



Expressive suppression to pain in others reduces negative emotion but not vicarious pain in the observer

Steven R. Anderson¹ · Wenxin Li² · Shihui Han² · Elizabeth A. Reynolds Losin¹

Accepted: 31 January 2021 / Published online: 24 March 2021
© The Psychonomic Society, Inc. 2021

Abstract

Although there are situations where it may be appropriate to reduce one's emotional response to the pain of others, the impact of an observer's emotional expressivity on their response to pain in others is still not well understood. In the present study, we examined how the emotion regulation strategy *expressive suppression* influences responses to pain in others. Based on prior research findings on expressive suppression and pain empathy, we hypothesized that expressive suppression to pain expression faces would reduce neural representations of negative emotion, vicarious pain, or both. To test this, we applied two multivariate pattern analysis (MVPA)-derived neural signatures to our data, the Picture Induced Negative Emotion Signature (PINES; Chang, Gianaros, Manuck, Krishnan, and Wager (2015)) and a neural signature of facial expression induced vicarious pain (Zhou et al., 2020). In a sample of 60 healthy individuals, we found that viewing pain expression faces increased neural representations of negative emotion and vicarious pain. However, expressive suppression to pain faces reduced neural representations of negative emotion only. Providing support for a connection between neural representations of negative emotion and pain empathy, PINES responses to pain faces were associated with participants' trait-level empathy and the perceived unpleasantness of pain faces. Findings suggest that a consequence of suppressing one's facial expressions in response to the pain of others may be a reduction in the affective aspect of empathy but not the experience of vicarious pain itself.

Keywords Emotion regulation · Expressive suppression · Empathy · Pain expression · Multivariate pattern expression

Introduction

There are many contexts in which it may be appropriate to suppress one's emotional expressivity to the pain of others. For example, parents may attempt to decrease their emotional responses to pain in their children in order to not upset them (Caes et al., 2014; Vervoort et al., 2011). Physicians are trained to regulate their emotional responses to the pain of their patients, consistent with a view of empathy as a form of “emotional labor” in clinical contexts that requires energy, resources, and support to maintain (Larson & Yao, 2005). The regulation

of empathic responses to the pain of patients may function to not exhaust physicians' own cognitive and emotional resources (Cheng et al., 2007; Decety, Yang, & Cheng, 2010). Despite the utility of understanding how the suppression of emotional expressivity influences responses to pain in others, the relationship between an observer's emotional expressivity and their responses to pain in others is still not well understood. Understanding the effect of the reduction of emotional expressivity on responses to others' pain has important implications. For example, if physicians' suppression of emotional expressivity to pain in others decreases the degree to which they feel the pain of their patients, this could have consequences for the quality of the physician-patient relationship and may even influence subsequent treatment decisions (Weilenmann et al., 2018).

The suppression of emotional expressivity has been primarily studied in experimental settings using the emotion regulation strategy *expressive suppression* (Gross & Levenson, 1997). This emotion regulation strategy is achieved by instructing participants to voluntarily inhibit their facial expressions in response to negative affective stimuli (Gross & John, 2003). The use of expressive suppression has been

✉ Steven R. Anderson
steven.anderson@miami.edu

✉ Elizabeth A. Reynolds Losin
e.losin@miami.edu

¹ Department of Psychology, University of Miami, 5665 Ponce de Leon Boulevard, Coral Gables, FL 33146-0751, USA

² School of Psychological and Cognitive Sciences, PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, 100080, China

associated with a number of negative social, cognitive, and affective outcomes, including an increase in negative affect and depressive symptomatology (Moore & Zoellner, 2012) and lower social support and social satisfaction (Cutuli, 2014; Goldin, McRae, Ramel, & Gross, 2008; John & Gross, 2004; Sheldon, Ryan, Rawsthorne, & Ilardi, 1997; Srivastava, Tamir, McGonigal, John, & Gross, 2009). In experimental studies, expressive suppression is associated with increased stress responses (Butler et al., 2003), worse social communication (Butler, Lee, & Gross, 2007), worse executive functioning (Niermeyer, Ziemnik, Franchow, Barron, & Suchy, 2019), and impaired memory (Moore & Zoellner, 2012; Richards & Gross, 2000).

An additional negative consequence of expressive suppression may be a reduction in empathy when viewing pain in others. Two nonmutually exclusive hypotheses of the effect of expressive suppression on pain empathy are suggested by the literature. First, expressive suppression may decrease the affective aspect of pain empathy. Observing others in pain is known to elicit negative emotion in the observer (Avenanti, Minio-Paluello, Sforza, & Aglioti, 2009; Franck, Cox, Allen, & Winter, 2004; Hadjistavropoulos et al., 2011; Yamada & Decety, 2009), and this experience of negative emotion may be one way that empathy arises (Lamm & Majdandžić, 2015; Simon, Craig, Gosselin, Belin, & Rainville, 2008). The possibility that engaging in expressive suppression reduces the affective aspect of empathy is supported by studies demonstrating that inhibiting facial expressivity impairs negative emotion perception (Schneider, Hempel, & Lynch, 2013) and the associated neural correlates of negative emotion (Dörfel et al., 2014; Ohira et al., 2006; Vrticka et al., 2013). Although several studies have found that expressive suppression reduces brain activity related to negative emotion, for example in the amygdala (Dörfel et al., 2014; Ohira et al., 2006) and anterior insula (aINS) (Vrticka et al., 2013), others have reported an increase in brain activity related to negative emotion during expressive suppression (Goldin et al., 2008; Hayes et al., 2010; Vanderhasselt, Baeken, Schuerbeek, Luybaert, & Raedt, 2013).

Expressive suppression to pain in others also may decrease the degree to which an observer shares the pain experience (i.e., experiences vicarious pain). Viewing someone in pain is associated with behavioral changes consistent with nociceptive pain (Loggia, Mogil, & Bushnell, 2008). Individuals with congenital insensitivity to nociceptive pain also have been found to have deficits in empathic responding, namely a reduction in aversive emotional responses to pain in others (Danziger, Prkachin, & Willer, 2006). Neuroimaging studies have found that viewing others in pain increases activity in brain regions involved in nociceptive pain, including the aINS and dorsal anterior cingulate cortex (dACC) (Corradi-Dell'Acqua, Hofstetter, & Vuilleumier, 2011; Lamm, Nusbaum, Meltzoff, & Decety, 2007). These “shared neural representations” between observed and experienced pain are

theorized to underlie empathic responses (Decety & Meyer, 2008; Singer et al., 2004), although it is likely that there are multiple neural mechanisms for eliciting empathy (Lamm & Majdandžić, 2015). Suggesting a potential link between the inhibition of facial expressivity and the downregulation of vicarious pain, Han, Luo, and Han (2016) reported that the physical inhibition of facial expressivity (via a pencil held in the mouth) reduced event-related potentials (ERPs) associated with empathic neural responses.

Functional magnetic resonance imaging (fMRI) can help to differentiate between the potential effects of expressive suppression on responses to pain in others. Specifically, the use of multivariate fMRI analysis techniques that treat the brain as the predictor, such as multivariate pattern analysis (MVPA), may help to delineate the specific neural processes involved in viewing pain in others (Zaki, Wager, Singer, Keyser, & Gazzola, 2016). Recent studies have used MVPA approaches to develop brain signatures of processes implicated in pain empathy: negative emotion and vicarious pain. These signatures are the Picture Induced Negative Emotion Signature (PINES) (Chang, Gianaros, Manuck, Krishnan, & Wager, 2015) and a neural signature of facial expression induced vicarious pain (Zhou et al., 2020; referred herein as the FEIVPS), which were developed using MVPA techniques (Wager, Atlas, Leotti, & Rilling, 2011) to predict behavioral responses from multivariate patterns of brain activity. This MVPA approach has the advantage of taking into account activity that is distributed across cortical and subcortical regions throughout the entire brain and thus may offer insights into emerging patterns of activity (Kragel & LaBar, 2016). Accordingly, this MVPA approach has been found to outperform functional connectivity-derived network maps and region-of-interest (ROI) analysis approaches (Chang et al., 2015). Furthermore, comparing two predictive brain patterns may identify useful similarities and differences between the underlying neural representations (Chikazoe, Lee, Kriegeskorte, & Anderson, 2014; López-Solà, Koban, Krishnan, & Wager, 2017a). Comparing expression of the PINES and FEIVPS patterns can provide evidence of which specific aspects of pain empathy are influenced by expressive suppression—negative emotion, vicarious pain, or both.

The primary goal of the research project from which this manuscript was developed was to examine the relationship between expressive suppression and responses to pain in others. A secondary goal was to investigate whether there are cross-national differences in this relationship, as prior literature suggests that East Asian individuals may experience fewer negative consequences of expressive suppression compared to Western individuals (Butler et al., 2007; Butler, Lee, & Gross, 2009; Soto, Perez, Kim, Lee, & Minnick, 2011), possibly due to cultural variability in the value placed on emotional suppression (Markus & Kitayama, 1991; Soto, Levenson, & Ebling, 2005). As a result, for this project we collected a sample of individuals in China and a matching sample in the United States. All

participants underwent fMRI while passively viewing and engaging in expressive suppression to pain and neutral expression faces (Han et al., 2016; Feng Sheng & Han, 2012; Sheng, Liu, Li, Fang, & Han, 2014) and negative and neutral pictures from the International Affective Picture System (IAPS) (Lang, Bradley, & Cuthbert, 1995).

To examine the influence of expressive suppression on responses to pain in others, we tested whether expressive suppression reduced expression of the PINES (negative affect), the FEIVPS (vicarious pain), or both. To test whether there were cross-national differences in the effect of expressive suppression on negative emotion or vicarious pain, we tested for interactions between the expressive suppression manipulation and nationality (U.S.-China). For completeness and to compare to prior pain empathy and expressive suppression literatures, we also conducted a standard whole-brain univariate general linear model (GLM) analysis in which we looked for brain regions that were more active when viewing pain faces or emotionally aversive pictures compared with neutral faces or pictures, and those that exhibited a change in activation during expressive suppression.

For our primary research goal, we predicted that if expressive suppression decreases negative emotional responses to others' pain, then the PINES should decrease when engaging in expressive suppression to pain faces. Similarly, we predicted that if expressive suppression decreases the vicarious experience of others' pain, then the FEIVPS should decrease when engaging in expressive suppression to pain faces. As a test of the specificity of our effects to pain facial expressions, we predicted that if the reduction of PINES or FEIVPS responses during expressive suppression is specific to pain facial expressions, then we should not see a reduction during expressive suppression to IAPS pictures. For our secondary research goal, we predicted that U.S. compared with Chinese participants would demonstrate greater negative consequences of expressive suppression in the form of a larger reduction in PINES or FEIVPS responses during expressive suppression to pain in others. In individual difference measures analyses, we predicted that if PINES or FEIVPS responses were associated with pain empathy then trait-level empathy measures would be positively associated with the unpleasantness or pain intensity of the pain faces. We also predicted that self-reported habitual use of expressive suppression and trait-level empathy would be negatively associated in our sample, consistent with our hypothesis that a reduction in pain empathy may be an additional negative consequence of expressive suppression.

Methods

Participants

Participants were a total of 60 individuals (50% female, 96.7% right-handed) between the ages of 18–29 years ($M = 21.2$, $SD = 3.31$). Because one of the goals of this research project was

to examine cross-national differences in the effect of expressive suppression on pain empathy, participants were recruited and scanned in Peking University in Beijing, China and the University of Miami in Miami, United States. Demographic characteristics for the samples are available in Table 1. All Chinese participants were born in China and identified their race/ethnicity as Han Chinese. All U.S. participants were born in the United States and identified their race/ethnicity as non-Hispanic white/Caucasian. Participants were excluded from participating in the study if they had a current or past psychiatric diagnosis, claustrophobia, or the presence of illness on the day of the scan. All participants had normal vision or vision that was normal after correction. During scanning at the U.S. site, one participant's T1 structural image revealed an incidental finding (brain abnormality). This participant was immediately notified of the incidental finding and their data was replaced in the final sample. The study protocol was approved by the ethics committees at Peking University and the University of Miami. Written informed consent was obtained according to the Declaration of Helsinki (1991; p. 1194) for all research participants.

Training

Before scanning, participants were trained on the fMRI tasks in two parts, closely following training procedures used in previous emotion regulation studies (Dörfel et al., 2014; Ochsner & Gross, 2005). In the first part, participants were given a brief description of how to respond to the images they would see in the scanner, and then given the opportunity to view a series of example pictures while the experimenter instructed them on how to respond. Participants were instructed to look at each picture carefully and try not to distract themselves with irrelevant or positive thoughts. Before viewing each picture, participants viewed an instruction cue telling them to either "Look" or "Suppress" to the picture that followed. For the "Look" cue, participants were instructed to *look at the picture directly and respond naturally*. For the "Suppress" cue, participants were instructed to *keep your face still while looking at the picture so that someone watching your face would not be able to know what you are feeling*. Participant instructions were taken from previous expressive suppression studies (Dörfel et al., 2014; Goldin et al., 2008).

Emotion regulation cues corresponded to the following task conditions (Figure 1). For the Face Task, the task conditions were passively viewing neutral expression faces (NEU_{Face} , preceded by "Look" cue), passively viewing pain expression faces ($PAIN_{Face}$, preceded by "Look" cue), or engaging in expressive suppression to pain expression faces ($eSUP_{Face}$, preceded by "Suppress" cue). For the IAPS Task, the task conditions were passively viewing neutral IAPS pictures (NEU_{IAPS} , preceded by "Look" cue), passively viewing negative IAPS pictures (NEG_{IAPS} , preceded by "Look" cue),

Table 1 Sample demographic characteristics

	Beijing, China ($n = 30$)	Miami, United States ($n = 30$)	Overall ($N = 60$)
Gender			
Female	15 (50.0%)	15 (50.0%)	30 (50.0%)
Male	15 (50.0%)	15 (50.0%)	30 (50.0%)
Age (yr)			
Mean (<i>SD</i>)	22.3 (3.14)	20.1 (3.17)	21.2 (3.31)
Median [Min, Max]	21.0 [19.0, 29.0]	19.0 [18.0, 29.0]	20.0 [18.0, 29.0]

or engaging in expressive suppression to negative IAPS pictures (eSUP_{IAPS}, preceded by “Suppress” cue).

In the second part of the training, participants completed a practice block for each fMRI task. The pictures used during the practice blocks were not shown to participants during scanning. At the end of the practice blocks, the experimenter confirmed that participants understood how to respond to the emotion regulation cues and answered any questions or concerns. All study text and instruction during training, and throughout each study session, was delivered in English for the U.S. sample and Simplified Mandarin Chinese for the Chinese sample.

Stimuli

Task stimuli were delivered using Presentation software (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA). For the Face Task, participants viewed 108 pain or neutral expression faces that have been used in prior studies of pain perception in

others (Han et al., 2016; Feng Sheng & Han, 2012; Sheng et al., 2014). Given prior evidence of racial in-group biases in empathic neural responses to others’ pain (Xu, Zuo, Wang, & Han, 2009), face stimuli were race and ethnicity-matched to participants. For the IAPS Task, participants viewed 108 negative or neutral pictures taken from the International Affective Picture System (Lang & Bradley, 2007). The IAPS pictures in the task conditions were matched for complexity and content, as well as for the presence of the color red and the presence of human faces, because these have been shown to influence neural and psychological responses (Elliot, Maier, Moller, Friedman, & Meinhardt, 2007; Hill & Barton, 2005). To control for any unanticipated aspects of the negative IAPS pictures that might influence emotion regulation, the IAPS pictures associated with the NEG_{IAPS} and eSUP_{IAPS} conditions were counterbalanced across participants.

Mean valence and arousal values on a scale from 1 = most unpleasant/calm to 9 = most pleasant/aroused for the IAPS pictures in the task conditions were calculated to be similar

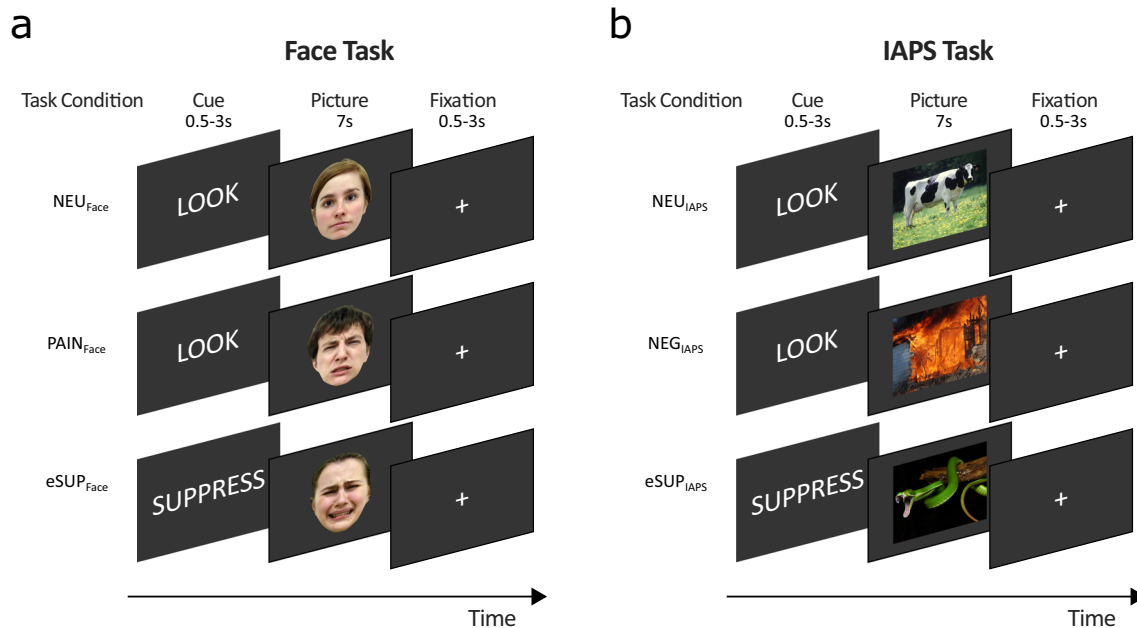


Fig. 1 Design of fMRI tasks. The order that the two tasks were presented was counterbalanced such that half of the participants completed the Face Task (a) first and half completed the IAPS Task (b) first. For the Face Task, participants viewed 108 pain or neutral expression faces. For the

IAPS Task, participants viewed 108 negative or neutral pictures taken from the International Affective Picture System. Face Task stimuli were race and ethnicity-matched to participants

to those reported in prior expressive suppression studies (Che, Luo, Tong, Fitzgibbon, & Yang, 2015; Dörfel et al., 2014) and were as follows: NEU_{IAPS}: Valence = 5.08, Arousal = 3.69; NEG_{IAPS}: Valence = 2.53, Arousal = 5.82; eSUP_{IAPS}: Valence = 2.53, Arousal = 5.60. Testing for differences in the mean valence and arousal values between the task conditions indicated that the values for the pictures in the NEG_{IAPS} and eSUP_{IAPS} conditions were significantly different from those in the NEU_{IAPS} condition ($p < 0.001$), while the values for the pictures in the NEG_{IAPS} and eSUP_{IAPS} conditions were not significantly different from each other ($p > 0.20$). Participants viewed the stimuli in each task on a rear-projected LCD screen via an angled mirror located on the head-coil. Stimuli were presented in the center of a gray background subtending a visual angle of $4.65^\circ \times 5.80^\circ$ and a viewing distance of 134 cm.

fMRI Tasks

Scanning consisted of six functional runs of approximately 7 minutes each. Participants completed three runs containing pain and neutral expression face stimuli (Face Task; Figure 1a) and three runs containing negative and neutral IAPS pictures (IAPS Task; Figure 1b). The order that the two tasks were presented was counterbalanced such that half of the participants completed the Face Task first and half completed the IAPS Task first. As both males and females were recruited, counterbalancing was performed such that an equal number of males and females were assigned to each task counterbalance order.

Each trial of the fMRI task began with a jittered cue word (0.5–3 seconds) instructing participants to either “Look” or “Suppress” to the upcoming picture. Participants then viewed a 7-second picture presentation followed by a jittered fixation cross (0.5–3 seconds). Each fMRI run began and ended with a 20-second fixation cross for an implicit baseline. The presentation of trials was pseudorandomized so that no more than two trials in the same task condition were displayed in a row. Participants saw an equal number of trials in each task condition.

Procedures

Upon arrival at the study location, participants completed informed consent and completed questionnaires assessing demographic characteristics. Next, participants were trained by the experimenter on the fMRI tasks following procedures used in previous emotion regulation studies (Dörfel et al., 2014; Ochsner & Gross, 2005). Participants then completed three functional runs containing pain and neutral expression face stimuli (Face Task) and three runs containing negative and neutral IAPS pictures (IAPS Task), with the order of the tasks counterbalanced across participants. Immediately after

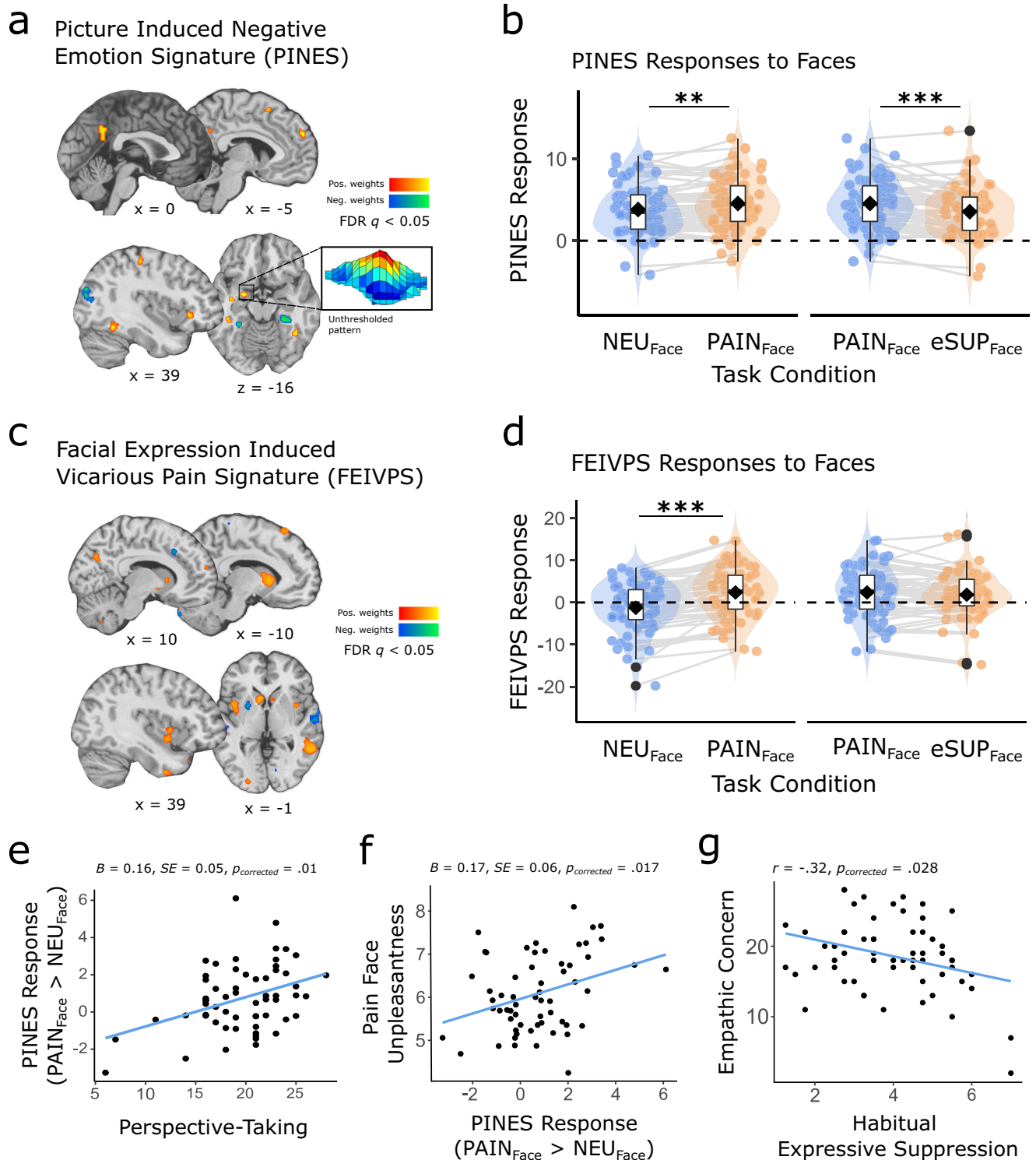
scanning, participants completed ratings of each picture they viewed during the scan, similar to prior emotion regulation studies (Dörfel et al., 2014), rather than on a trial-by-trial basis, due to evidence that conducting ratings in the scanner environment may interrupt amygdala activity (Lieberman, Inagaki, Tabibnia, & Crockett, 2011). Participants were instructed to try to remember how each picture made them feel when they viewed it in the scanner. Participants rated how unpleasant each picture made them feel (1 = Very unpleasant to 9 = Very pleasant; anchors reversed for subsequent analyses so that higher values indicated more unpleasantness) and, for Face Task stimuli only, how much pain intensity they perceived in each face (0 = No pain to 10 = Most intense pain imaginable).

Next, participants completed self-report questionnaires assessing the habitual use of expressive suppression and trait-level empathy, as well as several cultural constructs. Measures included the Emotion Regulation Questionnaire (ERQ), a 10-item measure of the habitual use of emotion regulation, with separate subscales for Expressive Suppression and Cognitive Reappraisal (Gross & John, 2003). Each item is rated on a 7-point scale (1 = Strongly disagree to 7 = Strongly agree). Participants also completed the Interpersonal Reactivity Index (IRI), a 28-item self-report measure of trait-level empathy comprising subscales for Empathic Concern, Perspective-Taking, Fantasy, and Personal Distress (Davis, 1983). Each item is rated on a 5-point scale (1 = Does not describe me well to 5 = Describes me very well).

Several additional measures of cultural constructs were administered. These included a measure to assess the strength of participants’ self-construal, a measure of self- or other-focus, which has been found to differ between individuals in East Asian and Western cultural contexts (Markus & Kitayama, 1991). To measure self-construal, participants completed the 24-item Self-Construal Scale (SCS) (Singelis, 1994). The SCS contains separate subscales for independent and interdependent self-construal, each rated on a 7-point scale. Assessing beliefs about appropriate pain behavior, which have been found to differ between U.S. and East/South Asian individuals (Hobara, 2005; Nayak, Shiflett, Eshun, & Levine, 2000), participants completed the Appropriate Pain Behavior Questionnaire (APBQ), a 28-item measure in which participants rate how much they agree with the appropriateness of various pain behaviors for men and women on a 4-point scale (Nayak et al., 2000).

fMRI acquisition and processing

Scanning at the U.S. site occurred at the Neuroimaging Facility at the University of Miami (Miami, United States) on a GE Discovery MR750 3.0T MR scanner. Scanning at the China site occurred at the IDG McGovern Institute for



Brain Research at Peking University (Beijing, China) on a comparable GE Discovery MR750 3.0T MR scanner, using the same scanner parameters and equivalent 32-channel head coils as the U.S. site. Thirty-two axial slices covering the whole brain volume were acquired, with a total of 216 volumes collected per functional run (TR = 2,000

ms, TE = 30 ms, flip angle = 90°, FOV = 240 mm, 64 × 64 matrix, slice thickness = 4 mm, voxel size = 3.75 × 3.75 × 4 mm³). A T1-weighted anatomical image was acquired from each participant using a 3D BRAVO sequence (TI = 450 ms, sagittal orientation, flip angle = 12°, FOV = 240 mm, 256 × 256 matrix, slice thickness = 1 mm). Data preprocessing and

◀ **Fig. 2** Multivariate pattern expression and individual difference analyses results. **a)** Picture Induced Negative Emotion Signature (PINES), threshold at $FDR q < 0.05$. Image modified and reproduced from Chang et al. (2015); **b)** PINES responses by task condition in the Face Task; **c)** Facial Expression Induced Vicarious Pain Signature (FEIVPS), threshold at $FDR q < 0.05$; **d)** FEIVPS responses by task condition in the Face Task; **e)** Participants' trait-level perspective-taking (IRI-PT) predicting the observed increase in PINES responses when viewing pain versus neutral faces ($PAIN_{Face} > NEU_{Face}$); **f)** Participants' observed increase in PINES responses when viewing pain versus neutral faces ($PAIN_{Face} > NEU_{Face}$) predicting ratings of the unpleasantness of the pain faces conducted immediately after the scan; **g)** Negative correlation between participants' habitual use of expressive suppression and trait-level empathic concern. *P* values Bonferroni corrected for multiple comparisons within family. Combined violin and boxplots depict raw data, while asterisks depict the results of linear mixed effects models. Lines connecting each violin and boxplot indicate within-subject changes in signature responses due to task condition. Abbreviations: IRI – PT = Interpersonal Reactivity Index – Perspective Taking Subscale; IRI – EC = Interpersonal Reactivity Index – Empathic Concern Subscale, ERQ – eSUP = Emotion Regulation Questionnaire – Expressive Suppression Subscale. ** $p < 0.01$; *** $p < 0.001$.

statistical analyses were conducted using FSL version 5.0.9. Five volumes were discarded from each run to ensure that the scanner had reached a steady state for the remaining volumes of interest. The anatomical images were preprocessed with the following steps: reorientation to standard MNI orientation, cropping, registration to standard space using FLIRT (Greve & Fischl, 2009; Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001), and brain-extraction. Preprocessing steps applied to functional data included brain extraction, slice timing correction to the middle slice for interleaved EPI, spatial smoothing using a full-width at half maximum (FWHM) 6-mm Gaussian smoothing kernel, and removal of low frequency drift (Smith et al., 2004). Functional images were co-registered to structural images and transformed into standard MNI space. Temporal autocorrelation was estimated and corrected via prewhitening using FMRIB's Improved Linear Model (Woolrich, Brady, & Smith, 2001).

Motion correction was performed using FSL's MCFLIRT (Jenkinson et al., 2002), which corrects for excessive head motion using rigid-body transformation across six standard motion parameters (rotations and translations along x, y, and z axes). Per recommendations for characterizing head motion in fMRI data (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012), we calculated the mean framewise displacement (FD) or the average of rotation and translation parameter differences (using weighted scaling) for each subject. No subjects in either the Face Task or IAPS Task had a mean FD > 0.50 mm, a cutoff recommended for motion scrubbing in task fMRI data (Siegel et al., 2014). As a result, no subjects were excluded from subsequent analyses due to excessive head motion.

Multivariate pattern expression analysis

To test whether expressive suppression to pain faces decreases neural representations of negative emotion, vicarious pain, or both, we applied two previously developed neural signatures to our data using Matlab version 8.6.0.267246 (R2015b).

Picture Induce Negative Emotion Signature (PINES) The PINES (Fig. 2a) was developed in a separate fMRI study (Chang et al., 2015) in which participants were scanned while viewing a series of negative and neutral pictures from the International Affective Picture System (IAPS) (Lang & Bradley, 2007). Chang et al. (2015) used a machine learning algorithm, Least Absolute Shrinkage and Selection Operator and Principle Components Regression (LASSO-PCR) (Wager et al., 2011; Wager et al., 2013) to develop the PINES. LASSO-PCR acts to first reduce the large amount of brain data per participant using principle components analysis, with each component of the principle components analysis representing a set of voxels that share a similar pattern of activation. These components are then used in a least squares regression with an L-1 regularization (i.e., a penalty on the parameters which increases predictive performance (Tibshirani, 1996)). This MVPA approach enabled the authors to identify a neural signature of negative emotion—the PINES—which consists of a map of positive and negative weights for each voxel in the brain that is predictive of negative emotion ratings.

Importantly, the PINES pattern contains non-zero or near-zero predictive weights in cortical and subcortical voxels across the entire brain. Brain regions with positive predictive weights (i.e., increased activity predicts more negative emotion) in a thresholded ($FDR q < 0.05$) PINES predictive map include the amygdala, anterior insula (aINS), dorsomedial prefrontal cortex (dmPFC), presupplementary motor area (pre-SMA), and posterior cingulate cortex (PCC), all regions that have been previously implicated in negative emotion (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). The PINES pattern also contains negative predictive weights (i.e., increased activity predicts less negative emotion) in the parahippocampal gyrus, superior temporal gyrus, temporoparietal junction (TPJ), and caudate, among other regions. The PINES has been demonstrated to be sensitive and specific to experiences of negative emotion in new subjects (Chang et al., 2015; Shahane, Lopez, & Denny, 2019).

Facial Expression Induced Vicarious Pain Signature (FEIVPS)

The FEIVPS (Fig. 2c) was developed in a previous fMRI study (Zhou et al., 2020) in which participants were scanned while viewing painful or nonpainful noxious stimulation of body limbs (Meng et al., 2012), pain or neutral facial expressions (the same face stimuli used in the present study) (Feng

Sheng & Han, 2012), negative or neutral IAPS pictures, and nonpainful control images. The FEIVPS is similar to the previously developed Vicarious Pain Signature (VPS) (Krishnan et al., 2016), which was trained to predict participants' ratings of vicarious pain while they viewed painful limb stimuli (see [Supplemental Materials](#) for multivariate pattern expression results for the VPS, which did not respond to the pain expression face stimuli in our study). Preprocessed (motion corrected and high-pass filtered) fMRI data from the tasks was then subjected to linear support vector machines (SVMs) to train separate multivariate pattern classifiers of vicarious pain induced by noxious stimulation and pain facial expressions. SVM classification was evaluated using a 10-fold cross-validation procedure (Hsu, Chang, & Lin, 2003). The resulting FEIVPS pattern has non-zero or near-zero predictive weights in voxels throughout the entire brain. Positive predictive weights (i.e., increased activity predicts more vicarious pain) in a thresholded (FDR $q < 0.05$) FEIVPS pattern are located in the precuneus, PCC, medial frontal gyrus, caudate, superior frontal gyrus, middle temporal gyrus, and insula. Negative predictive weights (i.e., increased activity predicts less vicarious pain) in a thresholded FEIVPS pattern are located in the midcingulate cortex (MCC), ventromedial prefrontal cortex (vmPFC), putamen, and superior temporal gyrus.

In order to calculate the pattern expression in our own data, we conducted a first-level general linear model (GLM) analysis in FSL with the onsets and durations of the following events modeled as regressors: the instruction cue for each task condition, the stimulus presentation period for each task condition, and the fixation periods between trials. Six standard motion parameters were included in the GLM as nuisance regressors. For the second-level analysis, parameter estimates from the individual contrast images in the first-level analysis were combined across functional runs using a fixed-effects analysis. Then, using Matlab software, we calculated the expression of each pattern in each of the task conditions versus an implicit baseline for each subject. As detailed in prior studies (López-Solà, Geuter, Koban, Coan, & Wager, 2019; López-Solà, Woo, et al., 2017b; Wager et al., 2013), the pattern expression was calculated by taking the dot product of the unthresholded, vectorized activation contrast z statistic image with the predictive weight map.

Calculating the expression of each pattern in new data, as we did in the present study, yields a single numeric value, the *signature response*. The signature response represents the degree to which the new subject's brain data resembles the predictive pattern, with higher response values indicating greater similarity between the new subject's brain data and the predictive pattern. Because the neural signatures were developed specifically to predict participant ratings, we can interpret a given signature response such that a higher signature response value indicates higher amounts of a given neural representation.

Statistical analysis

The resulting vectors of signature response values calculated in Matlab were then imported into R Version 3.3.2 (R Core Team, 2019). In R, we used linear mixed effects models (Magezi, 2015) to estimate the effect of our task conditions (passively viewing faces/images and expressive suppression) on participants' signature response values (one signature response value for each task condition and for each subject). Our use of linear mixed effects models allowed us to specify a random intercept for each participant, which accounts for individual differences in baseline brain activity to the stimuli in each task (Costafreda, 2009; Mumford & Nichols, 2006). Models were specified using the lme4 package in R (Bates, M, Bolker, & Walker, 2015). Significance values for fixed effects were calculated using Satterthwaite approximation (lmerTest R package) of degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2015). Separate models were specified for each neural signature (PINES and FEIVPS) and fMRI task (Face and IAPS).

To investigate whether expressive suppression to pain faces influenced neural representations of negative emotion, vicarious pain, or both, we specified linear mixed effects models with task condition (passively viewing faces/images and expressive suppression) predicting signature responses. Participant ID was included as random effect with a random intercept. Participant age and gender were included as fixed effects in each model for statistical control due to evidence of gender and age-related variability in pain empathy (Christov-Moore et al., 2014; Grünh, Rebusal, Diehl, Lumley, & Labouvie-Vief, 2008). Nationality was included as a fixed effect of no interest in these models to control for potential cross-national differences in signature responses. To address our secondary research goal, which was whether there were cross-national differences in the effect of expressive suppression, we specified separate linear mixed effects models with a task condition-by-nationality interaction, which allowed us to assess cross-national differences in the magnitude of signature response changes due to passively viewing faces/images and expressive suppression. Gender and age were included as covariates, and participant was included as a random effect with a random intercept in each model.

To investigate whether changes in signature responses due to viewing pain faces were related to pain empathy, we first tested in linear regression models whether trait-level empathy was associated with signature responses when viewing pain versus neutral faces. In separate linear regression models, we next tested whether signature responses when viewing pain versus neutral faces were associated with stimulus ratings of the unpleasantness or pain intensity of pain faces. Finally, we examined whether the habitual use of expressive suppression and trait-level empathy were negatively related in our sample, consistent with our hypothesis of a negative relationship

between expressive suppression and pain empathy. We tested this hypothesis using a Pearson r correlation analysis conducted between the ERQ – Expressive Suppression subscale and the IRI subscales found to be significantly related to signature responses. To correct for multiple comparisons, we subjected p values to Bonferroni correction within each family of models.

To interpret key null effects estimated using null-hypothesis significance testing, we estimated Bayes Factors which quantify the likelihood of the data under the null and alternative hypotheses (Morey & Rouder, 2011). To estimate Bayes Factors for one-sample t -tests, we first computed difference scores between our task conditions (passively viewing faces/images and expressive suppression). We then utilized the *ttestBF* function in the BayesFactor R package (v0.9.12), which utilizes the Jeffreys–Zellner–Siow (JZS) prior (scale factor = 0.707). This prior combines the Cauchy distribution on the standardized effect size and a noninformative Jeffreys prior on the variance of the normal population (Rouder, Speckman, Sun, Morey, & Iverson, 2009).

Univariate general linear model (GLM) analysis

We conducted an exploratory univariate general linear model (GLM) analysis as an additional source of information to that provided by the multivariate pattern expression analysis. The within-subject fixed-effects analysis that was used as the input for the pattern expression analysis was also used in the GLM analysis. In the GLM analysis, the within-subject analysis was followed by a third-level mixed-effects analysis which combined data across participants using FMRIB Local Analysis of Mixed Effects (FLAME1; Beckmann et al., 2003; Woolrich et al., 2004). We calculated a group mean across all participants and included participant nationality as an additional covariate for statistical control, orthogonalized with respect to the group mean. Resulting z statistic images were false discovery rate (FDR) corrected using a critical FDR threshold of $q < 0.05$ and $k > 50$, which controls the expected proportion of false positives among suprathreshold voxels (Benjamini & Hochberg, 1995; Genovese, Lazar, & Nichols, 2002). The FDR threshold is determined from the observed p value distribution and is therefore adaptive to the amount of signal in the data. The q value is the FDR corrected p value for that image. Thus, images thresholded at FDR $q < 0.05$ are significant at FDR corrected $p < 0.05$. This would mean that, on average, 5% of the observed results would be false positives. Anatomical brain regions were identified using the Harvard-Oxford Cortical and Subcortical Structural Atlases in FSLView.

The contrast $\text{PAIN}_{\text{Face}} > \text{NEU}_{\text{Face}}$ was calculated to assess increases in activation when passively viewing pain versus neutral expression faces in the Face Task. The contrast $\text{eSUP}_{\text{Face}} > \text{PAIN}_{\text{Face}}$ was calculated to assess increases in

activation related to engaging in expressive suppression to pain faces versus passively viewing pain faces in the Face Task. The contrast $\text{NEG}_{\text{IAPS}} > \text{NEU}_{\text{IAPS}}$ was calculated to assess increases in activation when passively viewing negative versus neutral pictures in the IAPS Task. The contrast $\text{eSUP}_{\text{IAPS}} > \text{NEG}_{\text{IAPS}}$ was calculated to assess increases in activation related to engaging in expressive suppression to negative pictures versus passively viewing negative pictures in the IAPS Task.

Results

Stimulus rating analysis results

Pain faces rated as more unpleasant and higher in pain intensity than neutral faces As a manipulation check of our fMRI tasks, we first examined participant stimulus ratings to the face and IAPS stimuli made immediately after scanning. As expected, participants rated pain faces in the Face Task as significantly more unpleasant ($M = 6.09$, $SD = 0.88$) than neutral faces ($M = 4.85$, $SD = 0.78$), $t(59) = 8.12$, $p < 0.001$. Participants also rated pain faces as higher in pain intensity ($M = 5.45$, $SD = 1.58$) than neutral faces ($M = 0.60$, $SD = 0.77$), $t(59) = 23.66$, $p < 0.001$. For the IAPS Task stimuli, as expected, participants rated negative pictures as significantly more unpleasant ($M = 7.08$, $SD = 0.81$) than neutral pictures ($M = 4.56$, $SD = 0.68$), $t(58) = 21.07$, $p < 0.001$ (one subject's data was lost due to software error). These results indicate that participants perceived the pain faces and negative IAPS pictures as unpleasant, the pain faces as reflecting pain, and were able to differentiate between the stimuli in the task conditions, suggesting that our fMRI tasks functioned as intended.

Multivariate pattern expression analysis results

Viewing pain faces increases neural representations of negative emotion and vicarious pain As a manipulation check of the multivariate pattern expression analysis, we first confirmed the effect of viewing pain faces on signature responses separate from engaging in expressive suppression. In linear mixed effects models predicting PINES responses elicited by passively viewing pain and neutral expression faces, we found that PINES responses significantly increased when passively viewing pain faces ($\text{PAIN}_{\text{Face}}$) versus passively viewing neutral faces (NEU_{Face}), $B = 0.73$, $SE = 0.26$, $t(118) = 2.85$, $p = 0.005$, $d = 0.52$ (Figure 2b, left). Similarly, FEIVPS responses significantly increased when passively viewing pain versus neutral faces, $B = 3.66$, $SE = 0.47$, $t(118) = 7.78$, $p < 0.001$, $d = 1.43$ (Figure 2d, left). These findings suggest that neural representations of negative emotion and vicarious pain were increased by viewing pain facial expressions.

Expressive suppression to pain faces reduces neural representations of negative emotion but not vicarious pain We next tested whether expressive suppression to pain faces reduced neural representations associated with negative emotion, vicarious pain, or both. In linear mixed effects models, we found that PINES responses significantly decreased when engaging in expressive suppression to pain faces (eSUP_{Face}) versus passively viewing pain faces (PAIN_{Face}), $B = -1.005$, $SE = 0.26$, $t(118) = -3.90$, $p < 0.001$, $d = -0.72$ (Figure 2b, right; Table 2). In contrast, FEIVPS responses did not significantly decrease when engaging in expressive suppression to pain faces, $B = -0.53$, $SE = 0.47$, $t(118) = -1.14$, $p = 0.258$, $d = -0.21$ (Figure 2d, right). Follow-up Bayes Factor estimation for the null FEIVPS effect revealed moderate evidence for equivalence between expressive suppression and passive viewing in FEIVPS responses (eSUP_{Face} > PAIN_{Face}: BF₁₀ = 0.26). Together, these findings suggest that expressive suppression to pain faces reduced neural representations of negative emotion but not vicarious pain.

Specificity of effect of expressive suppression to pain facial expressions To test whether the observed reduction in neural representations of negative emotion due to expressive suppression is specific to pain facial expressions, we examined the effects of expressive suppression on PINES and FEIVPS responses when viewing IAPS pictures, which are also emotionally aversive but do not consistently contain pain facial expressions. PINES responses significantly increased when passively viewing negative IAPS pictures (NEG_{IAPS}) versus passively viewing neutral IAPS pictures (NEU_{IAPS}), $B = 2.37$, $SE = 0.21$, $t(118) = 11.47$, $p < 0.001$, $d = 2.11$. Although FEIVPS responses also significantly increased when

passively viewing negative IAPS pictures ($B = 2.07$, $SE = 0.46$, $t(118) = 4.54$, $p < 0.001$, $d = 0.84$), mean responses were negative in both conditions. Previous evidence suggests potential nonmeaningful variation of neural signatures in negative-response or near-zero response ranges (Reddan & Wager, 2018; Wager et al., 2013).

We also found that PINES responses significantly decreased when engaging in expressive suppression to negative IAPS pictures (eSUP_{IAPS}) versus passively viewing negative IAPS pictures (NEG_{IAPS}), $B = -0.85$, $SE = 0.21$, $t(118) = -4.12$, $p < .001$, $d = -0.76$. In contrast, FEIVPS responses did not significantly decrease when engaging in expressive suppression versus passively viewing negative IAPS pictures, $B = 0.003$, $SE = 0.46$, $t(118) = 0.007$, $p = 0.994$, $d = 0.001$. Follow-up Bayes Factor estimation for the null FEIVPS effect revealed moderate evidence for equivalence between expressive suppression and passive viewing in FEIVPS responses (eSUP_{IAPS} > NEG_{IAPS}: BF₁₀ = 0.14).

These findings suggest that the observed reduction in neural representations of negative emotion due to expressive suppression is not specific to pain facial expressions. Additionally, these results provide further validation for the constructs these signatures represent, as the PINES was trained on negative and neutral IAPS pictures (Chang et al., 2015) and the FEIVPS was trained on the same pain and neutral facial expression stimuli used in the present study (Feng Sheng & Han, 2012; Zhou et al., 2020).

Effect of expressive suppression on neural representations of negative emotion does not differ by gender or nationality As a follow-up analysis, we tested whether the effect of expressive suppression or passive viewing of faces differed by participant gender given prior evidence of gender differences in empathic responding (Christov-Moore et al., 2014). Although female participants had significantly higher overall FEIVPS responses to pain and neutral face stimuli compared to male participants ($B = 3.01$, $SE = 1.36$, $t(56) = 2.22$, $p = 0.030$, $d = 0.59$), the effect of participant gender was nonsignificant in all other tested models ($ps > 0.70$).

Next, addressing our secondary research aim, we tested whether the observed effect of expressive suppression on PINES responses differed by participant nationality. We specified in linear mixed effects models a task condition-by-nationality interaction, which allowed us to assess any cross-national differences in the magnitude of PINES response changes due to expressive suppression. PINES response changes due to task condition did not significantly differ by participant nationality ($ps > 0.20$). However, we did find that U.S. participants demonstrated a larger increase in FEIVPS responses when engaging in expressive suppression versus passively viewing negative IAPS pictures compared with Chinese participants ($B = 2.23$, $SE = 0.89$, $t(116) = 2.49$, $p = 0.014$, d

Table 2 Results of linear mixed effects model predicting PINES responses in the Face Task

Predictor	Estimate	SE	<i>p</i>	<i>d</i>
(Intercept)	9.50	2.70	<0.001	
Task Condition [NEU _{Face}]	-0.73	0.26	0.005	-0.52
Task Condition [eSUP _{Face}]	-1.005	0.26	<0.001	-0.72
Gender [male]	-0.04	0.75	0.960	-0.01
Age	-0.24	0.12	0.051	-0.53
Nationality [U.S.]	0.35	0.78	0.423	0.12
Random effects				
σ^2	1.99			
τ_{00} Subject	7.45			
ICC	0.79			
<i>N</i> Subject	60			
Observations	180			
Marginal R ² / conditional R ²	0.090 / 0.808			

Bolded *p*-values indicate significance at $p < 0.05$ level. Results are presented with the PAIN_{Face} task condition as the reference variable.

= 0.46). This suggests that U.S. participants may have been less successful at reducing vicarious pain during expressive suppression to IAPS pictures than Chinese participants.

Finally, given our finding that expressive suppression reduced neural representations of negative emotion but not vicarious pain, we tested whether PINES and FEIVPS responses were uncorrelated in our data. FEIVPS responses have been previously found to be not sensitive to negative emotion (Zhou et al., 2020). PINES and FEIVPS responses to faces were uncorrelated in each task condition in our data (NEU_{Face} : $r = -0.18$, $p = 0.162$; $PAIN_{Face}$: $r = 0.08$, $p = 0.537$; $eSUP_{Face}$: $r = -0.09$, $p = 0.496$). Similarly, PINES and FEIVPS responses to IAPS pictures were uncorrelated in each task condition (NEU_{IAPS} : $r = -0.09$, $p = 0.512$; NEG_{IAPS} : $r = -0.11$, $p = 0.401$; $eSUP_{IAPS}$: $r = -0.14$, $p = 0.276$). As a result, we can assume that the PINES and FEIVPS are likely measuring separable neural representations in our data.

Taken together, these findings suggest that our observed effect of expressive suppression on neural representations of negative emotion did not differ by participant gender or nationality, increasing the generalizability of our findings.

Individual difference analysis results

Neural representations of negative emotion positively associated with trait-level empathy and pain face unpleasantness ratings Because we observed both a significant increase in neural representations of negative emotion (as measured by PINES responses) during passive viewing of pain faces and a significant decrease during expressive suppression, we next investigated the relationship between PINES responses and individual difference measures of empathy and expressive suppression. Testing whether the increase in PINES responses due to viewing pain faces was related to trait-level empathy, we found that trait-level perspective-taking (IRI – PT) was positively associated with the increase in PINES responses due to viewing pain faces, $B = 0.16$, $SE = 0.05$, Bonferroni-corrected $p = 0.01$ (Figure 2e). That is, as participants' trait-level perspective-taking increased, the magnitude of their PINES responses to viewing pain faces increased. We also found a marginal positive association between trait-level empathic concern (IRI – EC) and PINES responses to viewing pain faces, $B = 0.11$, $SE = 0.04$, $p_{corrected} = 0.06$. The relationship between personal distress (IRI – PD) and change in PINES responses was not significant, $B = 0.03$, $SE = 0.05$, corrected $p_{corrected} > 1.00$. Similarly, the relationship between fantasy (IRI – FS) and change in PINES responses was not significant, $B = -0.05$, $SE = 0.04$, $p_{corrected} > 1.00$.

Next, testing whether the increase in PINES responses due to viewing pain faces was associated with the unpleasantness or pain intensity of the pain faces, we found that PINES responses were positively associated with the unpleasantness of pain faces, $B = 0.17$, $SE = 0.06$, Bonferroni-corrected $p = 0.017$ (Figure 2f).

That is, as the magnitude of participants' PINES responses due to viewing pain faces increased, their ratings of the unpleasantness induced by the pain faces increased. The relationship between PINES responses due to viewing pain faces and ratings of the pain intensity of pain faces was not significant, $B = 0.05$, $SE = 0.12$, $p_{corrected} > 1.00$.

Habitual use of expressive suppression negatively associated with trait-level empathy

To examine whether individual difference measures of expressive suppression and empathy were negatively related in our data, we tested whether the habitual use of expressive suppression was negatively associated with the trait-level empathy measures that we identified as related to the increase in PINES responses due to viewing pain faces. Consistent with our hypothesis, we found that the habitual use of expressive suppression (ERQ – eSUP) was negatively associated with trait-level empathic concern (IRI – EC), $r = -0.32$, $t = -2.53$, 95% confidence interval (CI) $[-0.53, -0.07]$, Bonferroni-corrected $p = 0.028$ (Figure 2g). In contrast, the habitual use of expressive suppression was not associated with trait-level perspective-taking (IRI – PT), $r = -0.03$, $t = -0.26$, 95% CI $[-0.29, 0.22]$, $p_{corrected} > 1.00$.

These findings indicate that the significant increase in PINES responses that we observed when participants viewed pain faces was associated with trait-level empathy and ratings of the unpleasantness of the pain faces. Additionally, the habitual use of expressive suppression and trait-level empathy were negatively related, suggesting that individuals who more frequently engaged in expressive suppression had lower trait-level empathy in our data.

Univariate GLM analysis results

Viewing pain faces elicits brain activity in regions previously associated with empathic neural responses We conducted an exploratory univariate GLM analysis as a complement to our multivariate pattern expression analysis and to provide a point of comparison with prior fMRI studies of empathy and expressive suppression. We found that several brain regions were more active when participants viewed pain versus neutral faces ($PAIN_{Face} > NEU_{Face}$), the contrast commonly used to investigate neural correlates of pain empathy in previous studies (Lamm, Decety, & Singer, 2011). Active regions in this contrast included the bilateral anterior insula (aINS), inferior frontal gyrus (IFG), dorsal anterior cingulate cortex (dACC), supplementary motor area (SMA), and precuneus, consistent with regions that have been previously implicated in empathic neural responses to pain (Fan, Duncan, de Greck, & Northoff, 2011; Lamm, Decety, & Singer, 2011) (Figure 3a; Table 3).

Expressive suppression to pain faces associated with activity in prefronto-parietal regions involved in emotion regulation

We next examined which brain regions were associated with expressive suppression as a point of comparison with prior

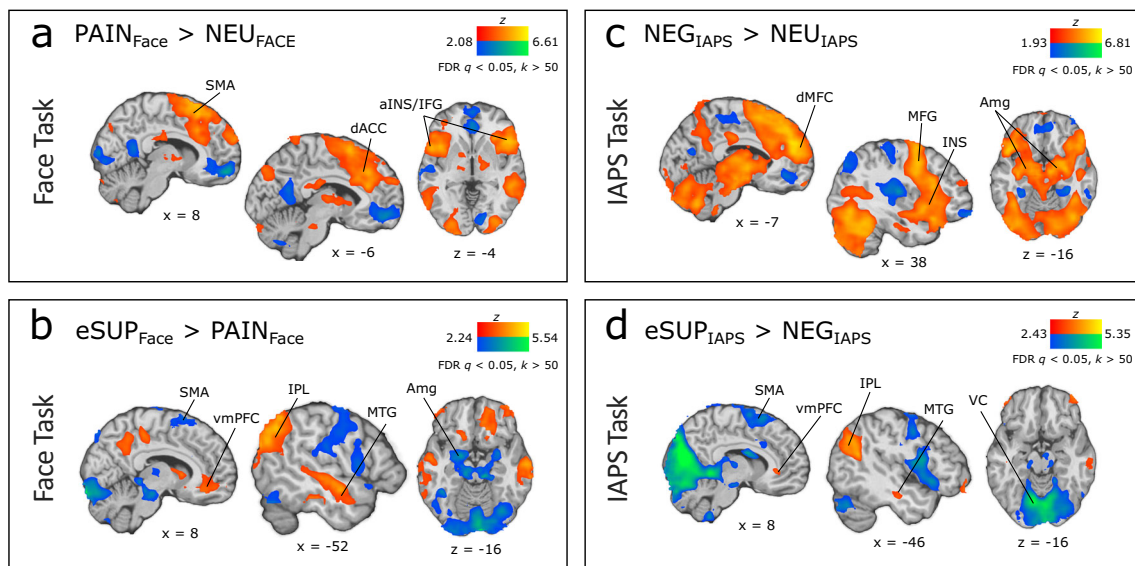


Fig. 3 Univariate general linear model (GLM) analysis results. Numbers displayed below each brain view indicate MNI coordinates. All maps thresholded at $FDR\ q < 0.05, k > 50$. **a)** Whole-brain activity when passively viewing pain versus neutral expression faces ($PAIN_{Face} > NEU_{Face}$); **b)** Whole-brain activity when engaging in expressive suppression to pain faces versus passively viewing pain faces ($eSUP_{Face} > PAIN_{Face}$); **c)** Whole-brain activity when passively viewing negative versus neutral IAPS pictures ($NEG_{IAPS} > NEU_{IAPS}$); **d)** Whole-brain activity

when engaging in expressive suppression to negative IAPS pictures versus passively viewing negative IAPS pictures ($eSUP_{IAPS} > NEG_{IAPS}$). SMA = supplementary motor area; dACC = dorsal anterior cingulate cortex; aINS = anterior insula; IFG = inferior frontal gyrus; dmFC = dorsal medial frontal cortex; MFG = middle frontal gyrus; MTG = middle temporal gyrus; Amg = amygdala; VC = visual cortex; vmPFC = ventromedial prefrontal cortex; IPL = inferior parietal lobule.

studies. Regions that showed increased activity when engaging in expressive suppression to pain faces ($eSUP_{Face} > PAIN_{Face}$) included the ventromedial prefrontal cortex (vmPFC), bilateral IPL, and right middle temporal gyrus (MTG), consistent with regions previously implicated in emotion regulation and expressive suppression specifically (Cutuli, 2014; Dörfel et al., 2014; Goldin et al., 2008; Hayes et al., 2010; Ochsner, Silvers, & Buhle, 2012) (Fig. 3b, Table 3). Regions that showed decreased activity when engaging in expressive suppression to pain faces ($PAIN_{Face} > eSUP_{Face}$) included regions involved in motor control and negative emotion, including the SMA, visual cortex, bilateral amygdala, bilateral IPL, central operculum, and putamen.

Expressive suppression to negative IAPS pictures elicits similar activity to pain faces As a test of the specificity of our whole-brain findings to pain facial expressions, we conducted a whole-brain GLM analysis to examine activity differences between negative and neutral IAPS pictures. When passively viewing negative versus neutral IAPS pictures ($NEG_{IAPS} > NEU_{IAPS}$), there were increases in activity in the right middle frontal gyrus, dorsomedial prefrontal cortex, bilateral insular cortex, superior frontal gyrus, middle temporal gyrus, caudate, and bilateral amygdala, consistent with prior studies of negative emotion elicited by negative IAPS pictures (Aldhafeeri, Mackenzie, Kay, Alghamdi, & Sluming, 2012; Britton, Taylor, Sudheimer, & Liberzon, 2006) (Figure 3c; Table 3).

Examining expressive suppression-related activity in response to IAPS pictures ($eSUP_{IAPS} > NEG_{IAPS}$), which is the most frequently examined stimulus type in emotion regulation studies (Kohn et al., 2014), revealed increased activity in many of the same brain regions that we found to be activated when viewing pain faces. These included the right middle frontal gyrus, middle temporal gyrus, bilateral IPL, and ventrolateral/ventromedial prefrontal cortex (Figure 3d; Table 3). Regions that showed decreased activity when engaging in expressive suppression to negative IAPS pictures ($NEG_{IAPS} > eSUP_{IAPS}$) included the SMA, visual cortex, parietal cortex, left insular cortex, and medial/anterior cingulate cortex.

Together, these findings suggest that our tasks robustly activated brain regions reported in prior studies of pain empathy and expressive suppression, providing a manipulation check of our tasks. Our finding that similar brain regions were activated by passive viewing and expressive suppression to pain faces and negative IAPS pictures is consistent with our multivariate pattern expression findings, in which we found that the observed reduction in neural representations of negative emotion due to expressive suppression was not specific to pain facial expressions. Our results, particularly with regard to increases in emotion regulation-related brain activity during expressive suppression, also provide a complement to the specific neural signatures used in our multivariate pattern expression analysis, which may not fully capture increases in brain activity associated with emotion regulation processes.

Table 3 List of peak coordinates from significant clusters in the univariate general linear model (GLM) analysis (FDR $q < 0.05$, $k > 50$)

Brain region	Cluster size	Peak Z	Peak MNI (x, y, z)
PAIN_{Face} > NEU_{Face}			
L frontal pole	18179	6.37	-34, 56, 14
L cerebellum	6223	4.55	-38, -62, -38
R inferior frontal gyrus (pars triangularis)	6066	5.41	52, 20, 0
L supramarginal gyrus (posterior division)	3900	4.94	-58, -50, 22
R middle temporal gyrus	3245	4.85	56, -48, 0
L cuneal cortex	537	3.57	-4, -84, 38
R posterior cingulate gyrus	344	3.78	4, -18, 24
postcentral gyrus	191	3.55	0, -50, 76
R lingual gyrus	91	2.73	26, -54, -8
brainstem	52	3.04	-10, -28, -28
eSUP_{Face} > PAIN_{Face}			
R frontal pole	5203	5.37	46, 54, -10
R lateral occipital cortex	3661	4.8	56, -62, 42
L lateral occipital cortex	2871	4.89	-50, -60, 48
L superior temporal gyrus	1940	4.06	-54, -8, -12
R posterior cingulate gyrus	1334	3.83	4, -32, 46
R middle temporal gyrus	736	4.36	58, -16, -20
R precentral gyrus	612	3.39	24, -26, 48
L middle frontal gyrus	540	3.7	-44, 18, 42
L frontal pole	267	3.67	-42, 56, -8
L precentral gyrus	104	2.79	-32, -26, 56
L frontal pole	98	3.04	-10, 62, -10
L frontal orbital cortex	96	3.16	-16, 22, -14
NEG_{IAPS} > NEU_{IAPS}			
R superior frontal gyrus	86620	6.55	6, 14, 58
R superior parietal lobule	248	3.05	26, -54, 38
L frontal pole	72	3.79	-4, 62, -22
eSUP_{IAPS} > NEG_{IAPS}			
R lateral occipital cortex	1998	4.31	52, -60, 28
L lateral occipital cortex	1247	5.2	-60, -64, 26
R middle frontal gyrus	1074	4.34	36, 26, 38
R frontal pole	247	3.62	44, 42, 8
R frontal pole	93	3.85	40, 60, -14
R middle temporal gyrus	84	3.05	64, -22, -20
L inferior temporal gyrus	66	3.16	-50, -18, -22
R anterior cingulate gyrus	63	3.11	10, 34, -2
L frontal pole	59	3.06	-42, 54, -14
L middle temporal gyrus	58	3.55	-72, -36, -2

Discussion

Summary of findings

The suppression of emotional expressivity when viewing pain in others may be appropriate in many different contexts. However, the effect that this suppression of expressivity has on the pain observer is still not well understood. In the present study, we tested whether engaging in the emotion regulation

strategy *expressive suppression* in response to pain in others reduced neural representations of negative emotion (as measured by the PINES), vicarious pain (as measured by the FEIVPS), or both. We found that viewing pain faces increased neural representations of negative emotion and vicarious pain, yet expressive suppression to pain faces decreased neural representations of negative emotion only. Tests of the specificity of this effect demonstrated that it was not specific to pain facial expressions, as similar effects were observed in

response to IAPS pictures. Connecting neural representations of negative emotion and trait-level empathy, PINES responses were positively associated with participants' perspective-taking and pain face unpleasantness ratings. Addressing our secondary research goal and speaking to the generalizability of these effects, the effect of expressive suppression on PINES responses did not differ by participant gender or nationality.

Expressive suppression to pain in others reduced neural representations of negative emotion but not vicarious pain

Prior evidence links the experience of negative emotion when viewing pain in others to the experience of pain empathy (Simon et al., 2008), consistent with current theories of pain empathy as a primarily affective phenomenon (Lamm et al., 2011; Lamm & Majdandžić, 2015). We found that the pain faces in our study increased neural representations of both negative emotion and vicarious pain. Yet engaging in expressive suppression to pain faces reduced neural representations of negative emotion only. This suggests that expressive suppression to pain faces was associated with a decrease in the affective aspect of pain empathy. Our findings that PINES responses were positively associated with perspective-taking and pain face unpleasantness ratings, and that individual difference measures related to expressive suppression and empathy were negatively related in our sample, provide further support for a link between the neural representations of negative emotion examined in our study and the affective aspect of empathy. Our replication of prior whole-brain findings of increased aINS and dACC activity when viewing pain in others (Lamm et al., 2011) provides support for a link between viewing the pain faces in our study and empathy. Despite these findings, it is important to note that empathy is a multidimensional construct, with multiple definitions and contributing neural processes (Klimecki, Leiberg, Ricard, & Singer, 2014; Lamm & Majdandžić, 2015). Similar to previous studies using MVPA-derived neural signatures to investigate empathy (López-Solà, Koban, et al., 2017), although we utilized the previously developed PINES and FEIVPS signatures to understand the influence of expressive suppression on different aspects of empathic responding, we do not suggest that these are the only neural signatures of negative emotion, vicarious pain, or empathy, or that these represent complete models of the brain systems contributing to these constructs. In sum, our findings suggest that suppressing facial expressivity when viewing pain in others reduces the negative affective aspect of pain empathy, but not the experience of vicarious pain itself.

Prior research on expressive suppression indicates that it is associated with largely negative social (Butler et al., 2007), cognitive (Moore & Zoellner, 2012; Niermeyer et al., 2019; Richards & Gross, 2000), and affective (Butler et al., 2003;

Schneider et al., 2013) consequences. Our findings implicate a reduction in the affective aspect of pain empathy as an additional negative consequence of expressive suppression. However, it is important to note that the consequences of expressive suppression may rely largely on the context in which it is employed (Butler et al., 2003). For example, physician training often focuses on specifically downregulating the affective or emotionally distressing aspects of conducting medical care (Nightingale, Yarnold, & Greenberg, 1991), and empathy has been found to decline as medical training increases (Bellini & Shea, 2005; Hojat et al., 2009). In experimental studies, evidence indicates that prefronto-parietal regions involved in emotion regulation may be utilized by physicians during pain observation to reduce the associated negative arousal that may inhibit appropriate helping behaviors (Decety et al., 2010). In conclusion, although our study did not include a sample of physicians, our findings suggest that inhibiting facial expressivity in response to pain in others may be an effective strategy for reducing the negative emotion associated with observing pain in others within medical contexts.

Effect of expressive suppression on neural representations of negative emotion did not differ by nationality

Addressing our secondary research aim, which investigated whether there were cross-national differences in the effect of expressive suppression on viewing pain in others, we found similar magnitudes of PINES responses when viewing and engaging in expressive suppression to pain faces and IAPS pictures in both Chinese and U.S. participants. This suggests that although the PINES pattern was developed using U.S. participants, the neural representations associated with negative emotion that it measures may generalize to other national groups. Although our findings suggest that the specific negative consequence of expressive suppression examined in the present study—reduced brain activity related to empathy when viewing pain in others—may not differ between the specific cultural groups studied, it is possible that there are other sociocultural or contextual variables that may influence how and when expressive suppression leads to negative consequences. Future research is needed to investigate the sociocultural and contextual factors influencing the effects of expressive suppression on pain in others.

Potential neural mechanisms of effect of expressive suppression on viewing pain in others

Emotion regulation is thought to involve increases in brain activity in prefrontal cortex regions involved in cognitive control and decreases in subcortical regions involved in emotion, particularly the amygdala (Ochsner et al., 2012). In a fMRI

study comparing multiple emotion regulation strategies, including expressive suppression, Dörfel et al. (2014) reported that detachment, expressive suppression, and distraction were all associated with increased activity in a right prefronto-parietal emotion regulation network and decreased activity in the left amygdala. Our multivariate and univariate results are consistent with these and other prior findings of reduced activity in motor and affective-related brain regions during expressive suppression (Dörfel et al., 2014; Ohira et al., 2006; Vrticka et al., 2013). Specifically, we found that expressive suppression to pain faces was associated with increased activity in the vmPFC and bilateral IPL, as well as decreased activity in the left amygdala. In terms of our multivariate pattern findings, it is noteworthy that the left amygdala is prominently represented with positive predictive weights (i.e., increased activity predicts more negative emotion) in the PINES pattern. Changing activity in the left amygdala may have thus contributed to the responsiveness of the PINES to the pain faces, as well as driven subsequent reductions in PINES responses during expressive suppression. Alternatively, the affective salience of the pain faces may have been the primary driver of amygdala activity given the amygdala's role in assigning affective salience to environmental stimuli (Phelps & LeDoux, 2005; Yu et al., 2020). The IPL is implicated in self-face recognition (Sugiura et al., 2005) and distinguishing between self and other (Uddin, Molnar-Szakacs, Zaidel, & Iacoboni, 2006). Activation in this region during expressive suppression may thus reflect participants' continuous awareness of their own facial expressions (Dörfel et al., 2014). The IPL also may function to direct attention to perceptual inputs held in working memory (Culham & Kanwisher, 2001; Ochsner et al., 2004). At the multivariate pattern level, the IPL is represented with negative predictive weights (i.e., increased activity predicts less negative emotion) in the PINES pattern. Together, these results suggest that increased activity in the vmPFC and IPL, along with decreased activity in the left amygdala, may be potential neural mechanisms underlying the effect of suppressing facial expressivity when viewing pain in others.

Limitations

The results of the present study should be considered in the light of several limitations. First, consistent with prior emotion regulation studies (Dörfel et al., 2014; Staudinger, Erk, Abler, & Walter, 2009; Staudinger, Erk, & Walter, 2011; Walter et al., 2009), we chose to have participants complete stimulus ratings immediately after the fMRI scan, rather than during scanning, given prior indication that behavioral ratings of negative affective stimuli conducted during scanning may interrupt amygdala activity (Hariri, Mattay, Tessitore, Fera, & Weinberger, 2003; Lieberman et al., 2011). However, due to

time constraints on the overall length of each study session, participants passively viewed each picture they rated but did not also engage in expressive suppression to each picture. As a result, we were unable to test how expressive suppression affected stimulus ratings. Second, our emotion regulation task did not have a condition in which participants engaged in expressive suppression to neutral expression faces. Therefore, we could not assess brain activity related to expressive suppression separate from the emotional aspects of the stimuli. Prior emotion regulation studies have typically excluded neutral emotion regulation conditions, however, likely because emotion regulation studies have been historically concerned with how responses to negative affective stimuli are regulated (Gross, 1998). Finally, we did not include an objective measure of facial muscle movement, such as facial electromyography (EMG), to confirm that participants successfully suppressed their facial expressivity when cued in the scanner. However, in-person training conducted prior to each fMRI scan included practice trials for participants to demonstrate expressive suppression in front of the experimenter. Additionally, a camera feed (without recording ability) allowed the experimenters to observe participants' facial expressions during the scan at the scanning site in China, although this was not available at the U.S. scanning site.

Conclusions

The present study is important for connecting two research domains, emotion regulation and pain empathy, whose interactions have been to date understudied. The present study further utilizes an MVPA approach that encompasses activity distributed across cortical and subcortical regions throughout the brain to understand the relationship between emotion regulation and pain empathy. Our use of two different stimulus types—faces and IAPS pictures—had the advantage of allowing us to assess the effects of expressive suppression independent of a single stimulus type. In addition to replicating prior neuroimaging findings on the neural correlates of pain empathy, this study also adds to our understanding of the neural correlates of expressive suppression, an emotion regulation strategy that has been considerably less examined to date using neuroimaging than other emotion regulation strategies (Cutuli, 2014). Our findings suggest that a consequence of suppressing facial expressivity in response to the pain of others may be a reduction in the affective aspect of pain empathy but not the experience of vicarious pain itself. Applied to medical settings, our findings suggest that expressive suppression may be an effective strategy for reducing the negative emotion associated with observing pain in others. Future studies with physician and patient samples are needed to understand the neurobiological and behavioral

consequences of engaging in expressive suppression when viewing pain in others in medical settings.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13415-021-00873-1>.

Acknowledgments The authors thank Xiaochun Han for her helpful assistance in data collection. The authors also thank Lucina Uddin and Ekaterina Denkova for providing feedback on this project. For generous support, the authors thank the U.S. National Science Foundation East Asia and Pacific Summer Institute Program (NSF EAPSI; Award Number: 1713969), which supported the first author while conducting research in Beijing, China, the University of Miami Antonio Orlando Award and institutional startup funds, which covered scanning and participant costs in the United States, and the National Natural Science Foundation of China (Projects 31661143039, 31421003), which covered scanning and participant costs in China. Research reported in this publication was supported by the National Institute On Drug Abuse of the National Institutes of Health under Award Number K01DA045735. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declarations

Conflict of Interest The authors have no conflicts of interest to disclose.

Open Practices Statement Brain data used in this study is available on NeuroVault: <https://neurovault.org/collections/XKWLLVBQ/>.

Analyses conducted in this study are available in an R Markdown file on Open Science Framework: https://osf.io/ewq6k/?view_only=d97b82b67a804d31b7cc407bb8374874.

References

- Aldhafeeri, F. M., Mackenzie, I., Kay, T., Alghamdi, J., & Sluming, V. (2012). Regional brain responses to pleasant and unpleasant IAPS pictures: different networks. *Neuroscience letters*, 512(2), 94–98.
- Avenanti, A., Minio-Paluello, I., Sforza, A., & Aglioti, S. M. (2009). Freezing or escaping? Opposite modulations of empathic reactivity to the pain of others. *Cortex*, 45(9), 1072–1077.
- Bates D, M. M., Bolker B and Walker S (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multi-level linear modeling for group analysis in fMRI. *NeuroImage*, 20(2), 1052–1063.
- Bellini, L. M., & Shea, J. A. (2005). Mood change and empathy decline persist during three years of internal medicine training. *Academic medicine*, 80(2), 164–167.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.
- Britton, J. C., Taylor, S. F., Sudheimer, K. D., & Liberzon, I. (2006). Facial expressions and complex IAPS pictures: common and differential networks. *Neuroimage*, 31(2), 906–919. Retrieved from https://ac.els-cdn.com/S1053811905025565/1-s2.0-S1053811905025565-main.pdf?_tid=bd192d96-0e95-11e8-b11e-00000aab0f01&acdnat=1518289872_b7d2764e8530e080f06b63f35b7e236
- Butler, E. A., Egloff, B., Wilhelm, F. H., Smith, N. C., Erickson, E. A., & Gross, J. J. (2003). The social consequences of expressive suppression. *Emotion*, 3(1), 48.
- Butler, E. A., Lee, T. L., & Gross, J. J. (2007). Emotion regulation and culture: are the social consequences of emotion suppression culture-specific? *Emotion*, 7(1), 30.
- Butler, E. A., Lee, T. L., & Gross, J. J. (2009). Does expressing your emotions raise or lower your blood pressure? The answer depends on cultural context. *Journal of Cross-Cultural Psychology*, 40(3), 510–517.
- Caes, L., Vervoort, T., Devos, P., Verlooy, J., Benoit, Y., & Goubert, L. (2014). Parental distress and catastrophic thoughts about child pain: Implications for parental protective behavior in the context of child leukemia-related medical procedures. *The Clinical journal of pain*, 30(9), 787–799.
- Chang, L. J., Gianaros, P. J., Manuck, S. B., Krishnan, A., & Wager, T. D. (2015). A sensitive and specific neural signature for picture-induced negative affect. *PLoS biology*, 13(6), e1002180. <https://doi.org/10.1371/journal.pbio.1002180>
- Che, X., Luo, X., Tong, D., Fitzgibbon, B. M., & Yang, J. (2015). Habitual suppression relates to difficulty in regulating emotion with cognitive reappraisal. *Biological psychology*, 112(Supplement C), 20–26. <https://doi.org/10.1016/j.biopsycho.2015.09.011>
- Cheng, Y., Lin, C.-P., Liu, H.-L., Hsu, Y.-Y., Lim, K.-E., Hung, D., & Decety, J. (2007). Expertise modulates the perception of pain in others. *Current Biology*, 17(19), 1708–1713. <https://doi.org/10.1016/j.cub.2007.09.020>
- Chikazoe, J., Lee, D. H., Kriegeskorte, N., & Anderson, A. K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nature neuroscience*, 17(8), 1114.
- Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: gender effects in brain and behavior. *Neuroscience & Biobehavioral Reviews*, 46, 604–627.
- Corradi-Dell'Acqua, C., Hofstetter, C., & Vuilleumier, P. (2011). Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. *Journal of Neuroscience*, 31(49), 17996–18006.
- Costafreda, S. G. (2009). Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. *Frontiers in neuroinformatics*, 3, 33.
- Culham, J. C., & Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion in Neurobiology*, 11(2), 157–163.
- Cutuli, D. (2014). Cognitive reappraisal and expressive suppression strategies role in the emotion regulation: an overview on their modulatory effects and neural correlates. *Frontiers in Systems Neuroscience*, 8, 175. <https://doi.org/10.3389/fnsys.2014.00175>
- Danziger, N., Prkachin, K. M., & Willer, J. C. (2006). Is pain the price of empathy? The perception of others' pain in patients with congenital insensitivity to pain. *Brain*, 129, 2494–2507. <https://doi.org/10.1093/brain/awl155>
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, 44(1), 113–126.
- Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and psychopathology*, 20(04), 1053–1080.
- Decety, J., Yang, C.-Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: an event-related brain potential study. *NeuroImage*, 50(4), 1676–1682.
- Dörfel, D., Lamke, J.-P., Hummel, F., Wagner, U., Erk, S., & Walter, H. (2014). Common and differential neural networks of emotion regulation by Detachment, Reinterpretation, Distraction, and Expressive Suppression: A comparative fMRI investigation. *Neuroimage*, 101(Supplement C), 298–309. <https://doi.org/10.1016/j.neuroimage.2014.06.051>
- Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on

- performance attainment. *Journal of Experimental Psychology: General*, 136(1), 154.
- Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, 35(3), 903–911.
- Franck, L., Cox, S., Allen, A., & Winter, I. (2004). Parental concern and distress about infant pain. *Archives of disease in childhood-Fetal and Neonatal Edition*, 89(1), F71–F75.
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4), 870–878.
- Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological psychiatry*, 63(6), 577–586. Retrieved from http://ac.els-cdn.com/S0006322307005926/1-s2.0-S0006322307005926-main.pdf?_tid=8178a764-28fd-11e6-ac3b-00000aabb0f26&acdnat=1464898174_8b7a5f66c54cb60410bd5d55b9517e47
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*, 48(1), 1095–9572 (Electronic).
- Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, 74(1), 224.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>
- Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, 106(1), 95.
- Grühn, D., Rebusal, K., Diehl, M., Lumley, M., & Labouvie-Vief, G. (2008). Empathy across the adult lifespan: Longitudinal and experience-sampling findings. *Emotion*, 8(6), 753.
- Hadjistavropoulos, T., Craig, K. D., Duck, S., Cano, A., Goubert, L., Jackson, P. L., ... de C Williams, A. C. (2011). A biopsychosocial formulation of pain communication. *Psychological Bulletin*, 137(6), 910.
- Han, X., Luo, S., & Han, S. (2016). Embodied neural responses to others' suffering. *Cognitive Neuroscience*, 7(1–4), 114–127.
- Hariri, A. R., Mattay, V. S., Tessitore, A., Fera, F., & Weinberger, D. R. (2003). Neocortical modulation of the amygdala response to fearful stimuli. *Biological psychiatry*, 53(6), 494–501.
- Hayes, J. P., Morey, R. A., Petty, C. M., Seth, S., Smoski, M. J., McCarthy, G., & LaBar, K. S. (2010). Staying cool when things get hot: emotion regulation modulates neural mechanisms of memory encoding. *Frontiers in Human Neuroscience*, 4.
- Hill, R. A., & Barton, R. A. (2005). Psychology: red enhances human performance in contests. *Nature*, 435(7040), 293.
- Hobara, M. (2005). Beliefs about appropriate pain behavior: cross-cultural and sex differences between Japanese and Euro-Americans. *Eur J Pain*, 9(4), 389–393. <https://doi.org/10.1016/j.ejpain.2004.09.006>
- Hojat, M., Vergare, M. J., Maxwell, K., Brainard, G., Herrine, S. K., Isenberg, G. A., ... Gonnella, J. S. (2009). The devil is in the third year: a longitudinal study of erosion of empathy in medical school. *Academic Medicine*, 84(9), 1182–1191.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification. In: Taipei.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2), 825–841.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med Image Anal*, 5(2), 143–156.
- John, O. P., & Gross, J. J. (2004). Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development. *Journal of personality*, 72(6), 1301–1334. Retrieved from <http://onlinelibrary.wiley.com/store/10.1111/j.1467-6494.2004.00298.x/asset/j.1467-6494.2004.00298.x.pdf?v=1&t=jekaotgo&s=12080cd7af4ba4c5659e0dc6b082e718080487>
- Klimecki, O. M., Leiberg, S., Ricard, M., & Singer, T. (2014). Differential pattern of functional brain plasticity after compassion and empathy training. *Social Cognitive and Affective Neuroscience*, 9(6), 873–879.
- Kohn, N., Eickhoff, S. B., Scheller, M., Laird, A. R., Fox, P. T., & Habel, U. (2014). Neural network of cognitive emotion regulation—an ALE meta-analysis and MACM analysis. *Neuroimage*, 87, 345–355.
- Kragel, P. A., & LaBar, K. S. (2016). Decoding the nature of emotion in the brain. *Trends in Cognitive Sciences*, 20(6), 444–455.
- Krishnan, A., Woo, C. W., Chang, L. J., Ruzic, L., Gu, X., Lopez-Sola, M., ... Wager, T. D. (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *Elife*, 5, 1–42. <https://doi.org/10.7554/eLife.15166>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest: Tests in Linear Mixed Effects Models (Version R package version 2.0-29). Retrieved from <http://CRAN.R-project.org/package=lmerTest>
- 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>
- Lamm, C., & Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy—A critical comment. *Neuroscience Research*, 90, 15–24.
- Lamm, C., Nusbaum, H. C., Meltzoff, A. N., & Decety, J. (2007). What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE*, 2(12), e1292.
- Lang, P., Bradley, M., & Cuthbert, B. (1995). International Affective Picture System (IAPS): NIMH Center for the Study of Emotion and Attention. *University of Florida, Gainesville, Florida, USA*.
- Lang, P., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. *Handbook of emotion elicitation and assessment*, 29.
- Larson, E. B., & Yao, X. (2005). Clinical empathy as emotional labor in the patient-physician relationship. *JAMA*, 293(9), 1100–1106.
- Lieberman, M. D., Inagaki, T. K., Tabibnia, G., & Crockett, M. J. (2011). Subjective responses to emotional stimuli during labeling, reappraisal, and distraction. *Emotion*, 11(3).
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences*, 35(3), 121
- Loggia, M. L., Mogil, J. S., & Bushnell, M. C. (2008). Empathy hurts: compassion for another increases both sensory and affective components of pain perception. *PAIN*, 136(1–2), 168–176.
- López-Solà, M., Geuter, S., Koban, L., Coan, J. A., & Wager, T. D. (2019). Brain mechanisms of social touch-induced analgesia in females. *Pain*, 160(9), 2072–2085.
- López-Solà, M., Koban, L., Krishnan, A., & Wager, T. D. (2017a). When pain really matters: A vicarious-pain brain marker tracks empathy for pain in the romantic partner. *Neuropsychologia*, 145
- López-Solà, M., Woo, C.-W., Pujol, J., Deus, J., Harrison, B. J., Monfort, J., & Wager, T. D. (2017b). Towards a neurophysiological signature for fibromyalgia. *Pain*, 158(1), 34–47. <https://doi.org/10.1097/j.pain.0000000000000707>
- Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). *Frontiers in psychology*, 6, 2.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253. Retrieved from <file://localhost/Users/elizabethRlosin/Documents/Sente/My%20Library.sente6lib/Contents/Attachments/1991/>

- Markus%201991%20Culture%20and%20the%20self%20Implications%20for.pdf
- Meng, J., Hu, L., Shen, L., Yang, Z., Chen, H., Huang, X., & Jackson, T. (2012). Emotional primes modulate the responses to others' pain: an ERP study. *Experimental Brain Research*, 220(3-4), 277-286.
- Moore, S. A., & Zoellner, L. A. (2012). The effects of expressive and experiential suppression on memory accuracy and memory distortion in women with and without PTSD. *Journal of Experimental Psychopathology*, 3(3), jep. 024411.
- Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, 16(4), 406.
- Mumford, J. A., & Nichols, T. (2006). Modeling and inference of multisubject fMRI data. *IEEE Engineering in Medicine and Biology Magazine*, 25(2), 42-51.
- Nayak, S., Shiflett, S. C., Eshun, S., & Levine, F. M. (2000). Culture and gender effects in pain beliefs and the prediction of pain tolerance. *Cross-Cultural Research*, 34(2), 135-151.
- Niermeyer, M. A., Ziemnik, R. E., Franchow, E. I., Barron, C. A., & Suchy, Y. (2019). Greater naturally occurring expressive suppression is associated with poorer executive functioning and motor-sequence learning among older adults. *Journal of Clinical and Experimental Neuropsychology*, 41(2), 118-132.
- Nightingale, S. D., Yarnold, P. R., & Greenberg, M. S. (1991). Sympathy, empathy, and physician resource utilization. *Journal of General Internal Medicine*, 6(5), 420-423.
- Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, 9(5), 242-249.
- Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage*, 23(2), 483-499.
- Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, 1251(1), E1-E24.
- Ohira, H., Nomura, M., Ichikawa, N., Isowa, T., Iidaka, T., Sato, A., ... Yamada, J. (2006). Association of neural and physiological responses during voluntary emotion suppression. *Neuroimage*, 29(3), 721-733. Retrieved from http://ac.els-cdn.com/S1053811905006038/1-s2.0-S1053811905006038-main.pdf?_tid=c3077e8c-1cad-11e6-8bf9-00000aab0f02&acdnat=1463544511_9cb4fafa13be80c5124f49ab92aa838
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, 48(2), 175-187. Retrieved from http://ac.els-cdn.com/S0896627305008238/1-s2.0-S0896627305008238-main.pdf?_tid=c0ba1726-0637-11e5-aaad-00000aacb361&acdnat=1432927403_01b87874689e0718661a4b186e9b33ed
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59(3), 2142-2154.
- R Core Team. (2019). A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*. Retrieved from <https://www.R-project.org/>
- Reddan, M. C., & Wager, T. D. (2018). Modeling pain using fMRI: from regions to biomarkers. *Neuroscience Bulletin*, 34(1), 208-215.
- Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *Journal of personality and social psychology*, 79(3), 410.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225-237.
- Schneider, K. G., Hempel, R. J., & Lynch, T. R. (2013). That "poker face" just might lose you the game! The impact of expressive suppression and mimicry on sensitivity to facial expressions of emotion. *Emotion*, 13(5), 852.
- Shahane, A. D., Lopez, R. B., & Denny, B. T. (2019). Implicit reappraisal as an emotional buffer: Reappraisal-related neural activity moderates the relationship between inattention and perceived stress during exposure to negative stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, 19(2), 355-365.
- Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Iardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of Personality and Social Psychology*, 73(6), 1380.
- Sheng, F., & Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *Neuroimage*, 61(4), 786-797. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1053811912004247>
- Sheng, F., Liu, Q., Li, H., Fang, F., & Han, S. H. (2014). Task modulations of racial bias in neural responses to others' suffering. *Neuroimage*, 88, 263-270. <https://doi.org/10.1016/j.neuroimage.2013.10.017>
- Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., & Petersen, S. E. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Human brain mapping*, 35(5), 1981-1996.
- Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2008). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *Pain*, 135(1-2), 55-64.
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, 20(5), 580. <https://doi.org/10.1177/0146167294205014>
- Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157-1162. Retrieved from <http://science.sciencemag.org/content/sci/303/5661/1157.full.pdf>
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, 208-219.
- Soto, J. A., Levenson, R. W., & Ebling, R. (2005). Cultures of moderation and expression: emotional experience, behavior, and physiology in Chinese Americans and Mexican Americans. *Emotion*, 5(2), 154.
- Soto, J. A., Perez, C. R., Kim, Y.-H., Lee, E. A., & Minnick, M. R. (2011). Is expressive suppression always associated with poorer psychological functioning? A cross-cultural comparison between European Americans and Hong Kong Chinese. *Emotion*, 11(6), 1450.
- Srivastava, S., Tamir, M., McGonigal, K. M., John, O. P., & Gross, J. J. (2009). The social costs of emotional suppression: A prospective study of the transition to college. *Journal of Personality and Social Psychology*, 96(4), 883. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4141473/pdf/nihms614281.pdf>
- Staudinger, M. R., Erk, S., Abler, B., & Walter, H. (2009). Cognitive reappraisal modulates expected value and prediction error encoding in the ventral striatum. *Neuroimage*, 47(2), 713-721. <https://doi.org/10.1016/j.neuroimage.2009.04.095>
- Staudinger, M. R., Erk, S., & Walter, H. (2011). Dorsolateral prefrontal cortex modulates striatal reward encoding during reappraisal of reward anticipation. *Cerebral Cortex*, 21(11), 2578-2588.
- Sugiura, M., Watanabe, J., Maeda, Y., Matsue, Y., Fukuda, H., & Kawashima, R. (2005). Cortical mechanisms of visual self-recognition. *Neuroimage*, 24(1), 143-149.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.

- Uddin, L. Q., Molnar-Szakacs, I., Zaidel, E., & Iacoboni, M. (2006). rTMS to the right inferior parietal lobule disrupts self-other discrimination. *Social Cognitive and Affective Neuroscience*, *1*(1), 65–71. <https://doi.org/10.1093/scan/nsl003>
- Vanderhasselt, M.-A., Baeken, C., Van Schuerbeek, P., Luypaert, R., & De Raedt, R. (2013). Inter-individual differences in the habitual use of cognitive reappraisal and expressive suppression are associated with variations in prefrontal cognitive control for emotional information: An event related fMRI study. *Biological psychology*, *92*(3), 433–439. <https://doi.org/10.1016/j.biopsycho.2012.03.005>
- Vervoort, T., Caes, L., Trost, Z., Sullivan, M., Vangronsveld, K., & Goubert, L. (2011). Social modulation of facial pain display in high-catastrophizing children: an observational study in schoolchildren and their parents. *PAIN®*, *152*(7), 1591–1599.
- Vrticka, P., Simioni, S., Fornari, E., Schlupe, M., Vuilleumier, P., & Sander, D. (2013). Neural substrates of social emotion regulation: a fMRI study on imitation and expressive suppression to dynamic facial signals. *Frontiers in Psychology*, *4*, 95.
- Wager, T. D., Atlas, L. Y., Leotti, L. A., & Rilling, J. K. (2011). Predicting individual differences in placebo analgesia: contributions of brain activity during anticipation and pain experience. *J Neurosci*, *31*(2), 439–452. <https://doi.org/10.1523/JNEUROSCI.3420-10.2011>
- Wager, T. D., Atlas, L. Y., Lindquist, M. A., Roy, M., Woo, C. W., & Kross, E. (2013). An fMRI-based neurologic signature of physical pain. *The New England Journal of Medicine*, *368*(15), 1388–1397. <https://doi.org/10.1056/NEJMoa1204471>
- Walter, H., von Kalckreuth, A., Schardt, D., Stephan, A., Goschke, T., & Erk, S. (2009). The temporal dynamics of voluntary emotion regulation. *PLoS ONE*, *4*(8), e6726.
- Weilenmann, S., Schnyder, U., Parkinson, B., Corda, C., Von Kaenel, R., & Pfaltz, M. C. (2018). Emotion transfer, emotion regulation, and empathy-related processes in physician–patient interactions and their association with physician well-being: a theoretical model. *Frontiers in psychiatry*, *9*, 389.
- Woolrich, M., Brady, M., & Smith, S. M. (2001). Hierarchical fully Bayesian spatio-temporal analysis of FMRI data. *Neuroimage*, *13*(6), 1053–8119.
- Xu, X. J., Zuo, X. Y., Wang, X. Y., & Han, S. H. (2009). Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses. *Journal of Neuroscience*, *29*(26), 8525–8529. <https://doi.org/10.1523/Jneurosci.2418-09.2009>
- Yamada, M., & Decety, J. (2009). Unconscious affective processing and empathy: an investigation of subliminal priming on the detection of painful facial expressions. *PAIN*, *143*(1–2), 71–75.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–670
- Yu, H., Koban, L., Chang, L. J., Wagner, U., Krishnan, A., Vuilleumier, P., . . . Wager, T. D. (2020). A generalizable multivariate brain pattern for interpersonal guilt. *Cerebral Cortex*, *30*(6), 3558–3572.
- Zaki, J., Wager, T. D., Singer, T., Keysers, C., & Gazzola, V. (2016). The Anatomy of Suffering: Understanding the Relationship between Nociceptive and Empathic Pain. *Trends in Cognitive Sciences*, *20*(4), 249–259. <https://doi.org/10.1016/j.tics.2016.02.003>
- Zhou, F., Li, J., Zhao, W., Xu, L., Zheng, X., Fu, M., . . . Becker, B. (2020). Empathic pain evoked by sensory and emotional-communicative cues share common and process-specific neural representations. *Elife*, *9*, e56929.

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