Endogenous opioid receptor system mediates costly altruism in the human brain

Jinglu Chen[1], Vesa Putkinen[1,2], Kerttu Seppälä[1], Jussi Hirvonen[1,3], Kalliopi Ioumpa[4], Valeria Gazzola[4,5], Christian Keysers[4,5] & Lauri Nummenmaa[1,6,7]

[1]Turku PET Centre, Turku University Hospital and University of Turku, Turku, Finland
[2]Turku Institute for Advanced Studies, University of Turku, Turku, Finland
[3]Department of Radiology, Turku University Hospital, Turku, Finland
[4]The Netherlands Institute for Neuroscience, KNAW research institute, Amsterdam, The Netherlands
[5]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
[6]Turku University Hospital, Turku, Finland
[7]Department of Psychology, University of Turku, Finland
Keywords: Pain, helping, opioids, positron emission tomography

Address for correspondence
Jinglu Chen
Turku PET Centre c/o Turku University Hospital
FI-20520 Turku, Finland
Email: jinglu.chen@utu.fi

Abstract

Functional neuroimaging studies suggest that a large-scale brain network transforms others' pain into its vicarious representation in the observer, potentially modulating helping behaviour. The neuromolecular basis of individual differences in vicarious pain and helping, however, have remain poorly understood. Here we investigated the role of the endogenous μ -opioid receptor (MOR) system – known for its role in analgesia and sociability – in altruistic costly helping. MOR density was measured with high-affinity agonist radioligand [^{11}C]carfentanil. In a separate fMRI experiment, participants could choose to donate money to reduce the pain of the confederate who was subjected to electric shocks of varying intensity. We found that subjects were in general willing to engage in costly helping, and haemodynamic activity in amydala, anterior insula, anterior cingulate cortex, striatum, primary motor cortex, primary somatosensory cortex, and thalamus increased when participants witnessed the pain of others. These haemodynamic responses were negatively associated with MORs availability in the striatolimbic and cortical emotion circuits. In turn, haemodynamic responses during helping were positively associated with MOR availability in the anterior cingulate cortex and hippocampus. Altogether these data suggest that the endogenous MOR system modulates the processing of altruistic behaviour and costly helping in the human brain.

Introduction

Prosocial behavior is prevalent in humans and animals. Social animals share emotions with each other, comfort peer's distress, and help others in times of need (De Waal & Preston, 2017; Keysers et al., 2022). Humans help each other in daily life even without genetic relatedness or obvious direct profit (Batson, 2010, 2011; Batson et al., 1997). Such altruistic behavior by definition benefits the receiver, but also has a range of positive outcomes for the helper, including improved academic performance and social preferences (Caprara et al., 2000), increased acceptance by peers at school (Deković & Gerris, 1994), and higher life satisfaction (Caprara & Steca, 2005). The evolutionary routes of altruism are a topic of enduring interest (Dovidio & Penner, 2003). For more information, see review Laguna et al., 2020. A prominent hypothesis for the proximate causes for altruistic helping holds that if we empathically activate our own pain when witnessing the pain of others, mechanisms that have evolved to motivate us to prevent damage to ourselves will also motivate us to prevent pain and damage to others. Accordingly, empathy can be hypothesized to drive prosociality (Batson et al., 2014; Decety et al., 2016; Keysers & Gazzola, 2014; Smith, 1759).

The neurocognitive link between empathy and prosociality has been investigated by quantifying responses in brain regions associated with empathy while witnessing the pain of others, and correlating these measures with individual differences in helping. Studies leveraging a range of different methods have consistently shown that seeing or hearing

others in distress compared to neutral states engages a core network including the anterior insula (aIns) and adjacent frontal operculum, the mid- to anterior- cingulate cortex, amygdala and, when the somatic source of the pain is salient, the primary and secondary somatosensory cortices (SI and SII) (Bufalari et al., 2007; De Waal & Preston, 2017; Keysers et al., 2010; Krishnan et al., 2016; Lamm et al., 2011; Paradiso et al., 2021). Based on the details of the task and the nature of how the pain is witnessed, this extends into a larger network also encompassing primary motor cortex (M1), extensive areas in frontal lobe and parietal lobe, dorsolateral prefrontal cortex (dlPFC), ventromedial frontal cortex (vmPFC), fusiform gyrus, temporal pole, precuneus, thalamus, caudate, and putamen(Jauniaux et al., 2019; Timmers et al., 2018). Particularly the ACC and aIns are activated both when subjects experience first-hand pain and when they see others experiencing pain, suggesting that these regions underlie the transformation of others' pain from visual and auditory inputs into sensorimotor formats (Bufalari et al., 2007; Jackson et al., 2005, 2006; Krishnan et al., 2016; Lamm et al., 2011; Saarela et al., 2007). The ACC has been shown to contain single neurons that respond to both witnessing and experiencing pain in rodents(Carrillo et al., 2019).

To investigate whether this recruitment of regions involved in the witnesses' own pain plays a role in motivating helping, some studies have gone beyond simply measuring brain activity while passively witnessing the pain of others and provided the participant with an opportunity to help. One study offered participants the opportunity to relieve other people's pain by taking some of that pain onto themselves and found that anterior insular activity predicted the willingness to help ingroup members and nucleus accumbens activity predicted reluctance to help outgroup members (Hein et al., 2010). Another found that reduced functional connectivity between the insula and ACC characterized participants that decide to help a person in need in a virtual reality scenario (Zanon et al., 2014). Activity in somatosensory and insular region while witnessing pain predicted later charitable donations, albeit in ways that depended on socioeconomic status (Ma et al., 2011). An approach that has been particularly successful in identifying brain regions with activity quantitatively associated with helping, involves offering participants a certain financial endowment, and enabling them to forfeit some of that endowment on a trial-by-trial basis to reduce the pain of a victim in an immersive experimental setting (FeldmanHall et al., 2012). These studies have found a variety of nodes also involved with self-pain to relate to individual or trial-by-trial differences in helping, including the dorsolateral prefrontal cortex, orbital frontal cortex (FeldmanHall et al., 2015), SI (Gallo et al., 2018), vmPFC (FeldmanHall et al., 2013; Fornari et al., 2023), insula, ACC, and amygdala (Caspar et al., 2022).

Although these financially costly helping paradigms thus start to provide some traction on which neural networks have BOLD activity that correlate with costly helping, the molecular basis of empathy and costly helping, and their individual differences, remain poorly understood. Multiple lines of evidence point towards the critical role of the endogenous opioid system and particularly the endogenous μ-opioid receptor (MOR)

system in the first-person experience of pain and in differences in empathy and prosociality. Among the three classes of opioid receptors (μ, δ, and ϰ), the μ receptors mediate the effects of endogenous β-endorphins, endomorphins, enkephalins, and various exogenous opioid agonists(Henriksen & Willoch, 2008). The predominant action of μ-opioids in the central nervous system is inhibitory, but they can also exert excitatory effects, and MORs are expressed widely throughout the human emotion circuits(Kantonen, Karjalainen, Isoj, et al., 2020; Nummenmaa & Tuominen, 2018). Opioids are well known for their role in antinociception, and genetic deletions of the MOR in mice abolishes the analgesic effect of opioid agonists (Darcq & Kieffer, 2018), demonstrating that this receptor is the sole pathway for the analgesic effects of these drugs. The effect of opioids however goes beyond nociception: opioid agonists decrease and antagonists increase social motivation in macaques (Fabre-Nys et al., 1982; Keverne et al., 1989), and in humans, the opioid system modulates both positive and negative emotions (Nummenmaa et al., 2020; Nummenmaa & Tuominen, 2018)

With regard to empathy, recent experiments have shown that activating the MOR system using placebo analgesia reduces first-hand pain ratings, how much pain participants perceive in others, and how unpleasant they find witnessing such pain (Rütgen, Seidel, Riečanský, et al., 2015; Rütgen, Seidel, Silani, et al., 2015), and this effect can be blocked using a MOR antagonist (Rütgen, Seidel, Silani, et al., 2015). Also, the long term use of opioids, known to dysregulate the MOR system, leads to reduced pain ratings in others (Kroll et al., 2021). One study also showed that placebo analgesia reduces helping (Hartmann et al., 2022). Positron emission tomography (PET) with [^{11}C] carfentanil, a synthetic, highly specific MOR agonist allows in vivo imaging of opioid receptors in humans. Studies with PET have demonstrated endogenous opioid release in the thalamus following acute pain in a dose-dependent manner (Bencherif et al., 2002). Individual differences in MOR availability also link with pain sensitivity: participants with lower MOR availability have a higher sensitivity to pain (Hagelberg et al., 2012). Finally, and critically with respect to the present study, PET studies have linked opioid receptor availability with vicarious pain perception and sociability. First, MOR availability is negatively associated with haemodynamic responses to seeing others in pain (Karjalainen et al., 2017). The MOR system is also activated during positive social interactions such as laughing together (Manninen et al., 2017), and MOR availability is positively correlated with prosocial motivation as indexed by social attachment styles (Nummenmaa et al., 2015; Turtonen et al., 2021). Against this background, it could be hypothesized that the MOR system would be a crucial molecular pathway for altruistic, costly helping, but currently this hypothesis lacks direct in vivo evidence.

The current study
Here we investigated whether individual differences in the MOR availability at rest translate into measurable differences in the willingness to forfeit money to reduce pain to others, using the established "costly" helping paradigm of Gallo et al.. In the fMRI experiment, participants could choose to donate money to reduce the pain of the confederate who was subjected to electric shocks of varying intensity. In a separate

scanning session, the participants underwent a baseline PET scan with the MOR specific radioligand [^{11}C]carfentanil. We found that people were in general willing to engage in costly helping; The fMRI results revealed that activity in amydala, aIns, anterior cingulate cortex, striatum, primary motor cortex, primary somatosensory cortex, and thalamus increased when participants saw the confederate in pain. These haemodynamic responses had amplitudes that differed across individuals in ways that correlated with the availability of MORs in the striatolimbic and cortical emotion circuits. Altogether these data suggest that endogenous MOR system contributes to altruistic behavior and its individual differences.

Materials and methods

Participants

Thirty healthy Finnish women (mean ± SD age: 24.7±5.65 years, range 19-42) with normal or corrected to normal vision were recruited in the study. To maximize statistical power of the study, only females were included because the MOR system shows a high level of sexual dimorphism (Kantonen, Karjalainen, Isoj, et al., 2020). In addition, research on sex differences in empathy and pain shows that women are better at judging emotional signals (Hall & Matsumoto, 2004), show higher emotional responsivity (Christov-Moore et al., 2014), evaluate others' pain as more intense (Robinson & Wise, 2003), are more empathic than men (Hoffman, 1977), and are not influenced by trait harm sensitivity (FeldmanHall et al., 2016). All subjects participated in the fMRI scan and 14 of them participated in the PET scan. PET and MRI scans were conducted on separate days. Exclusion criteria included medications affecting the central nervous system, mood or anxiety disorders, psychotic disorders or neurological conditions, substance abuse and standard MRI and PET exclusion criteria. Structural brain abnormalities that are clinically relevant or could bias the analyses were excluded by a consultant neuroradiologist. The study was approved by the ethics board of the hospital district of Southwest Finland and conducted according to Good Clinical Practice and the Declaration of Helsinki (Approval number: 51/1801/2019). All subjects signed written informed consent and were informed that they had the right to withdraw at any time without giving a reason. The participants were compensated for their time and travel costs.

Experimental design and stimuli for fMRI

The fMRI study was run using the costly helping paradigm of Gallo et al. (2018). The participants were led to believe that they were witnessing the distress of another person in real time. Before the experiment, participants met the confederate and were explained

that the experiment would be performed in two separate rooms connected by a video camera. They were invited to draw lots to determine who would undergo the fMRI measurement while seeing the other participant receiving the electric shocks. The confederate was always chosen to receive the electric shocks. During the scan, participants, believing to witness the pain of the victim through a closed-circuit television, actually they viewed pre-recorded videos of the confederate receiving the painful simulation. All participants were debriefed at the end of the experiment.

The task was run by Presentation software (https://www.neurobs.com/) and the trial structure of the experiment is shown in **Figure 1.** On each trial, subjects first saw a video of the confederate receiving a painful electric shock (1st video). The intensity of expressed pain ranged randomly from 2 (mild pain) to 6 (moderate pain) out of ten. After the video, subjects could decide how much money they were willing to donate on that trial to reduce the intensity of the second shock in that trial, to do that, one button was held by the participant in each hand, one representing increasing money and the other representing decreasing money. For each trial, they were given 6€. They knew that if they did not donate any money, the second shock would have the same intensity as the first, whilst, for every donated 1€, the second shock would be reduced by one point on the 10-point pain scale. Participants were told that they can keep the undonated money of all trials divided by 10 as their extra compensation after the experiment. The subjects then saw the 2nd video of the confederate receiving the second electric shock whose intensity reflected the first shock minus the donation (see supplementary for more information on videos). There were two imaging runs with 15 trials in each.
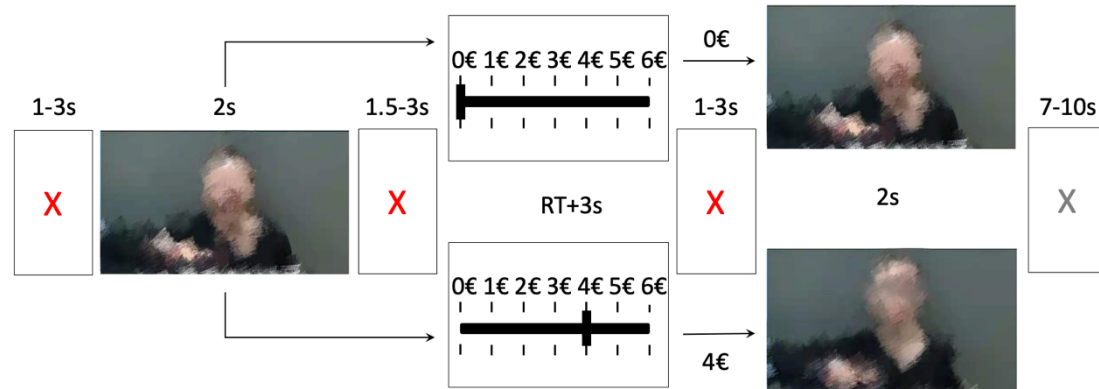

Figure 1. Trial structure. A red fixation cross was shown for 1-3s and was followed by 1st video for 2s. Then another red fixation cross was shown for 1.5-3s. The donation phase was self-paced and was followed by a red fixation cross (1-3 s). Next, the post-donation video was shown for 2s followed by a gray fixation cross (7-10s). The snapshots from the videos shown on the right illustrate two possible scenarios from the 6 possibilities.

MRI data acquisition and analysis

fMRI data were acquired with the 3T scanner (SIGNA, Premier, GE Healthcare, Waukesha, WI, USA) at the University Hospital of Turku. T2*-weighted functional images were collected with echo-planner imaging (EPI) sequence (45 slices; slice

thickness = 3 mm; TR = 2600 ms; TE = 30 ms; flip angle =75°; FOV = 24 mm; voxel size = 3×3×3 mm³). T1-weighted structural images were collected with voxel size of 1×1×1 mm³. MRI data were preprocessed with fMRIPrep 1.3.0.post2 (Esteban et al., 2019). The following preprocessing was performed on the T1-weighted (T1w) image: correction for intensity, skull-stripping, brain surfaces reconstruction, refined brain mask estimating, cortical gray-matter segmentations, spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c using nonlinear registration with antsRegistration (ANTs 2.2.0), and brain tissue segmentation. The following preprocessing was performed on the functional data: generation of reference volume and its skull-stripped version, co-registration to the T1w reference, slice-time correction, spatial smoothing with an isotropic, Gaussian Kernel of 6mm FWHM (full-width half-maximum), automatic removal of motion artifacts using ICA-AROMA (Pruim et al., 2015), and resampling to MNI152NLin2009cAsym standard space, and principal components are estimated after high-pass filtering the preprocessed BOLD time-series for the two CompCor variants: temporal and anatomical.

fMRI data were analyzed in SPM12 (Wellcome Trust Center for Imaging, London, UK, https://www.fil.ion.ucl.ac.uk/spm/). To investigate regions activated by i) seeing pain in the 1st video, ii) donating, and iii) seeing pain in the 2nd video, first-level general linear models (GLM) were estimated by modeling the 1st video, donation phase, and the 2nd video by using boxcar regressors in the design matrix. Donation size (trial-wise donations for each subject) was entered as parametric modulator for the 1st video. Subject-wise contrast images were then generated for main effects of 1st video, donation phase, 2nd video. Additionally, a subtraction contrast was computed for 1st vs. 2nd video and the parametric effect of donation size for 1st video. The contrast images were then subjected to second-level (random effects) analysis. Results are shown after FWE correction for cluster-size, by initially thresholding statistical maps at p_{unc}<0.001, identifying the FWEc minimum cluster-size value for FWE correction at the cluster-size level, and then thresholding the statistical maps again at p_{unc}<0.001 and k=FWEc-1.

PET data acquisition and analysis

PET data were acquired during resting baseline with a GE Discovery Molecular Insights DMI PET/CT, GE Healthcare, Waukesha WI in Turku PET Center. The high-affinity agonist radioligand [^{11}C]carfentanil was used to measure brain μ -opioid receptor availability. After intravenous radioligand injection (250.6±10.9 MBq), radioactivity in the brain was measured by the PET scanner for 51 minutes with increasing frame length (3×60s, 4×180s, 6×360s) using in-plane resolution of 3.77mm FWHM (Full Width Half Maximum) and tangential 4.00mm FWHM. All subjects lay supine in the PET scanner throughout the study. Data were corrected for dead-time, decay, and measured photon attenuation. In-house MAGIA pipeline was used to preprocess PET images (Kantonen, Karjalainen, Isojärvi, et al., 2020; Karjalainen et al., 2020).

Radiotracer binding was quantified using binding potential (BP_{ND}), calculated as the ratio of specific binding to non-displaceable binding in the tissue (Innis et al., 2007). This outcome measure is not confounded by differences in peripheral distribution or radiotracer metabolism, or alterations in brain perfusion. BP_{ND} is traditionally interpreted by target molecule density (B_{max}), although [^{11}C]carfentanil is also sensitive to endogenous neurotransmitter release. Accordingly, the BP_{ND} for the tracer should be interpreted as the density of the receptors unoccupied by endogenous ligand (i.e., receptor availability). Binding potential was calculated by applying the basis function method (Gunn et al., 1997) for each voxel using the simplified reference tissue model (Lammertsma & Hume, 1996), with occipital cortex serving as the reference region (Frost et al., 1989). The parametric images were spatially normalized to MNI-space via segmentation and normalization of T1-weighted anatomical images, and finally smoothed with an 8-mm FWHM Gaussian kernel. PET imaging with [^{11}C]carfentanil has high test-retest stability (Hirvonen et al., 2009). PET imaging always preceded fMRI to avoid potential impact of the fMRI tasks on measured MOR levels (mean ± SD day: 84±62 days, range 4-190).

ROI selection

The average tracer BP_{ND} was quantified in 17 anatomical *a priori* regions of interest (ROI) involved in vicarious pain and empathy: amygdala, caudate, cerebellum, dorsal anterior cingulate cortex, inferior temporal gyrus, insula, middle temporal gyrus, nucleus accumbens, orbitofrontal cortex, pars opercularis, posterior cingulate cortex, putamen, rostral anterior cingulate cortex, superior frontal gyrus, superior temporal gyrus, temporal pole, and thalamus. The selection was based on previous studies on the effects of MORs on vicarious pain and arousal (Karjalainen et al., 2017, 2019). The ROIs were derived separately for each subject from the T1-weighted MR images using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/).

PET-fMRI Fusion Analysis

Three approaches were taken to characterize the interactions between MOR availability and BOLD responses in pain perception and costly altruism. In the first two approaches, a principal component analysis (PCA) was used to reduce the dimensionality in the BP_{ND} values across our ROIs. This was done because regional MOR availability has high autocorrelation (Tuominen et al., 2014), thus PCA would increase the power of our analyses by reducing the multiple comparison correction that would otherwise reduce power. We found the first 3 PCs to explain >90% of the variance, with 61%, 22% and 7% of variance explained, respectively. To identify voxels with BOLD responses that depend on individual differences in the overall MOR signal, we used the first PC to predict the voxel-wise BOLD responses to the 1st video with donation size as a parametric modulator and the donation phase, separately. Specifically, we used the same parametric model as the fMRI analysis in the first-level model, and input the first PC in

the second-level model for 1st video and donation, in separate models. Second, to explore the relationship between individual MOR differences and responses in the vicarious pain observation network, we used the affective vicarious pain signature (AVPS) to dot-multiply the 1st level beta maps to 1st video with donation size as a parametric regressor for each participant (Zhou et al., 2020), thereby reducing each parameter estimate volume to a scalar value, and computed the correlation between the resulting value and the first 3 MOR PCs. Finally, to replicate previous studies on the links between MOR availability and haemodynamic responses to vicarious pain and arousal (Karjalainen et al., 2017, 2019), the voxel-wise BOLD responses to donation size in 1st video were predicted with ROI-wise [^{11}C]carfentanil binding potentials using whole-brain linear regression analysis with a statistical threshold set at $p < 0.05$, FWE-corrected at cluster-level. We then computed a cumulative map of the binarized MOR $\times$ BOLD beta maps to highlight regions whose BOLD responses were most consistently dependent on regional MOR availability.

Results

Behavioral results

Participants donated money in all intensity conditions (M = 2.826, SD = 1.964), and a linear mixed model confirmed that donations increased as a function of the shock intensity shown in the first video, $\beta = 0.877$, $SE = 0.023$, $t = 37.591$, $p < 0.001$, intercept is 0.198 (**Fig. 2**). The slope (beta=0.877) indicates that participants overall donate money to reduce shock intensity by 88%, adapting their donations very precisely to how much pain was at stake. In addition, in the linear mixed model, we added subject as a random effect. The variance of the effect from subject is 1.043, SD is 1.021.
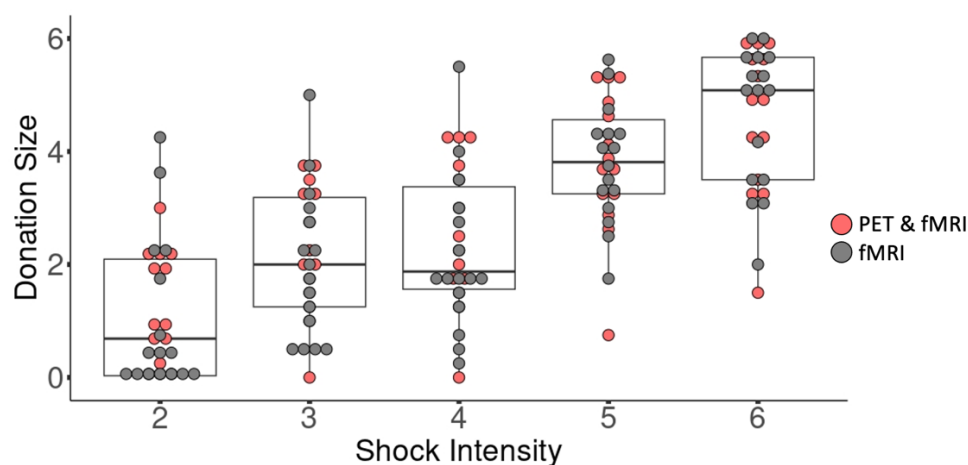

Figure 2. Donation size increased with increasing shock intensity. Each dot represents the average donation of a participant for all the stimuli of that intensity. N=30 participants. The box represents the 1st, 2nd and 3rd quartile, the whiskers the highest

and lowest value within 1.5 interquartile range. Red dots represent participants who undertook both PET and fMRI scans. Gray dots represent participants with only fMRI scans.

To explore whether donations depended on the MOR availability for the 14 participants who also underwent the PET scan, for each participant, we calculated the slope and intercept linking the participants donation with the intensity of shocks shown in 1st video (donation=slope*intensity+intercept). We then explored whether individual differences in slope or intercept were associated with individual differences in the MOR availability using the first 3 principal components of the MOR. We first did so using a multiple regression including all three PCs, which did not yield significant results for the slope ($F(3,10)=0.036$, $p=0.990$) or intercept ($F(3,10)=0.552$, $p=0.658$). Performing Bayesian correlation tests between the slope and intercept and each individual PCA (**Table 1**), confirms that this data is more likely under null hypothesis of no association. Together, this suggest that the actual donations do not depend on MOR availability in our ROIs. We therefore explored what brain circuits are involved in making the donation decisions, and whether these circuits vary based on MOR availability even if the outcomes of the decisions do not.

donation	MOR	Pearson's r	p	BF_{10}
slope	pca1	0.075	0.800	0.339
slope	pca2	0.013	0.965	0.329
slope	pca3	-0.070	0.813	0.338
intercept	pca1	-0.100	0.734	0.347
intercept	pca2	-0.136	0.642	0.363
intercept	pca3	0.337	0.239	0.621

Table 1. Lack of association between individual differences in slope or intercept and MOR PCs. For each of the 14 PET participants, we estimated the slope and intercept of the regression donation=slope*intensity+intercept. We then calculated the Pearson's r value between the 14 slope (top) or intercept (bottom rows) with Jasp (https://jasp-stats.org/).

BOLD-fMRI responses to vicarious pain perception

Full random effects results maps are available on NeuroVault (https://identifiers.org/neurovault.collection:14151). Consistent with previous studies, whole-brain general linear model analysis (GLM) revealed that anterior insula, anterior cingulate cortex (ACC) and thalamus were activated when seeing the confederate in pain in the 1st video (**Fig. 3**). Additional activations were observed in the amygdala and striatum, which are key nodes of the emotion and reward networks. We next modelled the BOLD signal with the donation size in the 1st video as a regressor. Consistent with previous research, insula, ACC activated more strongly as the donation increased (**Fig. 4**)

(Ioumpa et al., 2023; Jackson et al., 2005, 2006; Karjalainen et al., 2017; Lamm et al., 2011; Saarela et al., 2007; Singer et al., 2004; Soyman et al., 2022).
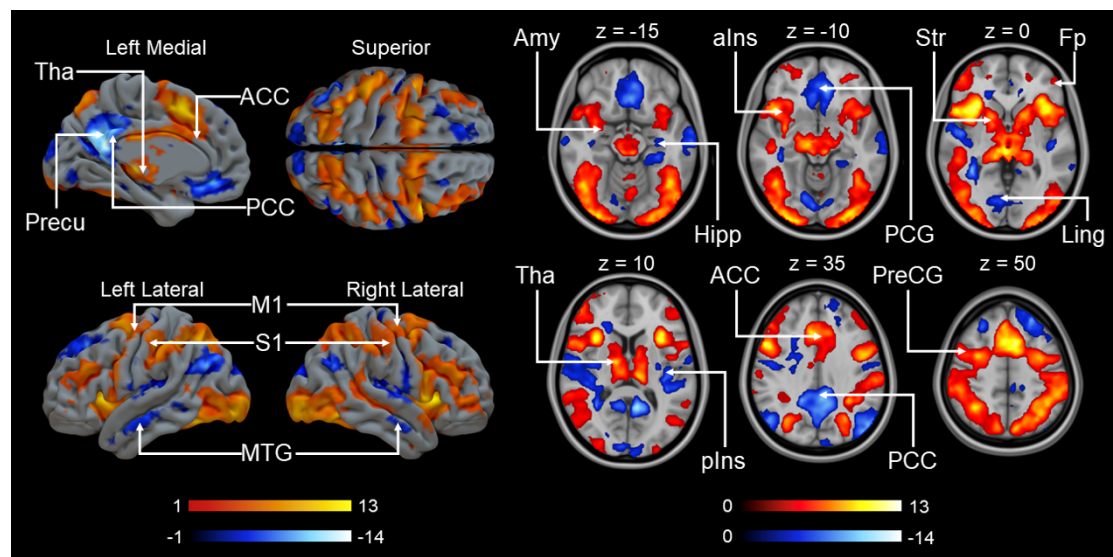

Figure 3. Main effect of brain responses to vicarious pain during the first video. The data are thresholded at p<0.001, FWE corrected at the cluster level (Positive: p_{unc}<0.001, k=FWEc=114 voxels, 3.40<t<13.17; Negative: p_{unc}<0.001, k=FWEc=97 voxels, 3.40<t<15.77). Colourbars indicate t statistic range. Tha = thalamus, Precu = precuneous cortex, ACC = anterior cingulate cortex, PCC = posterior cingulate cortex, M1 = primary motor cortex, S1 = primary somatosensory cortex, MTG = middle temporal gyrus, Amy = amygdala, Hipp = hippocampus, aIns = anterior insula, PCG = paracingulate gyrus, Str = striatum, Fp = frontal pole, Ling = lingual gyrus, pIns = posterior insula, PreCG = precentral gyrus.
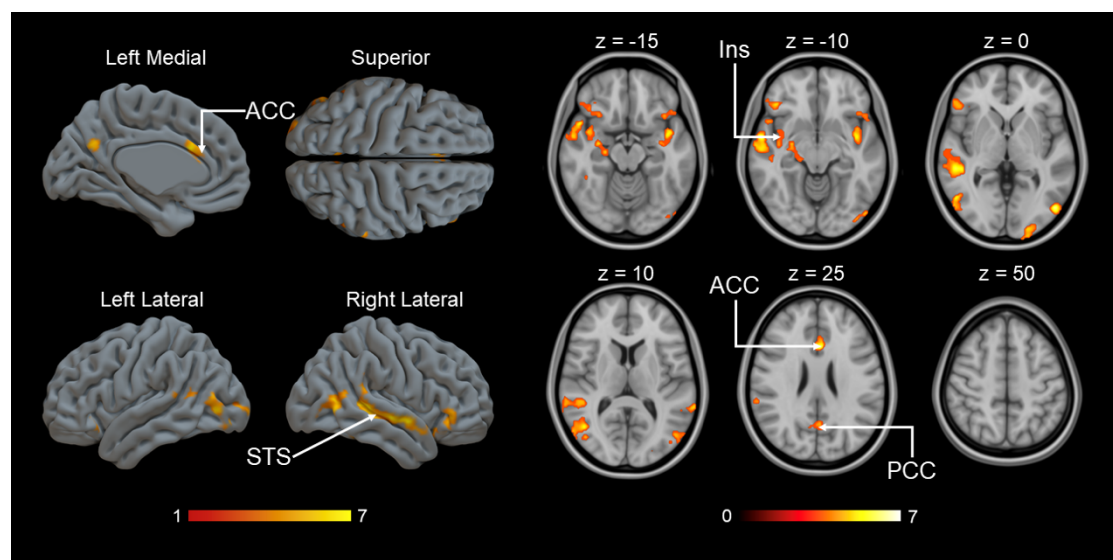

Figure 4. Brain regions with BOLD signals associated with donation size during the 1ˢᵗ video. Colourbars indicate t statistic range. The data are thresholded at p<0.001, FWE

corrected at the cluster level (p_{unc}<0.001, k=FWEc=163 voxels, 3.40<t<7.20). ACC = anterior cingulate cortex, PCC = posterior cingulate cortex, STS = superior temporal sulcus, Ins = insula.

BOLD-fMRI response changes following costly helping

In each trial, participants viewed two videos, one before their decision, and another following their decision. Comparing neural responses to the 1st and 2nd videos revealed decreased responses to the 2nd video in areas involved in vicarious pain perception such as, insula, thalamus and ACC. Additional effects were found in striatum, M1 and S1 (**Fig. 5**). Because subjects were on average willing to help, the 2nd video condition contained predominantly low shock intensity clips, thus direct comparison of 1st and 2nd could simply reflect differences in mean pain intensity across conditions. We subsequently restricted this analysis to low shock intensity levels (level 2 and level 3) in the 1st video. Even when pain intensity was thus approximately matched across the 1st and 2nd videos (Fig. 6), effects were similar to those in Fig. 5, with thalamus, striatum, ACC, M1 and S1 showing decreased responses to the 2nd video (**Fig. 6**).
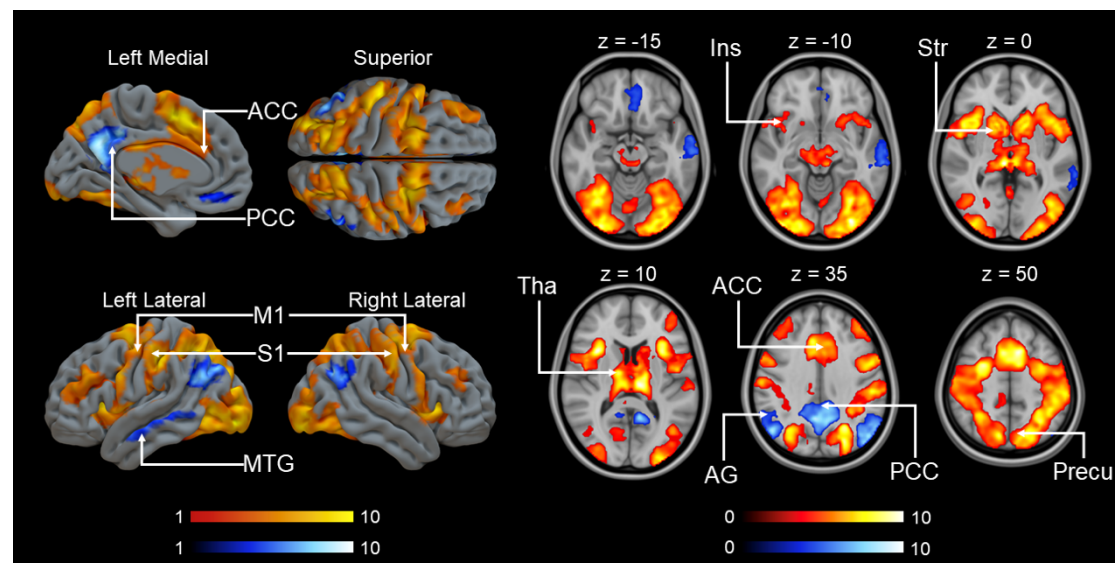

Figure 5. Brain activation for the 1st versus the 2nd video. The data are thresholded at p<0.001, FWE corrected at the cluster level (1st > 2nd : p_{unc}<0.001, k=FWEc=692 voxels, 3.40<t<10.85; 2nd > 1st : p_{unc}<0.001, k=FWEc=246 voxels, 3.40<t<10.24). Colourbars indicate t statistic range (red: 1st > 2nd video, blue: 1st < 2nd video). ACC = anterior cingulate cortex, PCC = posterior cingulate cortex, M1 = primary motor cortex, S1 = primary somatosensory cortex, MTG = middle temporal gyrus, Ins = insula, Str = striatum, Tha = thalamus, AG = angular gyrus, Precu = precuneus cortex.
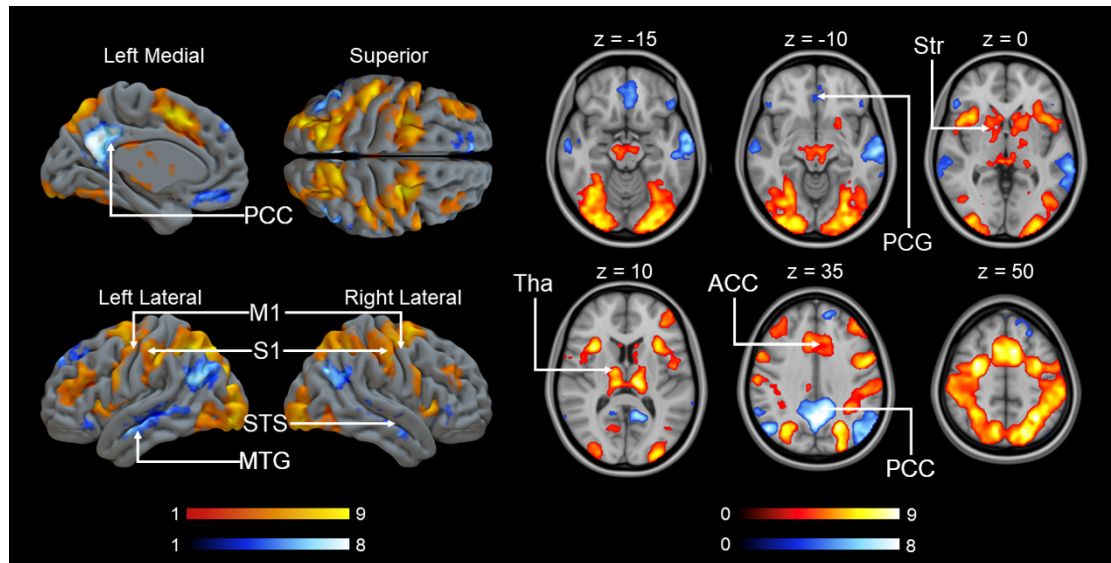
Figure 6. Brain activation for the low-level shock intensity trials in 1st video versus the 2nd video. The data are thresholded at p<0.001, FWE corrected at the cluster level (1st > 2nd : p_{unc}<0.001, k=FWEc=146 voxels, 3.40<t<9.99; 2nd > 1st : p_{unc}<0.001, k=FWEc=112 voxels, 3.40<t<10.40). Colourbars indicate t statistic range (red: 1st > 2nd video, blue: 1st < 2nd video). ACC = anterior cingulate cortex, PCC = posterior cingulate cortex, M1 = primary motor cortex, S1 = primary somatosensory cortex, STS = superior temporal sulcus, MTG = middle temporal gyrus, PCG = paracingulate gyrus, Str = striatum, Tha = thalamus.

MOR-dependent responses to empathy for pain

To test our hypothesis that individual differences in MOR availability would be associated with differences in the neural circuitry associated with taking helping decisions, we next modelled BOLD responses associated with donation size during the 1st video using regional MOR availabilities as predictors. We first used a PCA that captures the individual variability across the ROIs and extracted the first principal component that explained 60.7% of the individual variation in MOR binding (see **Table S1 for details**). The component scores were used as regressors to predict the haemodynamic response to donation size during the 1st video. We found a generally negative correlation between BOLD signal and MOR availability, mainly in amygdala, striatum, insula, hippocampus, thalamus, anterior cingulate cortex and posterior cingulate cortex (**Fig. 7**). This indicates that participants with reduced MOR availability show BOLD signals in these regions that are more sensitive to trial-by-trial differences in donation. Importantly, these differences in the association between brain activity and donation as a function of MOR availability were observed despite similar donations across participants with higher or lower MOR availability (**Table 1**), ascertaining that these differences are not related to systematic differences in the predictors inserted into the first level fMRI models (see Lebreton et al., 2019 for a discussion of the importance of this factor).

To focus more specifically on regions associated with witnessing the pain of others, we used the AVPS signature, which was found to be negatively associated with the third principal component (r = -0.659, p = 0.010, BF_{10}=6.5) (**Fig. S5**). We also generated a cumulative map of correlation between regional MOR availabilities and BOLD responses to donation size in the 1st video. The results replicated our prior study revealing a generally negative correlation between BOLD signal and MOR availability in vicarious pain related areas like the insula, anterior cingulate cortex and thalamus. Extensive association between MOR availability and brain activation was also observed in somatosensory areas, temporal gyrus, limbic regions and frontal cortex (**Fig. S2**) (Karjalainen et al., 2017).
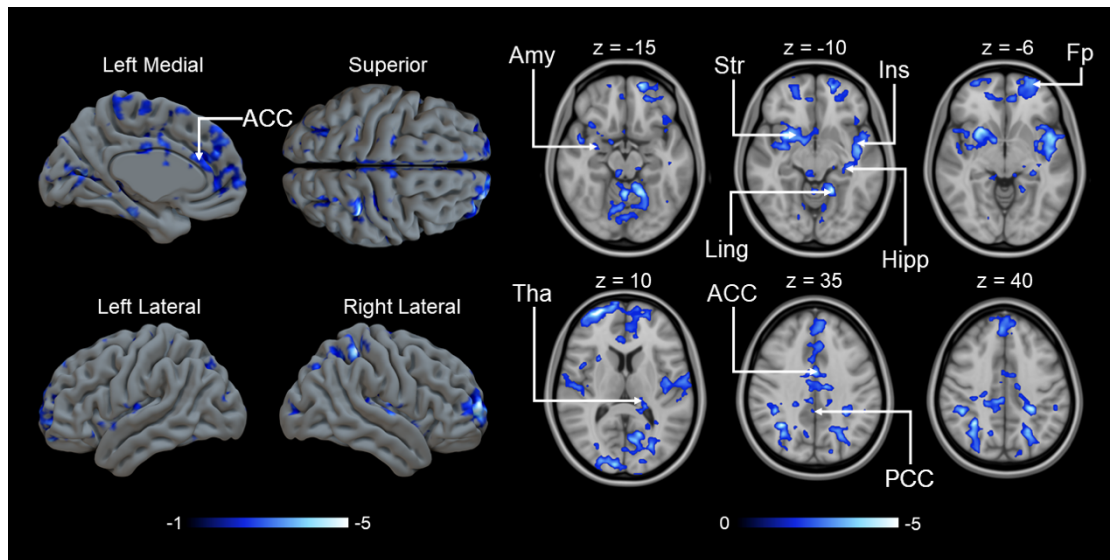

Figure 7. Negative correlation between the first component of MOR availability and the parametric modulator of BOLD response to donation size during 1st video (p<0.05, FEW corrected at cluster level, p_{unc}<0.05, k=FWEc=1800 voxels, 1.78<t<9.01). Amy = amygdala, Str = striatum, Ins = insula, Hipp = hippocampus, Ling = lingual gyrus, Fp = frontal pole, Tha = thalamus, ACC = anterior cingulate cortex, PCC = posterior cingulate cortex.

MOR-dependent responses to costly helping

Finally, we explored whether individual differences in MOR availability were associated with the BOLD activity during the donation phase, independently of what donation was given. As in the previous analysis, we used the 1st PC scores for the MOR availability maps to predict haemodynamic activation during the donation phase. Unlike the negative associations between MOR and the parametric modulator for donation size during the 1st video, this analysis revealed positive correlations between BOLD signal and MOR availability particularly in ACC and hippocampus (**Fig. 8**).
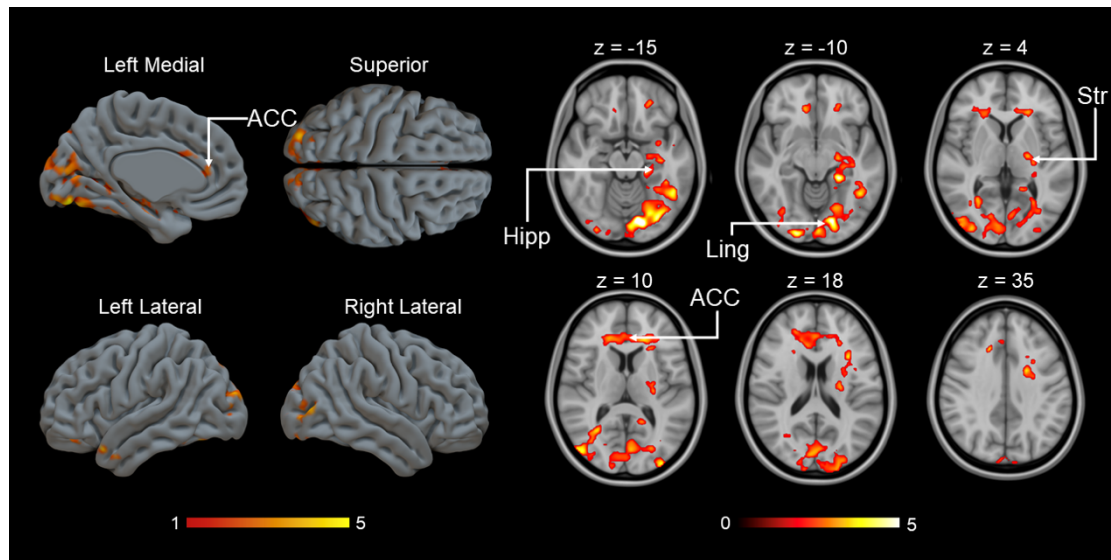
Figure 8. Positive correlation between MOR availability and BOLD response during the donation phase (p<0.05, FWE corrected at cluster level, Donation: p_{unc}<0.05, k=FWEc=10316 voxels, 1.78<t<8.44). Hipp = hippocampus, ACC = anterior cingulate cortex, Ling = lingual gyrus, Str = striatum.

Discussion

Our main finding was that individual differences in the endogenous opioid system tone do not directly alter participants' decisions to help others, but are linked with brain activity differences during vicarious pain and costly, altruistic helping. Our subjects were generally willing to dive up significant amounts of money to help others and donated larger sums when they saw the confederate experiencing stronger pain. Trials in which participants donated more money were associated with increased BOLD activity in regions associated with empathy and the central nociceptive system (Ins, ACC, PCC STS). In line with studies showing that reduced baseline MOR availability is associated with heightened sensitivity to pain (Karjalainen et al., 2017), baseline MOR availability was negatively correlated with the BOLD responses while witnessing the pain of others in 1st video associated with donating to help. Donation depended on how much pain was displayed in the 1st video, and participants with reduced MOR availability had activity in circuits associated with empathy and nociception that were more strongly associated with donation. Also, brain activity during the donation phase, after they seeing how much pain the other participants expressed, depended on endogenous MOR tone in hippocampus, ACC, and striatum. These data provide the first direct *in vivo* evidence for the engagement of the opioid system in the neural processes occurring during costly helping, significantly extending the role of MORs in human social behavior.

Neural basis of vicarious pain perception

As our goal was to model authentic altruistic helping that would also be costly to the participant, we used a naturalistic interactive setting in which we made participants believe that they were interacting with a real person whom they also met before the experiment. Behavioral data confirmed that the manipulation successfully induced costly helping: Participants donated more money when the shock intensity increased and the confederate expressed more intense pain. This finding accords with previous studies (Gallo et al., 2018) and suggests that people are willing to altruistically help strangers that they have met only recently. FMRI data revealed that witnessing the pain of the others during the 1st video evoked widespread cortical and subcortical activation, in regions associated with empathic pain (aIns, ACC, M1, S1, thalamus, amygdala, and striatum). The results are consistent with meta-analyses of brain regions associated with empathy for pain (Jauniaux et al., 2019; Lamm et al., 2011; Timmers et al., 2018). ACC, PCC, insula, and STS response during 1st video were stronger in trials in which participants later decided to donate more money. Within the limitations of fMRI, activity in these regions may thus have played a role in motivating costly helping. This dovetails with findings from previous studies that showed BOLD activity in similar networks of regions scaled with perceived pain in Hollywood-type videos with various intensities of pain (Karjalainen et al., 2017), and for the insula, recent intracranial recordings showing that the power in broadband gamma and the spiking of single neurons in this region scales with the perceived pain intensity for similar stimuli (Soyman et al., 2022). Importantly, this also matches findings using a similar task acquired in a different lab (Ioumpa et al., 2023).

Endogenous opioids and vicarious pain perception

We observed a negative association between cerebral MOR availability and the relationship between BOLD activity and donation while viewing 1st video across a wide range of brain regions involved in vicarious pain perception in three different ways. This association was significant in regions associated with empathy for pain (insula, ACC, thalamus, amygdala, and striatum). This negative association accords with previous findings associating lower MOR availability with higher sensitivity to pain (Hagelberg et al., 2012), more distress (Nummenmaa et al., 2020) as well as acute adverse emotions evoked by witnessing the pain of others (Karjalainen et al., 2017). Taken together these data suggest that lower MOR availability makes individuals more sensitive to suffer of others. These data also accord with the general role of the opioid system in maintaining social bonds and attachment (Manninen et al., 2017; Sun et al., 2022; Turtonen et al., 2021), which was here conceptualized as hemodynamic responses to seeing others in distress.

Brain basis of costly helping

A peculiarity of the paradigm developed by Gallo et al. (2018) employed here is that participants can directly monitor the effectiveness of their costly helping by comparing 1st video and 2nd video. Comparing the brain activity across the 1st and 2nd videos (i.e., before and after helping), we found that the 1st video evoked stronger activation in striatum, thalamus, ACC, M1, and S1, whether we compared all 1st video against all 2nd video, or repeat the analyses selecting instances in which 1st video was of low intensity. This pattern was spatially similar to that elicited by the 1st video only. One interpretation why the activity was lower in 2nd video than 1st video (even if the two videos showed similar levels of pain, Fig. 6) is that the opportunity to help decreased the vicarious pain response. This would accord with the negative state relief model for altruism, which states that helping others makes the helper feel better (Dovidio & Penner, 2003; Williamson & Clark, 1989). Altruistic behavior could thus be motivated by the (anticipated) alleviation of vicarious pain. Some have argued that altruism is driven by the rewarding nature of empathy and helping (Batson et al., 2014; Dovidio & Penner, 2003; FeldmanHall et al., 2015), which might predict helping-induced activations in the brain's striatal reward systems (**Figure 5** and **supplementary figure S3**), which we, however, failed to find in our contrast between the 1st and 2nd videos.

Accordingly, within the limits of reverse inference, the neural data are perhaps best explained by the notion that helping is motivated by an anticipation of reduced vicarious pain/distress rather than by the anticipation of reward. It should, however, be noted that several alternative explanations could account for the observed reduction in brain responses from 1st video to 2nd video. First, 1st video is relevant to the task given to the participants, and the association between the intensity of pain perceivable in 1st video and the donation confirms that participants adapted their responses to the content of 1st video. In contrast, 2nd video is task-irrelevant. Task relevance and the attention it commands, may thus account for the intensity of the response to 1st video. Second, the intensity of 1st video is unpredictable, while the intensity of 2nd video is predictable based on 1st video and the donation. Given that a network similar to the one we observed here has been shown to encode prediction errors while witnessing the pain of others (Fornari et al., 2023), this difference in predictability may also account for the differences in BOLD responses. Together, the observed differences in BOLD activity across 1st video and second should thus be interpreted with caution.

Opioid system modulates brain activity during costly helping

As mentioned in the introduction, networks supporting costly helping have been described, but the role played by the opioid system remains poorly understood. Here we demonstrate that individual differences in baseline μ-receptor availability also relate to the individual differences in brain activity while participants plan and report their costly-helping decision. We observed a predominantly positive correlation between cerebral

MOR availability and an extensive BOLD response in the helping phase (donation) in hippocampus, ACC, and striatum. Prior PET studies have illustrated the modulatory role of the opioid system in emotion, vicarious pain, positive social interaction, prosocial motivation in humans, and prosocial behavior in monkeys (Fabre-Nys et al., 1982; Karjalainen et al., 2017; Keverne et al., 1989; Nummenmaa & Tuominen, 2018). Our findings thus broaden our knowledge of the function of the opioid system, demonstrating its modulating role in the neural processing of costly helping. Importantly, MOR availability had, in general, opposite relationship with haemodynamic responses in the cortical and subcortical pain and emotion circuits during vicarious pain perception (negative) and deciding to help (positive). This suggests that during vicarious pain perception, the MORs may act as a buffer against the stress evoked by seeing others in distress. Yet when relieving others' stress via one's own decisions becomes possible (here during donation phase), individuals with high MOR tone show amplified hemodynamic responses. Thus, individuals with high MOR tone might shift their brain activity from the moment of witnessing other people's sufferance (Video1) to the moment of deciding to help them (decision phase). This suggests that MOR tone could be an important molecular pathway modulating altruism and sociability via multiple mechanisms.

Contrary to our findings that individual differences in brain activity during costly helping are associated with differences in MOR availability, we did not find a significant correlation between MOR availability and donation size. This is slightly unexpected given the association at the neural level with brain regions associated with donation size, but in general, brain-behaviour relationships require substantially larger samples than what can be achieved with invasive experimental PET studies (Marek et al., 2022). However, given the behavioural pharmacological studies pointing towards causal effects of opiates on helping (Hartmann et al., 2022) our data contributes to a nascent literature associating the MOR system with prosocial decision-making.

Limitations

In a single PET scan, it is impossible to demonstrate the exact molecule-level mechanism for altered receptor availability (Kantonen, Karjalainen, Isoj, et al., 2020). Our single PET scan study design only allowed quantifying differences in baseline receptor binding, but not the capacity for endogenous opioid release. We also only studied females so the results may also not generalize to males, given sex differences in empathy (Christov-Moore et al., 2014; Hoffman, 1977) and MOR availability (Kantonen, Karjalainen, Isoj, et al., 2020). Finally, PET and fMRI data were not measured simultaneously, yet prior studies have established that [^{11}C]carfentanil has excellent test-retest reliability even with multiple-months intervals.

Conclusion

Placebo analgesia studies have suggested that the opioid system may contribute to costly helping decisions. We provide evidence that individual differences in μ-opioid baseline availability can explain significant individual variability in how the brain processes the distress of others in a costly helping context, and how it processes the decision to help. Brain regions associated with empathic pain such as the anterior insula, ACC, thalamus, amygdala, striatum, were significantly activated while perceiving the pain of others, and more so on trials in which participants later decided to help more. Activation of these regions decreased following helping. MOR availability was negatively correlated with the processing of the pain of others but dominantly positively correlated with neural responses while making the decision to help. These results suggest that the opioid system is intimately involved in vicarious pain and neural processing of helping decisions.

Acknowledgements

References

Batson, C. D. (2010). Empathy-induced altruistic motivation. *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature.*, 15–34. https://doi.org/10.1037/12061-001

Batson, C. D. (2011). *Altruism in Humans.* Oxford University Press. https://doi.org/https://doi.org/10.1093/acprof:oso/9780195341065.001.0001

Batson, C. D., Lishner, A. D., & Stocks, L. E. (2014). The Empathy–Altruism Hypothesis. In *The Oxford Handbook of Prosocial Behavior.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399813.013.023

Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is Empathy-Induced Helping Due to Self-Other Merging? *Journal of Personality and Social Psychology*, *73*(3), 495–509. https://doi.org/10.1037/0022-3514.73.3.495

Bencherif, B., Fuchs, P. N., Sheth, R., Dannals, R. F., Campbell, J. N., & Frost, J. J. (2002). Pain activation of human supraspinal opioid pathways as demonstrated by [11C]-carfentanil and positron emission tomography (PET). *Pain*, *99*(3), 589–598. https://doi.org/10.1016/S0304-3959(02)00266-X

Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., & Aglioti, S. M. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cerebral Cortex*, *17*(11), 2553–2561. https://doi.org/10.1093/cercor/bhl161

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, G. P. (2000). PROSOCIAL FOUNDATIONS OF CHILDREN'S ACADEMIC ACHIEVEMENT. *Psychological Science*, *11*(4), 302–306. https://doi.org/10.1111/1467-9280.00260

Caprara, G. V., & Steca, P. (2005). Self-efficacy beliefs as determinants of prosocial behavior conducive to life satisfaction across ages. *Journal of Social and Clinical Psychology*, *24*(2), 191–217. https://doi.org/10.1521/jscp.24.2.191.62271

Carrillo, M., Han, Y., Migliorati, F., Liu, M., Gazzola, V., & Keysers, C. (2019). Emotional Mirror Neurons in the Rat's Anterior Cingulate Cortex. *Current Biology*, *29*(8), 1301-1312.e6. https://doi.org/10.1016/j.cub.2019.03.024

Caspar, E. A., Ioumpa, K., Arnaldo, I., Di Angelis, L., Gazzola, V., & Keysers, C. (2022). Commanding or Being a Simple Intermediary: How Does It Affect Moral Behavior and Related Brain Mechanisms? *ENeuro*, *9*(5), 1–24. https://doi.org/10.1523/ENEURO.0508-21.2022

Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience and Biobehavioral Reviews*, *46*(P4), 604–627. https://doi.org/10.1016/j.neubiorev.2014.09.001

Darcq, E., & Kieffer, B. L. (2018). Opioid receptors: Drivers to addiction? *Nature Reviews Neuroscience*, *19*(8), 499–514. https://doi.org/10.1038/s41583-018-0028-x

De Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: Behavioural manifestations and neural basis. *Nature Reviews Neuroscience*, *18*(8), 498–509. https://doi.org/10.1038/nrn.2017.72

Decety, J., Bartal, I. B. A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: Highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686). https://doi.org/10.1098/rstb.2015.0077

Deković, M., & Gerris, J. R. M. (1994). Developmental analysis of social cognitive and behavioral differences between popular and rejected children. *Journal of Applied Developmental Psychology*, *15*(3), 367–386. https://doi.org/10.1016/0193-3973(94)90038-8

Dovidio, J. F., & Penner, L. A. (2003). Helping and Altruism. In *Blackwell Handbook of Social Psychology: Interpersonal Processes*. https://doi.org/https://doi.org/10.1002/9780470998557.ch7

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Fabre-Nys, C., Meller, R. E., & Keverne, E. B. (1982). Opiate antagonists stimulate affiliative behaviour in monkeys. *Pharmacology, Biochemistry and Behavior*, *16*(4), 653–659. https://doi.org/10.1016/0091-3057(82)90432-4

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage*, *105*, 347–356. https://doi.org/10.1016/j.neuroimage.2014.10.043

FeldmanHall, O., Dalgleish, T., Evans, D., Navrady, L., Tedeschi, E., & Mobbs, D. (2016). Moral Chivalry: Gender and Harm Sensitivity Predict Costly Altruism. *Social Psychological and Personality Science*, *7*(6), 542–551. https://doi.org/10.1177/1948550616647448

FeldmanHall, O., Dalgleish, T., & Mobbs, D. (2013). Alexithymia decreases altruism in real social decisions. *Cortex*, *49*(3), 899–904. https://doi.org/10.1016/j.cortex.2012.10.015

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., & Mobbs, D. (2012). Differential neural circuitry and self-interest in real vs hypothetical moral decisions. *Social Cognitive and Affective Neuroscience*, *7*(7), 743–751. https://doi.org/10.1093/scan/nss069

Fornari, L., Ioumpa, K., Nostro, A. D., Evans, N. J., De Angelis, L., Speer, S. P. H., Paracampo, R., Gallo, S., Spezio, M., Keysers, C., & Gazzola, V. (2023). Neuro-computational mechanisms and individual biases in action-outcome learning under moral conflict. *Nature Communications*, *14*(1), 1–18. https://doi.org/10.1038/s41467-023-36807-3

Frost, J. J., Douglass, K. H., Mayberg, H. S., Dannals, R. F., Links, J. M., Wilson, A. A., Ravert, H. T., Crozier, W. C., & Wagner, H. N. (1989). Multicompartmental analysis of [11C]-carfentanil binding to opiate receptors in humans measured by positron emission tomography. *Journal of Cerebral Blood Flow and Metabolism*, *9*(3), 398–409. https://doi.org/10.1038/jcbfm.1989.59

Gallo, S., Paracampo, R., Müller-Pinzler, L., Severo, M. C., Blömer, L., Fernandes-Henriques, C., Henschel, A., Lammes, B. K., Maskaljunas, T., Suttrup, J., Avenanti, A., Keysers, C., & Gazzola, V. (2018). The causal role of the somatosensory cortex in prosocial behaviour. *ELife*, *7*, 1–31. https://doi.org/10.7554/eLife.32740

Gunn, R. N., Lammertsma, A. A., Hume, S. P., & Cunningham, V. J. (1997). Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage*, *6*(4), 279–287. https://doi.org/10.1006/nimg.1997.0303

Hagelberg, N., Aalto, S., Tuominen, L., Pesonen, U., Någren, K., Hietala, J., Scheinin, H., Pertovaara, A., & Martikainen, I. K. (2012). Striatal μ-opioid receptor availability predicts cold pressor pain threshold in healthy human subjects. *Neuroscience Letters*, *521*(1), 11–14. https://doi.org/10.1016/j.neulet.2012.05.042

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*(2), 201–206. https://doi.org/10.1037/1528-3542.4.2.201

Hartmann, H., Forbes, P. A. G., Rütgen, M., & Lamm, C. (2022). Placebo Analgesia Reduces Costly Prosocial Helping to Lower Another Person's Pain. *Psychological Science*, *33*(11), 1867–1881. https://doi.org/10.1177/09567976221119727

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149–160. https://doi.org/10.1016/j.neuron.2010.09.003

Henriksen, G., & Willoch, F. (2008). Imaging of opioid receptors in the central nervous system. *Brain*, *131*(5), 1171–1196. https://doi.org/10.1093/brain/awm255

Hirvonen, J., Aalto, S., Hagelberg, N., Maksimow, A., Ingman, K., Oikonen, V., Virkkala, J., Någren, K., & Scheinin, H. (2009). Measurement of central μ-opioid receptor binding in vivo with PET and [11C]carfentanil: A test-retest study in healthy subjects. *European Journal of Nuclear Medicine and Molecular Imaging*, *36*(2), 275–286. https://doi.org/10.1007/s00259-008-0935-6

Hoffman, M. L. (1977). Sex differences in empathy and related behaviors. *Psychological Bulletin*, *84*(4), 712–722. https://doi.org/10.1037/0033-2909.84.4.712

Innis, R. B., Cunningham, V. J., Delforge, J., Fujita, M., Gjedde, A., Gunn, R. N., Holden, J., Houle, S., Huang, S. C., Ichise, M., Iida, H., Ito, H., Kimura, Y., Koeppe, R. A., Knudsen,

G. M., Knuuti, J., Lammertsma, A. A., Laruelle, M., Logan, J., … Carson, R. E. (2007). Consensus nomenclature for in vivo imaging of reversibly binding radioligands. In *Journal of Cerebral Blood Flow and Metabolism* (Vol. 27, Issue 9, pp. 1533–1539). https://doi.org/10.1038/sj.jcbfm.9600493

Ioumpa, K., Gallo, S., Keysers, C., & Gazzola, V. (2023). Neural mechanisms of costly helping in the general population and mirror-pain synesthetes. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2023.03.09.531639

Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). *Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. 44*, 752–761. https://doi.org/10.1016/j.neuropsychologia.2005.07.015

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). *How do we perceive the pain of others ? A window into the neural processes involved in empathy. 24*, 771–779. https://doi.org/10.1016/j.neuroimage.2004.09.006

Jauniaux, J., Khatibi, A., Rainville, P., & Jackson, P. L. (2019). A meta-analysis of neuroimaging studies on pain empathy: Investigating the role of visual information and observers' perspective. *Social Cognitive and Affective Neuroscience*, *14*(8), 789–813. https://doi.org/10.1093/scan/nsz055

Kantonen, T., Karjalainen, T., Isoj, J., Rinne, J., Hietala, J., Kaasinen, V., Kalliokoski, K., Scheinin, H., Hirvonen, J., Vehtari, A., & Nummenmaa, L. (2020). *NeuroImage Interindividual variability and lateralization of μ -opioid receptors in the human brain. 217*(April). https://doi.org/10.1016/j.neuroimage.2020.116922

Kantonen, T., Karjalainen, T., Isojärvi, J., Nuutila, P., Tuisku, J., Rinne, J., Hietala, J., Kaasinen, V., Kalliokoski, K., Scheinin, H., Hirvonen, J., Vehtari, A., & Nummenmaa, L. (2020). Interindividual variability and lateralization of μ-opioid receptors in the human brain: Individual differences in the μ-opioid receptor system. *NeuroImage*, *217*(April). https://doi.org/10.1016/j.neuroimage.2020.116922

Karjalainen, T., Karlsson, H. K., Lahnakoski, J. M., Glerean, E., Nuutila, P., Jääskeläinen, I. P., Hari, R., Sams, M., & Nummenmaa, L. (2017). Dissociable roles of cerebral μ-opioid and type 2 dopamine receptors in vicarious pain: A combined PET-fMRI study. *Cerebral Cortex*, *27*(8), 4257–4266. https://doi.org/10.1093/cercor/bhx129

Karjalainen, T., Seppälä, K., Glerean, E., Karlsson, H. K., Lahnakoski, J. M., Nuutila, P., Jääskeläinen, I. P., Hari, R., Sams, M., & Nummenmaa, L. (2019). Opioidergic Regulation of Emotional Arousal: A Combined PET-fMRI Study. *Cerebral Cortex*, *29*(9), 4006–4016. https://doi.org/10.1093/cercor/bhy281

Karjalainen, T., Tuisku, J., Santavirta, S., Kantonen, T., & Bucci, M. (2020). *Magia : Robust Automated Image Processing and Kinetic Modeling Toolbox for PET Neuroinformatics. 14*(February), 1–13. https://doi.org/10.3389/fninf.2020.00003

Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology*, *14*(1–2), 155–161. https://doi.org/10.1016/0306-4530(89)90065-6

Keysers, C., & Gazzola, V. (2014). Dissociating the ability and propensity for empathy. *Trends in Cognitive Sciences*, *18*(4), 163–166. https://doi.org/10.1016/j.tics.2013.12.011

Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews Neuroscience*, *11*(6), 417–428. https://doi.org/10.1038/nrn2833

bioRxiv preprint doi: https://doi.org/10.1101/2023.12.06.569026; this version posted December 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Keysers, C., Knapska, E., Moita, M. A., & Gazzola, V. (2022). Emotional contagion and prosocial behavior in rodents. *Trends in Cognitive Sciences*, *26*(8), 688–706. https://doi.org/10.1016/j.tics.2022.05.005

Krishnan, A., Woo, C. W., Chang, L. J., Ruzic, L., Gu, X., López-Solà, M., Jackson, P. L., Pujo, J., Fan, J., & Wager, T. D. (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *ELife*, *5*(JUN2016), 1–42. https://doi.org/10.7554/eLife.15166

Kroll, S. L., Thayer, J. F., Williams, D. W. P., Pfabigan, D. M., Baumgartner, M. R., Lamm, C., & Quednow, B. B. (2021). Chronic non-medical prescription opioid use and empathy for pain: Does pain make the difference? *Psychophysiology*, *58*(4), 1–16. https://doi.org/10.1111/psyp.13776

Laguna, M., Mazur, Z., Kędra, M., & Ostrowski, K. (2020). Interventions stimulating prosocial helping behavior: A systematic review. *Journal of Applied Social Psychology*, *50*(11), 676–696. https://doi.org/10.1111/jasp.12704

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Lammertsma, A. A., & Hume, S. P. (1996). Simplified reference tissue model for PET receptor studies. *NeuroImage*. https://doi.org/10.1006/nimg.1996.0066

Lebreton, M., Bavard, S., Daunizeau, J., & Palminteri, S. (2019). Assessing inter-individual differences with task-related functional neuroimaging. *Nature Human Behaviour*, *3*(9), 897–905. https://doi.org/10.1038/s41562-019-0681-8

Ma, Y., Wang, C., & Han, S. (2011). Neural responses to perceived pain in others predict real-life monetary donations in different socioeconomic contexts. *NeuroImage*, *57*(3), 1273–1280. https://doi.org/10.1016/j.neuroimage.2011.05.003

Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E., Hari, R., Jääskeläinen, I. P., Sams, M., & Nummenmaa, L. (2017). Social laughter triggers endogenous opioid release in humans. *Journal of Neuroscience*, *37*(25), 6125–6131. https://doi.org/10.1523/JNEUROSCI.0688-16.2017

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654–660. https://doi.org/10.1038/s41586-022-04492-9

Nummenmaa, L., Karjalainen, T., Isojärvi, J., Kantonen, T., Tuisku, J., Kaasinen, V., Joutsa, J., Nuutila, P., Kalliokoski, K., Hirvonen, J., Hietala, J., & Rinne, J. (2020). Lowered endogenous mu-opioid receptor availability in subclinical depression and anxiety. *Neuropsychopharmacology*, *45*(11), 1953–1959. https://doi.org/10.1038/s41386-020-0725-9

Nummenmaa, L., Manninen, S., Tuominen, L., Hirvonen, J., Kalliokoski, K. K., Nuutila, P., Jääskeläinen, I. P., Hari, R., Dunbar, R. I. M., & Sams, M. (2015). Adult attachment style is associated with cerebral μ-opioid receptor availability in humans. *Human Brain Mapping*, *36*(9), 3621–3628. https://doi.org/10.1002/hbm.22866

Nummenmaa, L., & Tuominen, L. (2018). Opioid system and human emotions. *British Journal of Pharmacology*, *175*(14), 2737–2749. https://doi.org/10.1111/bph.13812

Paradiso, E., Gazzola, V., & Keysers, C. (2021). Neural mechanisms necessary for empathy-related phenomena across species. *Current Opinion in Neurobiology*, *68*, 107–115. https://doi.org/10.1016/j.conb.2021.02.005

Pruim, R. H. R., Mennes, M., Rooij, D. Van, Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). NeuroImage ICA-AROMA : A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, *112*, 267–277. https://doi.org/10.1016/j.neuroimage.2015.02.064

Robinson, M. E., & Wise, E. A. (2003). Gender bias in the observation of experimental pain. *Pain*, *104*(1–2), 259–264. https://doi.org/10.1016/S0304-3959(03)00014-9

Rütgen, M., Seidel, E. M., Riečanský, I., & Lamm, C. (2015). Reduction of empathy for pain by placebo analgesia suggests functional equivalence of empathy and first-hand emotion experience. *Journal of Neuroscience*, *35*(23), 8938–8947. https://doi.org/10.1523/JNEUROSCI.3936-14.2015

Rütgen, M., Seidel, E. M., Silani, G., Riečanský, I., Hummer, A., Windischberger, C., Petrovic, P., & Lamm, C. (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(41), E5638–E5646. https://doi.org/10.1073/pnas.1511269112

Saarela, M. V, Hlushchuk, Y., & C, A. C. De. (2007). *The Compassionate Brain : Humans Detect Intensity of Pain from Another ' s Face. January.* https://doi.org/10.1093/cercor/bhj141

Singer, T., Seymour, B., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, *303*, 1157–1162. https://doi.org/https://www.science.org/doi/10.1126/science.1093535

Smith, A. (1759). *The Theory of Moral Sentiments.* https://doi.org/https://doi.org/10.1017/CBO9780511800153

Soyman, E., Bruls, R., Ioumpa, K., Müller-Pinzler, L., Gallo, S., Qin, C., van Straaten, E. C. W., Self, M. W., Peters, J. C., Possel, J. K., Onuki, Y., Baayen, J. C., Idema, S., Keysers, C., & Gazzola, V. (2022). Intracranial human recordings reveal association between neural activity and perceived intensity for the pain of others in the insula. *ELife*, *11*, 1–34. https://doi.org/10.7554/eLife.75197

Sun, L., Lukkarinen, L., Putkinen, V., Karlsson, H. K., Hirvonen, J., Tiihonen, J., Lauerma, H., Scott, S., & Nummenmaa, L. (2022). Mu-opioid receptor system modulates responses to vocal bonding and distress signals in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1863). https://doi.org/10.1098/rstb.2021.0181

Timmers, I., Park, A. L., Fischer, M. D., Kronman, C. A., Heathcote, L. C., Hernandez, J. M., & Simons, L. E. (2018). Is empathy for pain unique in its neural correlates? A meta-analysis of neuroimaging studies of empathy. *Frontiers in Behavioral Neuroscience*, *12*(November), 1–12. https://doi.org/10.3389/fnbeh.2018.00289

Tuominen, L., Nummenmaa, L., Keltikangas-Järvinen, L., Raitakari, O., & Hietala, J. (2014). Mapping neurotransmitter networks with PET: An example on serotonin and opioid systems. *Human Brain Mapping*, *35*(5), 1875–1884. https://doi.org/10.1002/hbm.22298

Turtonen, O., Saarinen, A., Nummenmaa, L., Tuominen, L., Tikka, M., Armio, R. L., Hautamäki, A., Laurikainen, H., Raitakari, O., Keltikangas-Järvinen, L., & Hietala, J. (2021). Adult Attachment System Links With Brain Mu Opioid Receptor Availability In Vivo. *Biological*

Psychiatry: Cognitive Neuroscience and Neuroimaging, *6*(3), 360–369.
https://doi.org/10.1016/j.bpsc.2020.10.013

Williamson, G. M., & Clark, M. S. (1989). Providing Help and Desired Relationship Type as Determinants of Changes in Moods and Self-Evaluations. *Journal of Personality and Social Psychology*, *56*(5), 722–734. https://doi.org/10.1037/0022-3514.56.5.722

Zanon, M., Novembre, G., Zangrando, N., Chittaro, L., & Silani, G. (2014). Brain activity and prosocial behavior in a simulated life-threatening situation. *NeuroImage*, *98*, 134–146. https://doi.org/10.1016/j.neuroimage.2014.04.053

Zhou, F., Li, J., Zhao, W., Xu, L., Zheng, X., Fu, M., Yao, S., Kendrick, K. M., Wager, T. D., & Becker, B. (2020). Empathic pain evoked by sensory and emotional-communicative cues share common and process-specific neural representations. *ELife*, *9*, 1–27. https://doi.org/10.7554/ELIFE.56929

Supplementary Materials

Stimuli

All videos played in the experiment (each video lasted 2s) were recorded in advance to span different shock levels (maximum to minimum: 6 to 2). All videos featured the same black background and black table. The participants could see the confederate's face and right hand, and the intensity of the electric shock could only be detected by changes in the facial expression, mainly by change of eye brows and mouth. The location of the hand was kept similar cross movies using a mark on the table that was covered by the hand, and hence invisible to the viewers of the movies. When shooting the videos, rather than giving a real electric shock to the hand, we poked the actor on her thigh to induce her pain, as she reported this to be a more effective stimulus. Altogether, 17 of intensity 2, 12 of intensity 3, 9 of intensity 4, 7 of intensity 5 and 5 of intensity 6 were presented in the experiment randomly with fixed proportion in different intensity level. Pain levels were achieved by matching perceived pain levels to those used in Gallo et al. (2018) by author KI. Furthermore, there was significant positive relation between shock intensity in 1st video and donation indicating the stimuli worked well (r = 0.672, p < 0.001). For 1st video, the number of different shock intensity condition were similar, and reduced high shock intensity video in 2nd video caused by participants' donation, see **Figure S1**.
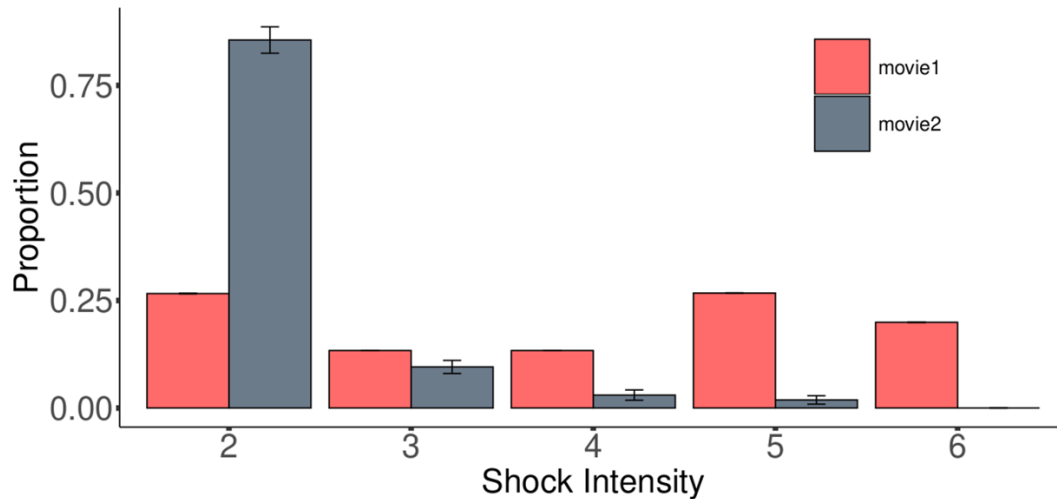
Figure S1. The proportion of shock intensity for 1st video and 2nd video. Error bar represents *SE*.

MOR-dependent responses to pain

Instead of using principal component analysis, we directly cumulate all the significate ROIs after corrected (**Fig. S2**).
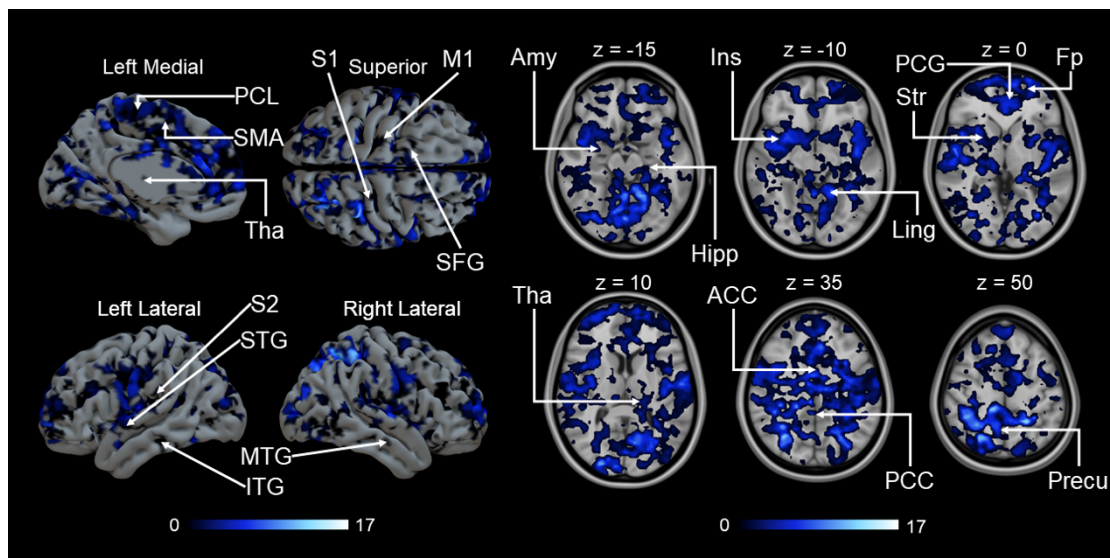

Figure S2. Cumulative maps showing the number of ROIs (out of 17) whose [^{11}C]carfentanil BP_{ND} was negatively correlated (p<0.05, FWE corrected at cluster level) with BOLD response to donation size. PCL = paracentral lobule, SMA = supplementary motor area, Tha = thalamus, S1 = primary somatosensory cortex, M1 = primary motor cortex, SFG = superior frontal gyrus, S2 = secondary somatosensory cortex, STG = superior temporal gyrus, MTG = middle temporal gyrus, ITG = inferior temporal gyrus, Amy = amygdala, Hipp = hippocampus, Ins = insula, Ling = lingual gyrus, PCG = paracingulate gyrus, Str = striatum, Fp = frontal pole, ACC = anterior cingulate cortex, PCC = posterior cingulate cortex, Precu = precuneous cortex.

Vicarious pain responses to 2nd video

After donation, BOLD response to 2nd video was different from 1st video. Anterior insula, and angular gyrus were activated. The anterior insula in particular has been associated with vicarious pain activated in previous studies (**Fig. S3**).
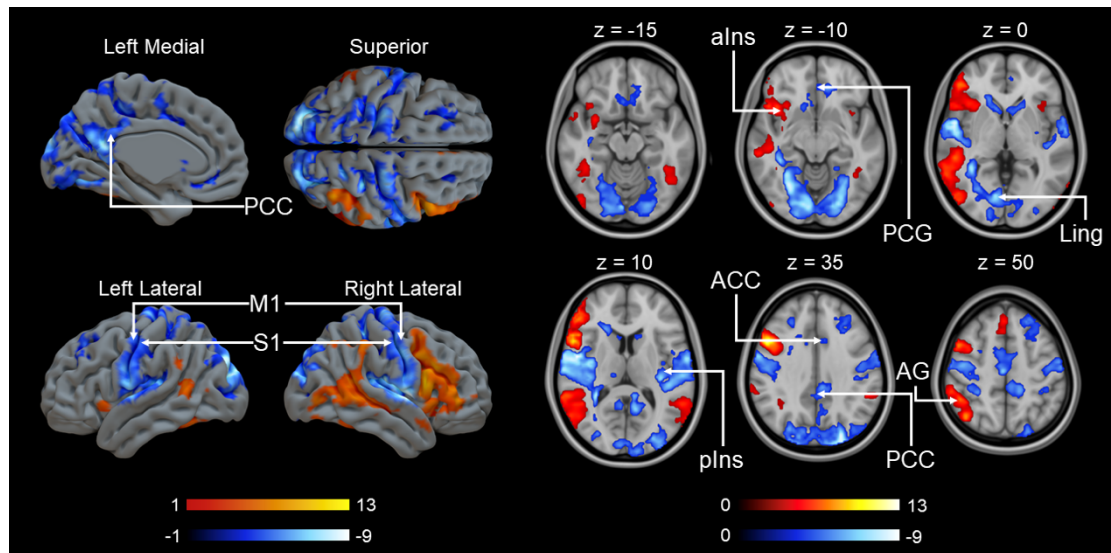

Figure S3. Main effect of pain perception in 2nd video. Colourbars indicate t statistic range. The data are thresholded at p<0.001, FWE corrected at the cluster level (Positive: p_{unc}<0.001, k=FWEc=110 voxels, 3.40<t<12.89; Negative: p_{unc}<0.001, k=FWEc=121 voxels, 3.40<t<10.62). PCC = posterior cingulate cortex, M1 = primary motor cortex, S1 = primary somatosensory cortex, aIns = anterior insula, PCG = paracingulate gyrus, Ling = lingual gyrus, pIns = posterior insula, ACC = anterior cingulate cortex, AG = angular gyrus.

Brain responses in donation phase

In donation phase, as shown in **figure S4**. Extensive brain areas were activated, including anterior insula, striatum, thalamus, anterior cingulate cortex, M1, S1.
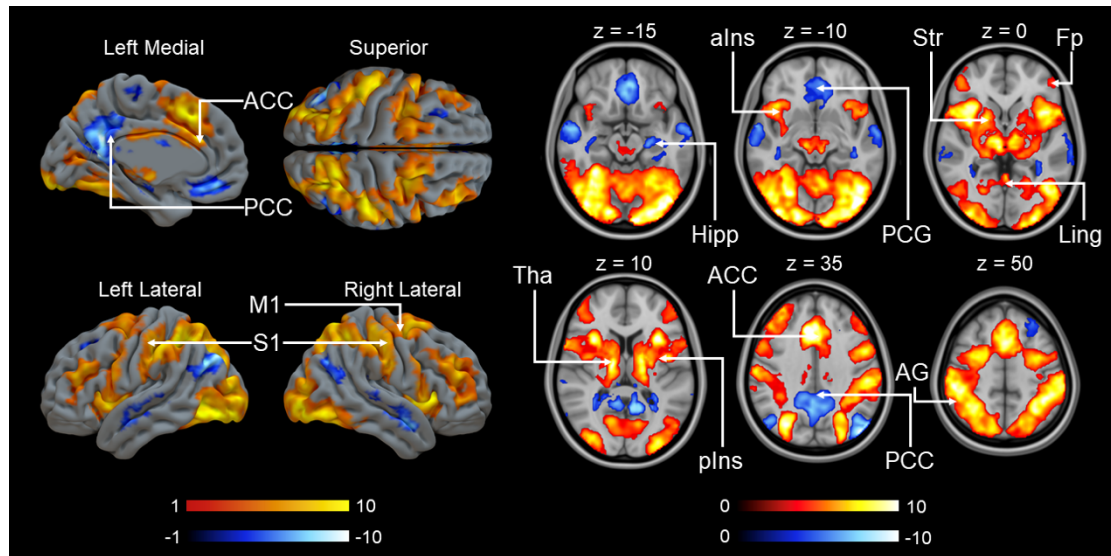
Figure S4. Main effect of brain activation to donation phase. Colourbars indicate t statistic range. The data are thresholded at p<0.001, FWE corrected at the cluster level (Positive: p_{unc}<0.001, k=FWEc=67990 voxels, 3.40<t<14.73; Negative: p_{unc}<0.001, k=FWEc=221 voxels, 3.40<t<11.12). ACC = anterior cingulate cortex, PCC = posterior cingulate cortex, M1 = primary motor cortex, S1 = primary somatosensory cortex, Hipp = hippocampus, aIns = anterior insula, PCG = paracingulate gyrus, Str = striatum, Fp = frontal pole, Ling = lingual gyrus, Tha = thalamus, pIns = posterior, AG = angular gyrus.

Principal component analysis results

PCA result shows that the first three component can explain 90.1% of the variance, from the first component to the third component, they respectively explained 60.74%, 22.13%, and 7.27% of the variance. See the loadings of each ROIs below.

ROI	PC1	PC2	PC3
amy	0.317	0.632	-0.207
cau	0.149	0.121	-0.386
cer	0.231	0.254	-0.069
dacc	0.312	0.004	-0.024
inftemp	0.219	-0.010	0.152
ins	0.214	0.212	-0.091
midtemp	0.200	-0.042	0.091
nacc	0.547	0.040	0.607
ofc	0.248	-0.385	0.184
parsop	0.154	-0.218	-0.084
pcc	0.169	-0.224	-0.091
put	0.110	-0.248	-0.276
racc	0.230	-0.248	-0.178

supfront	0.120	-0.171	-0.210
suptemp	0.121	-0.188	-0.131
tempol	0.168	-0.197	-0.224
tha	0.239	-0.030	-0.355

Table S1. Loadings of the first three principal component on the ROIs. Amy = amygdala, cau = caudate, cer = cerebellum, dacc = dorsal anterior cingulate cortex, inftemp = inferior temporal gyrus, ins = insula, midtempo = middle temporal gyrus, nacc = nucleus accumbens, ofc = orbitofrontal cortex, parsop = pars opercularis, pcc = posterior cingulate cortex, put = putamen, racc = rostral anterior cingulate cortex, supfront = superior frontal gyrus, suptemp = superior temporal gyrus, tempol = temporal pole, tha = thalamus.
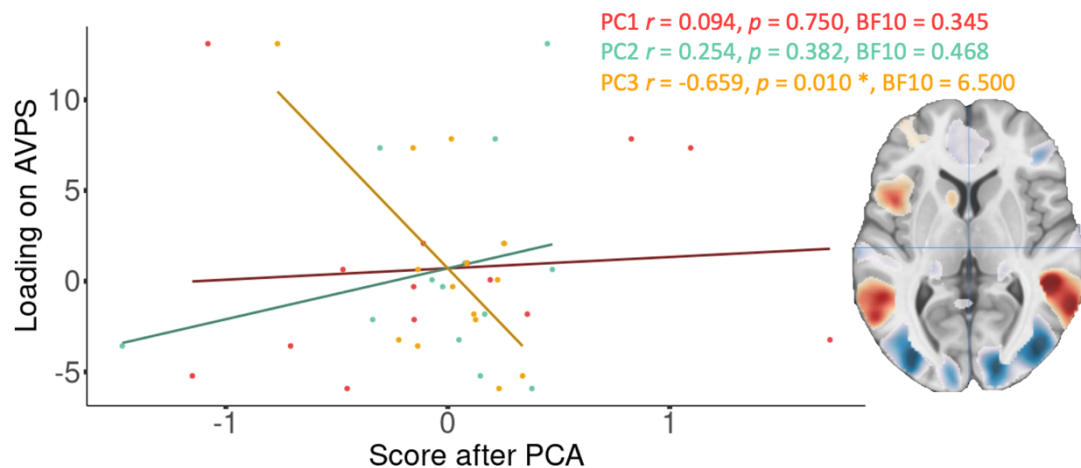

Figure S5. Correlation between the loading on AVPS of the parametric modulator donation size during 1st video and first three principal components. (Red=PC1, green=PC2, yellow=PC3). The sagittal slice on the right represents a thresholded version of the AVPS signature (https://identifiers.org/neurovault.collection:6332) that was dot-multiplied with each participants parametric modulator volume. Each dot represents the loading of one participant onto the AVPS.