodmann area
Attend negative > neutral (placebo > naltrexone)								
	Left ventromedial prefrontal cortex	−10	48	−10	376	3.83	0.00278	32/24/10
	Bilateral thalamus extending to left caudate nucleus	−6	2	4	238	4.02	0.0491	Not applicable
Regulate > attend negative (naltrexone > placebo)								
	Left lateral occipital cortex	−32	−84	10	260	3.87	0.0332	19

MNI, Montreal Neurological Institute; *k*, number of voxels.

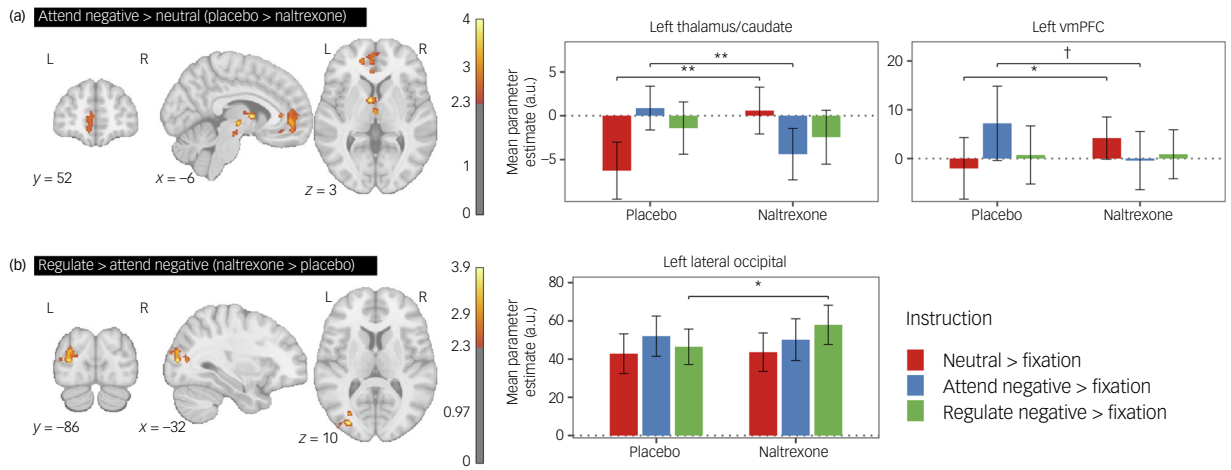

Fig. 1 Significant differences in relative activity between drug conditions in the emotional reappraisal task. Bars represent mean difference in parameter estimate from baseline (fixation cross). Error bars represent 95% confidence intervals. Paired *t*-tests were conducted to compare drug conditions by task instruction. (a) In the attend negative > neutral contrast, there was greater relative activity under placebo in the left thalamus extending into the caudate nucleus and in the left ventromedial prefrontal cortex. (b) In the regulate > attend negative contrast, there was greater relative activity under naltrexone in the left lateral occipital cortex. $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.07$ (trend).

Whole-brain analyses revealed that face-viewing relative to fixation activated regions, including the fusiform gyrus, bilateral amygdala, bilateral insula and orbitofrontal cortex, in both drug conditions. There were additionally relative differences in activation when viewing different facial expressions (see Supplementary Table 3). In the placebo condition, the fearful > happy contrast revealed relatively increased activity in ventromedial prefrontal cortex, whereas the happy > fearful contrast showed relatively increased activity in lateral prefrontal regions. In the naltrexone condition, there was relatively increased activity in temporo-occipito-parietal regions, including fusiform gyrus, when viewing fearful faces compared with happy or neutral faces. We did not find relative differences in activity in the opposite direction.

In the fearful > happy contrast, the naltrexone > placebo comparison revealed increased activity in bilateral supramarginal gyrus, right middle temporal gyrus and right superior frontal gyrus (Table 2). This appeared driven by greater activation when viewing happy and neutral faces under placebo in these regions, although the difference was only significant in the right superior frontal gyrus. The placebo > naltrexone comparison revealed increased activity in the ventromedial prefrontal cortex, which appeared driven at a trend level by reduced deactivation when viewing happy faces with naltrexone (see Fig. 2(a) and Supplementary Table 4).

In the fearful > neutral contrast, the naltrexone > placebo comparison revealed greater activation in the right frontal pole, right precuneus and left middle frontal gyrus. Examination of parameter estimates showed the precuneus effect was driven by a smaller reduction in precuneus activity for fearful faces under naltrexone, whereas the other effects were driven by increases in activity under naltrexone (see Fig. 2(b) and Supplementary Table 4). There were no significant clusters in the placebo > naltrexone comparison.

Discussion

We examined the effects of endogenous opioid antagonism on explicit cognitive emotional reappraisal and implicit emotional processing. Contrary to our hypotheses, we found that naltrexone was associated with less subjective distress in the emotional reappraisal task and did not affect participants' ability to apply cognitive reappraisal strategies. Patterns of fMRI activation in the regulate > attend contrasts were similar between drug conditions, apart from one cluster of relatively increased activation in the left lateral occipital cortex in the naltrexone condition. By contrast, there was relatively increased activation in the thalamus and caudate and in the ventromedial prefrontal cortex in the attend negative > neutral contrast with placebo compared with naltrexone. Similarly, placebo was associated with increased ventromedial prefrontal cortex activity when viewing fearful compared with happy faces, driven by a greater decrease in activity here when

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

Table 2 Clusters of significantly different activation between the naltrexone and placebo conditions in the faces task

Contrast	Anatomical region	Peak voxel (MNI) *x*	*y*	*z*	*k*	*Z*	*p*-value	Brodmann area
Fearful > happy (naltrexone > placebo)								
	Right middle temporal gyrus/ posterior supramarginal gyrus	66	−44	2	428	3.95	0.00099	22/40
	Right superior frontal gyrus	30	0	68	290	3.61	0.0155	6
	Left supramarginal gyrus	−62	−40	30	285	3.73	0.0172	40
Fearful > happy (placebo > naltrexone)								
	Bilateral ventromedial prefrontal cortex extending to pregenual anterior cingulate gyrus	−10	54	−2	349	3.93	0.00459	10/32
Fearful > neutral (naltrexone > placebo)								
	Right frontal pole	34	54	16	237	4.06	0.0237	10
	Right precuneus	6	−70	60	224	3.78	0.0330	7
	Left middle frontal gyrus	−40	36	34	219	3.90	0.0374	9

MNI, Montreal Neurological Institute; *k*, number of voxels.

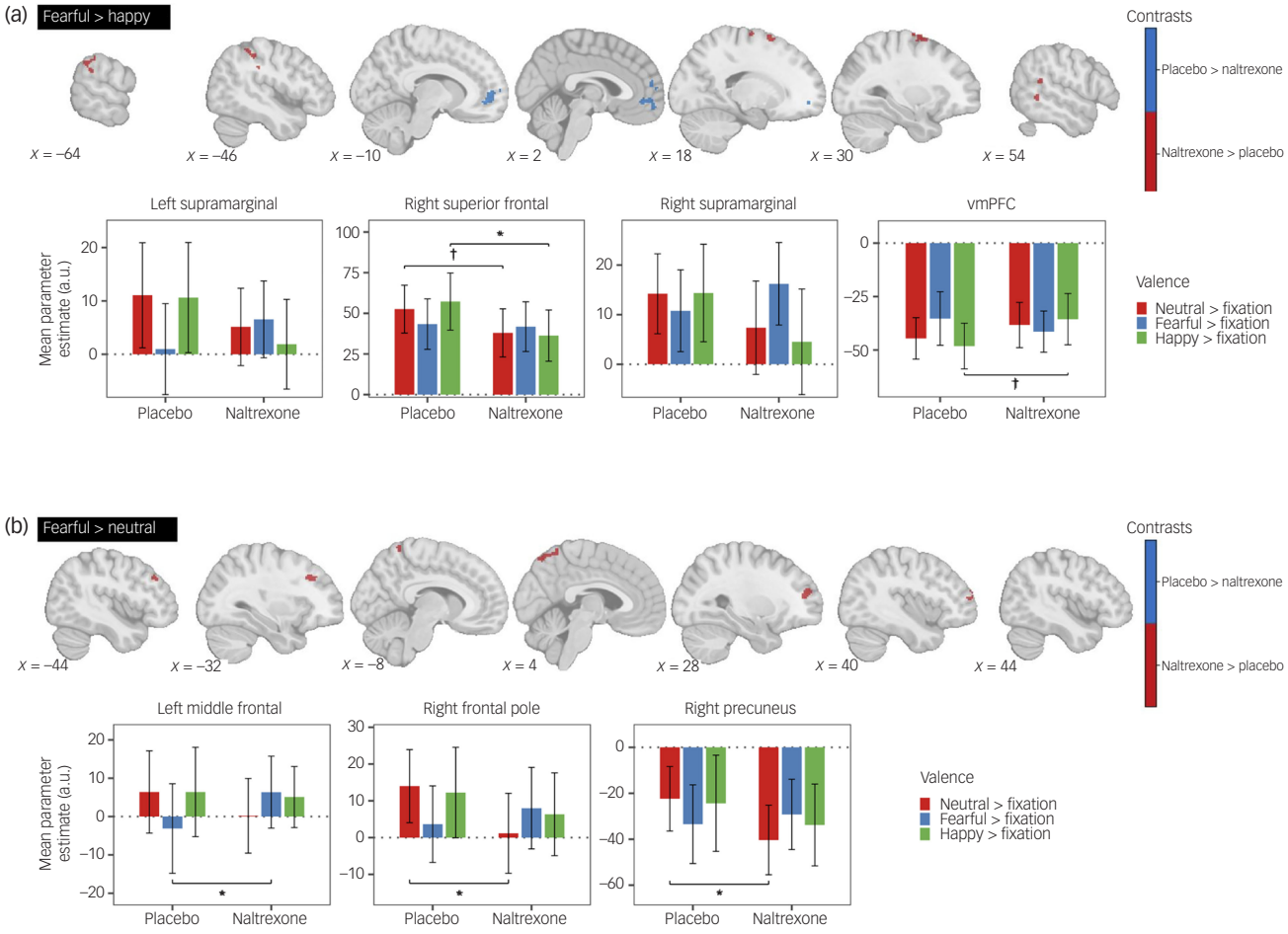

Fig. 2 Significant differences in relative activity between drug conditions in the face-viewing task. Bars represent mean difference in parameter estimate from baseline (fixation cross). Error bars represent 95% confidence intervals. Paired *t*-tests were conducted to compare drug conditions by valence. (a) In the fearful > happy contrast, there was greater relative activity under placebo in the ventromedial prefrontal cortex (vmPFC). By contrast, there was greater activity under naltrexone in bilateral supramarginal and right superior frontal gyrus. (b) In the fearful > neutral contrast, there was greater relative activity under naltrexone in the left middle frontal gyrus, right frontal pole and right precuneus. $^{*}p < 0.05$, $^{\dagger}p < 0.07$ (trend).

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

viewing happy faces. Meanwhile, activity in the naltrexone condition involved mostly temporal and parietal regions when viewing emotional faces.

Emotion reappraisal task findings

Attending to emotional images activated a group of regions encompassing bilateral occipito-temporal cortex including fusiform gyrus and precuneus/posterior cingulate cortex. This group of brain regions is broadly in line with those reported in the literature when healthy volunteers view emotional compared with neutral images,[30,31] suggesting our participants were attending to the emotional content as instructed. Interestingly, placebo was associated with relatively increased activity in the ventromedial prefrontal cortex, and thalamus extending to caudate nucleus when viewing negative images, whereas activity in these regions was instead increased for neutral images with naltrexone. The thalamus, striatum and amygdala seem involved in detecting salient emotional stimuli and deploying vigilant attention as required in response.[32] These regions interact with ventromedial prefrontal areas, likely with the thalamus as an integrative centre, to generate or regulate negative emotions depending on optimal behaviour.[33,34] Given that naltrexone was additionally associated with reduced subjective distress, our findings indicate there is perhaps a 'disruption' of the normal interaction between these regions and resulting emotional experience following administration of naltrexone.

In the 'regulate' blocks, participants were able to upregulate or downregulate their emotional response to positive and negative images, and neural correlates were in line with previous literature.[19,35] The only drug effect we found was that naltrexone was associated with increased left lateral occipital cortex activity in the regulate > attend negative contrast. Inspection of parameter estimates showed this was driven by a relative increase in activity in 'regulate' blocks. Accumulating evidence suggests that down-regulation of negative emotion via reappraisal mechanisms is achieved through lateral prefrontal cortex modulation of ventro-medial prefrontal cortex-to-amygdala modulation.[36] However, as described above, it appears naltrexone altered ventromedial prefrontal cortex responses to negative images. A hypothesis that follows is that participants perhaps relied more on perceptual operations within lateral occipital cortex to successfully carry out the reappraisal instructions.

Faces task findings

We found that regions activated by viewing emotional faces, compared with fixation cross, were broadly consistent with regions reported in previous implicit emotional face-processing tasks.[37] When contrasting activity between different facial expressions, we found some noteworthy results. When viewing happy faces, placebo was associated with greater deactivation of ventromedial prefrontal cortex compared with naltrexone. The medial prefrontal cortex is a node within the default mode network, which participants would have been required to disengage to carry out the task.[38] Reduced deactivation under naltrexone, therefore, might reflect reduced social or hedonic salience of happy faces, resulting in less task-related suppression of this region. Consistent with this, naltrexone was associated with reduced activity in right superior frontal gyrus when viewing happy faces, and at a trend level when viewing neutral faces. This right superior frontal region was within the premotor area, with activity here perhaps reflecting action planning and attentional orienting,[39] i.e. reduced task-positive processing with naltrexone.

In the fearful > neutral faces contrast, naltrexone was associated with increased activity in right lateral frontal pole, precuneus and left middle frontal gyrus. This result was driven by reduced activity in right frontal pole and greater deactivation in precuneus when viewing neutral faces, and increased activity in left middle frontal gyrus when viewing fearful faces, compared with placebo. Right lateral frontal pole is a task-positive region, thought to be involved in directed exploration.[40] The precuneus region found in this analysis corresponded with the precuneus visual area, which appears involved in identifying emotional faces over neutral objects.[41] Reduced activity here might suggest reduced attentional engagement with neutral face stimuli. Left middle frontal gyrus is usually deactivated during emotional face-viewing, with greater deactivation for negative emotional faces.[38] We saw this pattern in the placebo condition, but the opposite in the naltrexone condition, with greater activation for fearful faces compared with fixation cross. These findings broadly suggest naltrexone was associated with reductions in social salience for neutral faces and greater reliance on cognitive processing to achieve task goals for fearful faces. These interpretations are consistent with the trend towards slower reaction times in the task in the naltrexone condition.

Does naltrexone flatten affective intensity?

In sum, findings from both tasks suggest naltrexone altered automatic emotion generation and/or regulation, whereas explicit regulation remained intact. There are some, but few, similar findings reported in the literature. A 2020 study examined the effect of naltrexone on neural responses to contextual framing and subsequent identification of emotional facial expressions in 20 unmedicated patients with depression. The authors found that naltrexone diminished ventromedial prefrontal cortical moderation context effects on emotional identification, but enhanced lateral prefrontal cortical moderation.[42] In another study, healthy volunteers chose whether faces morphed on a continuum represented 'disgust' or 'pain' expressions. On average, participants chose 'pain' less frequently in the naltrexone session, and activity in right occipito-temporal and fusiform cortex was increased.[43] Furthermore, previous studies have shown naltrexone attenuates substance cue-reactivity in striatum and ventromedial prefrontal cortex in patients with alcohol use disorder[44] and in patients with opioid use disorder,[45] and attenuates activation to hedonic cues in the caudate in healthy volunteers.[46] Therefore, converging evidence appears to suggest naltrexone attenuates medial prefrontal-limbic activity in response to emotional salience.

We speculate that naltrexone administration results in a general blunting of affective intensity, mediated via blockade of mu-opioid receptors. Naltrexone is primarily a mu-opioid receptor antagonist. Converging evidence suggests that the mu-opioid system 'fine-tunes' behaviour in relation to both stress/fear and reward (for a review[47]). For example, the mu-opioid receptor antagonist naloxone modulates hedonic responses for both rewards and losses, and this is associated with reduced activity in cortico-limbic structures including ventral striatum, amygdala and medial prefrontal cortex activity.[48] In a dot-probe paradigm, naltrexone 50 mg reduces subjective arousal in response to emotional facial expressions, interferes with identification of fearful and sad faces, but also increases attention to emotional faces.[49] However, naltrexone is additionally a kappa-opioid antagonist. The kappa-opioid system is known to mediate the aversive affective component of pain and evidence suggests that antagonism of this system might be anxiolytic.[50] Indeed, buprenorphine (a partial mu-agonist and kappa-antagonist) has been shown to reduce recognition of fearful facial expressions[51] and reduce subjective feelings of threat in anticipation of, and salivary cortisol response to, the Trier Social Stress Test.[52] Thus, depending on task demands and dosage used, naltrexone may exhibit both anhedonic and anxiolytic

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

properties, as we have found in this study and in our recently published report on the effect of naltrexone on reward receipt.[14]

Given the above, the finding that naltrexone enhances fear responses conditioned through observational learning[10] appears counterintuitive. However, converging animal and human evidence suggests that reduced ventromedial prefrontal cortex activity is associated with overgeneralisation of conditioned fear, possibly via thalamo-cortical circuits.[34] The result is expression of fear responses to stimuli perceptually similar to the conditioned stimulus. Given that we found a reduction in ventromedial prefrontal cortex activity with naltrexone in response to negative stimuli, it is possible that the previously reported enhanced conditioned fear responses results from an overgeneralisation effect.

To summarise, we propose a model whereby endogenous mu-opioid tone modulates arousal signals in ventromedial prefrontal cortex-striatal-limbic circuits (akin to 'neural gain' for arousal). This is consistent with positron emission tomography–fMRI imaging showing that baseline mu-opioid receptor availability is associated with arousal but not valence when viewing movie scenes.[53] By blocking mu-opioid receptors, we suggest that naltrexone blunts normal modulation of arousal in ventromedial prefrontal cortex-striatal-limbic circuits in response to affective-salience cues. This processing is automatic, whereas explicit cognitive reappraisal is preserved, suggesting that lateral prefrontal-limbic circuits are not opioid-dependent. Kappa-opioid antagonism by naltrexone may additionally blunt negative-valence processing specifically,[49] but given naltrexone's non-selective profile, this requires selective probing to test.

Clinical implications and future research

Prefrontal-limbic dysregulation and altered implicit face-processing recur across anxiety disorders. Our pattern of results suggests a clinical hypothesis: mu-opioid blockade might reduce arousal to anxiogenic cues, but with the potentially negative consequence of dampened hedonic salience, in patients with anxiety disorders. The effects of baseline opioidergic tone might be key in determining this balance.[11,13] Future work should explore resting opioidergic tone across anxiety disorders, and how this interacts with behaviour and brain responses during emotional processing. In addition, the role of kappa-opioid mechanisms in shaping negative valence processing, both in healthy volunteers and in patients, needs further exploration. This could be achieved either with specific probes or potentially by exploiting dose-dependent profiles of non-selective antagonists such as naltrexone. Answering these questions would help elucidate whether endogenous opioid modulation is a tractable therapeutic target in anxiety disorders.

Limitations

Our findings should be considered in light of possible limitations. First, this study was powered to detect changes in neural activation owing to naltrexone. Statistical power was therefore limited to detect subtle drug effects or interactions on behaviour. Second, although we used a within-individual crossover design to reduce inter-individual variability, residual carry-over effects cannot be completely excluded despite the washout period. Third, we examined a single dose of naltrexone (50 mg) and a single time point. We could not assess dose–response relationships and temporal dynamics of naltrexone. Fourth, although fMRI provides valuable insights into functional activation, it cannot directly measure neurotransmitter activity, and our interpretations regarding mu- and kappa-opioid receptor mechanisms remain speculative. Finally, we did not collect physiological measures such as cortisol or blood pressure, limiting our ability to link neural changes to broader stress-response systems.

In conclusion, we found that endogenous opioid antagonism with naltrexone alters neural responses to emotional stimuli, reducing subjective distress without impairing explicit emotion regulation. Across two tasks, naltrexone was associated with altered patterns of automatic emotional processing. These findings highlight the complexity of the endogenous opioid system with regard to 'fine-tuning' automatic appraisal and regulation of emotional salience, and that blocking this system may blunt affective intensity rather than simply increasing distress. Future research should explore dose–response relationships with naltrexone and further delineate these mechanisms in clinical populations.

Nathan T. M. Huneke, School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; and Hampshire and Isle of Wight Healthcare NHS Foundation Trust, Calmore, UK; **Henk van Steenbergen**, Cognitive Psychology Unit, Institute of Psychology, Leiden University, The Netherlands; and Leiden Institute for Brain and Cognition, Leiden, The Netherlands; **Harry A. Fagan**, School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; and Hampshire and Isle of Wight Healthcare NHS Foundation Trust, Calmore, UK; **Laura Molteni**, School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; **Algirdas Midveris**, School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; **Naomi Phillips**, Hampshire and Isle of Wight Healthcare NHS Foundation Trust, Calmore, UK; **Angela Darekar**, Department of Medical Physics, University Hospital Southampton NHS Foundation Trust, Southampton, UK; **Nic J. A. van der Wee**, Department of Psychiatry, Leiden University Medical Center (LUMC), Leiden, The Netherlands; **Matthew Garner**, School of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK; **David S. Baldwin**, School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; and Hampshire and Isle of Wight Healthcare NHS Foundation Trust, Calmore, UK

Correspondence: Nathan T. M. Huneke. Email: n.huneke@soton.ac.uk

First received 22 Dec 2025, final revision 16 Feb 2026, accepted 18 Feb 2026, first published online 16 Apr 2026

Supplementary material

The supplementary material is available online at https://doi.org/10.1192/bjp.2026.10605

Data availability

Data that support the findings of this study are available at https://osf.io/x8rh3.

Acknowledgements

We would like to thank Chris Everitt, MRI Research Radiographer at University Hospital Southampton, Southampton, UK, for all her help with data collection. We would also like to thank Prof Marie-José van Tol of University Medical Centre Groningen, The Netherlands, for supplying the Emotional Reappraisal Task. H.A.F. was supported to contribute to this work through a National Institute for Health and Care Research (NIHR, UK) Academic Foundation Post. N.T.M.H., D.S.B. and N.J.A.v.d.W. are all members of the European College of Neuropsychopharmacology Anxiety Disorders Research Network (ECNP ADRN).

Author contributions

N.T.M.H. contributed to study conceptualisation, formal analysis, investigation, data curation, visualisation, funding acquisition and writing, reviewing and editing the manuscript. H.v.S. contributed to conceptualisation, formal analysis, supervision and reviewing and editing the manuscript. H.A.F., L.M., A.M. and N.P. contributed to study investigation, and reviewing and editing the manuscript. A.D. contributed to study conceptualisation, methodology, resources, and reviewing and editing the manuscript. N.J.A.v.d.W., M.G. and D.S.B. contributed to study conceptualisation, supervision, and reviewing and editing the manuscript.

Funding

This work was supported by a Medical Research Council, UK, grant awarded to N.T.M.H. (grant number MR/T000902/1).

Declaration of interest

N.T.M.H. is Deputy Director of Education for the British Association for Psychopharmacology, for which he receives an honorarium, and is a NIHR Clinical Lecturer. He has received a speaker honorarium from Focus Gulf Conferences. The views expressed in this publication are those of the authors and not necessarily those of the National Institute for Health and Care Research, NHS or the UK Department of Health and Social Care. D.S.B. declares current

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

research funding from the NIHR, for research in other areas, and support from the Office of Life Sciences, UK. A.D. is NIHR South Central RRDN Imaging Specialty Lead, for which her department receives funding. All other authors declare no conflicts of interest. N.T.M.H. and D.S.B. are guest editors and did not take part in the review or decision-making process of this paper.

References

1 Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialog Clin Neurosci* 2015; **17**: 327–35.

2 McDaid D, Park A-L, Davidson G, John A, Knifton L, McDaid S, et al. *The Economic Case for Investing in the Prevention of Mental Health Conditions in the UK*. Mental Health Foundation, 2022. (https://www.mentalhealth.org.uk/explore-mental-health/publications/economic-caseinvesting-prevention-mental-health-conditions-UK).

3 GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry* 2022; **9**: 137–50.

4 Springer KS, Levy HC, Tolin DF. Remission in CBT for adult anxiety disorders: a meta-analysis. *Clin Psychol Rev* 2018; **61**: 1–8.

5 Bereza BG, Machado M, Ravindran AV, Einarson TR. Evidence-based review of clinical outcomes of guideline-recommended pharmacotherapies for generalized anxiety disorder. *Can J Psychiatry* 2012; **57**: 470–8.

6 Karp JF, Mathew S, Todtenkopf MS, Ehrich EW, Zubieta J-K. Endogenous opioid system dysregulation in depression: implications for new therapeutic approaches. *Mol Psychiatry* 2019; **24**: 576–87.

7 Kang W, Wilson SP, Wilson MA. Overexpression of proenkephalin in the amygdala potentiates the anxiolytic effects of benzodiazepines. *Neuropsychopharmacology* 2000; **22**: 77–88.

8 Burghardt PR, Wilson MA. Microinjection of naltrexone into the central, but not the basolateral, amygdala blocks the anxiolytic effects of diazepam in the plus maze. *Neuropsychopharmacology* 2006; **31**: 1227–40.

9 Duka T, Millan MJ, Ulsamer B, Doenicke A. Naloxone attenuates the anxiolytic action of diazepam in man. *Life Sci* 1982; **31**: 1833–6.

10 Haaker J, Yi J, Petrovic P, Olsson A. Endogenous opioids regulate social threat learning in humans. *Nat Commun* 2017; **8**: 15495.

11 Nummenmaa L, Karjalainen T, Isojärvi J, Kantonen T, Tuisku J, Kaasinen V, et al. Lowered endogenous mu-opioid receptor availability in subclinical depression and anxiety. *Neuropsychopharmacology* 2020; **45**: 1953–9.

12 Picó-Pérez M, Radua J, Steward T, Menchón JM, Soriano-Mas C. Emotion regulation in mood and anxiety disorders: a meta-analysis of fMRI cognitive reappraisal studies. *Prog Neuro-Psychopharmacol Biol Psychiatry* 2017; **79**: 96–104.

13 Seppälä K, Putkinen V, Harju H, Rebelos E, Hirvonen J, Helin S, et al. Endogenous opioid system modulates proximal and distal threat signals in the human brain. *Mol Psychiatry* [Epub ahead of print] 24 Dec 2025. Available from: https://doi.org/10.1038/s41380-025-03385-3.

14 Gable P, Fagan H, Molteni L, Midveris A, Huneke N. Seeing the bigger picture: endogenous opioids mediate attentional broadening after reward receipt. *Psychol Med* 2025; **55**: e283.

15 Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a Structured Diagnostic Psychiatric Interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998; **59**: 22–33.

16 Veale JF. Edinburgh handedness inventory – short form: a revised version based on confirmatory factor analysis. *Laterality* 2014; **19**: 164–77.

17 Gonzalez JP, Brogden RN. Naltrexone. *Drugs* 1988; **35**: 192–213.

18 Eikemo M, Haaker J, Frost JJ, Leknes S. Opioid antagonism in humans: a primer on optimal dose and timing for central mu-opioid receptor blockade. *Neuropsychopharmacology* 2023; **48**: 299–307.

19 van Kleef RS, Müller A, van Velzen LS, Marie Bas-Hoogendam J, van der Wee NJA, L Schmaal, et al. Functional MRI correlates of emotion regulation in major depressive disorder related to depressive disease load measured over nine years. *NeuroImage Clin* 2023; **40**: 103535.

20 Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, et al. The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res* 2009; **168**: 242–9.

21 Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ. MRIQC: advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE* 2017; **12**: e0184661.

22 Markiewicz CJ, Esteban O, Goncalves M, Poldrack RA, Gorgolewski KJ. fMRIPrep: a robust preprocessing pipeline for functional MRI. *BioRxiv* [Preprint] 2025. Available from: https://www.biorxiv.org/content/10.1101/2025.05.14.654069v1.

23 Esteban O, Markiewicz CJ, Burns C, Goncalves M, Jarecka D, Ziegler E, et al. *nipy/nipype: 1.9.1*. Zenodo, 2025 (https://zenodo.org/records/15054184).

24 Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF. ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage* 2015; **112**: 267–77.

25 Singmann H, Bolker B, Westfall J, Aust F, Ben-Shachar MS, Højsgaard S, et al. *afex: Analysis of Factorial Experiments*. The Comprehensive R Archive Network, 2025 (https://cran.r-project.org/web/packages/afex/index.html).

26 Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage* 2012; **62**: 782–90.

27 Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage* 2001; **14**: 1370–86.

28 Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, Smith SM. Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage* 2004; **21**: 1732–47.

29 Worsley KJ. 14 Statistical analysis of activation images. In *Functional Magnetic Resonance Imaging: An Introduction to Methods* (eds P Jezzard, PM Matthews, SM Smith): 251–70. Oxford University Press, 2001.

30 Mansueto SP, Romeo Z, Angrilli A, Spironelli C. Emotional pictures in the brain and their interaction with the task: a fine-grained fMRI coordinate-based meta-analysis study. *NeuroImage* 2025; **305**: 120986.

31 Gerdes ABM, Wieser MJ, Mühlberger A, Weyers P, Alpers GW, MM Plichta, et al. Brain activations to emotional pictures are differentially associated with valence and arousal ratings. *Front Hum Neurosci* 2010; **4**: 175.

32 Barson JR, Mack NR, Gao W-J. The paraventricular nucleus of the thalamus is an important node in the emotional processing network. *Front Behav Neurosci* 2020; **14**: 598469.

33 Mair RG, Francoeur MJ, Krell EM, Gibson BM. Where actions meet outcomes: medial prefrontal cortex, central thalamus, and the basal ganglia. *Front Behav Neurosci* 2022; **16**: 928610.

34 Hiser J, Koenigs M. The multifaceted role of ventromedial prefrontal cortex in emotion, decision-making, social cognition, and psychopathology. *Biol Psychiatry* 2018; **83**: 638–47.

35 Buhle JT, Silvers JA, Wager TD, Lopez R, Onyemekwu C, Kober H, et al. Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cereb Cortex* 2014; **24**: 2981–90.

36 Steward T, Davey CG, Jamieson AJ, Stephanou K, Soriano-Mas C, Felmingham KL, et al. Dynamic neural interactions supporting the cognitive reappraisal of emotion. *Cereb Cortex* 2021; **31**: 961–73.

37 Fusar-Poli P, Placentino A, Carletti F, Landi P, Allen P, Surguladze S, et al. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J Psychiatry Neurosci* 2009; **34**: 418–32.

38 Sreenivas S, Boehm SG, Linden DEJ. Emotional faces and the default mode network. *Neurosci Lett* 2012; **506**: 229–34.

39 Nachev P, Kennard C, Husain M. Functional role of the supplementary and pre-supplementary motor areas. *Nat Rev Neurosci* 2008; **9**: 856–69.

40 Ramnani N, Owen AM. Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat Rev Neurosci* 2004; **5**: 184–94.

41 Dadario NB, Sughrue ME. The functional role of the precuneus. *Brain* 2023; **146**: 3598–607.

42 Chen J, Mizuno A, Lyew T, Karim HT, Karp JF, Dombrovski AY, et al. Naltrexone modulates contextual processing in depression. *Neuropsychopharmacology* 2020; **45**: 2070–8.

43 Zhao Y, Rütgen M, Zhang L, Lamm C. Pharmacological fMRI provides evidence for opioidergic modulation of discrimination of facial pain expressions. *Psychophysiology* 2021; **58**: e13717.

44 Bach P, Weil G, Pompili E, Hoffmann S, Hermann D, Vollstädt-Klein S, et al. FMRI-based prediction of naltrexone response in alcohol use disorder: a replication study. *Eur Arch Psychiatry Clin Neurosci* 2021; **271**: 915–27.

45 Shi Z, Wang A-L, Jagannathan K, Fairchild VP, O'Brien CP, Childress AR, et al. Effects of extended-release naltrexone on the brain response to drug-related stimuli in patients with opioid use disorder. *J Psychiatry Neurosci* 2018; **43**: 254–61.

46 Murray E, Brouwer S, McCutcheon R, Harmer CJ, Cowen PJ, McCabe C. Opposing neural effects of naltrexone on food reward and aversion: implications for the treatment of obesity. *Psychopharmacology* 2014; **231**: 4323–35.

47 Meier IM, Eikemo M, Leknes S. The role of Mu-opioids for reward and threat processing in humans: bridging the gap from preclinical to clinical opioid drug studies. *Curr Addict Rep* 2021; **8**: 306–18.

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press

48 Petrovic P, Pleger B, Seymour B, Klöppel S, De Martino B, Critchley H, et al. Blocking central opiate function modulates hedonic impact and anterior cingulate response to rewards and losses. *J Neurosci* 2008; **28**: 10509–16.

49 Wardle MC, Bershad AK, de Wit H. Naltrexone alters the processing of social and emotional stimuli in healthy adults. *Soc Neurosci* 2016; **11**: 579–91.

50 Emery MA, Akil H. Endogenous opioids at the intersection of opioid addiction, pain, and depression: the search for a precision medicine approach. *Annu Rev Neurosci* 2020; **43**: 355–74.

51 Ipser JC, Terburg D, Syal S, Phillips N, Solms M, Panksepp J, et al. Reduced fear-recognition sensitivity following acute buprenorphine administration in healthy volunteers. *Psychoneuroendocrinology* 2013; **38**: 166–70.

52 Bershad AK, Jaffe JH, Childs E, de Wit H. Opioid partial agonist buprenorphine dampens responses to psychosocial stress in humans. *Psychoneuroendocrinology* 2015; **52**: 281–8.

53 Karjalainen T, Seppälä K, Glerean E, Karlsson HK, Lahnakoski JM, Nuutila P, et al. Opioidergic regulation of emotional arousal: a combined PET–fMRI study. *Cereb Cortex* 2019; **29**: 4006–16.

https://doi.org/10.1192/bjp.2026.10605 Published online by Cambridge University Press