OXFORD

Cerebral Cortex, 2023, **33**, 3421–3436
https://doi.org/10.1093/cercor/bhac281
Advance access publication date 24 August 2022
Original Article

Doctor trustworthiness influences pain and its neural correlates in virtual medical interactions

Steven R. Anderson [1], Morgan Gianola[1], Natalia A. Medina[1], Jenna M. Perry[1], Tor D. Wager[2], Elizabeth A. Reynolds Losin[1,*]

[1]Department of Psychology, University of Miami, 5665 Ponce de Leon Boulevard, Coral Gables, FL 33146-0751, USA,
[2]Department of Psychological and Brain Sciences, Dartmouth College, 3 Maynard St, Hanover, NH 03755-3565, USA

*Corresponding author: Department of Psychology, University of Miami, 5665 Ponce de Leon Boulevard, Coral Gables, FL 33146-0751, USA. Email: e.losin@miami.edu

Trust is an important component of the doctor-patient relationship and is associated with improved patient satisfaction and health outcomes. Previously, we reported that patient feelings of trust and similarity toward their clinician predicted reductions in evoked pain in response to painful heat stimulations. In the present study, we investigated the brain mechanisms underlying this effect. We used face stimuli previously developed using a data-driven computational modeling approach that differ in perceived trustworthiness and superimposed them on bodies dressed in doctors' attire. During functional magnetic resonance imaging, participants (n = 42) underwent a series of virtual medical interactions with these doctors during which they received painful heat stimulation as an analogue of a painful diagnostic procedure. Participants reported increased pain when receiving painful heat stimulations from low-trust doctors, which was accompanied by increased activity in pain-related brain regions and a multivariate pain-predictive neuromarker. Findings suggest that patient trust in their doctor may have tangible impacts on pain and point to a potential brain basis for trust-related reductions in pain through the modulation of brain circuitry associated with the sensory-discriminative and affective-motivational dimensions of pain.

Key words: doctor–patient relationship; pain; fMRI; medical trust.

Introduction

A patient's trust in their doctor, or the feeling that the doctor will do what is best for them (Thom and Campbell 1997), is a key component of the doctor–patient relationship (Hall et al. 2001). A positive doctor–patient relationship is known to impact multiple patient health outcomes including increasing placebo response (Kaptchuk et al. 2008; Howe et al. 2017) and patient satisfaction (Platonova et al. 2008), reducing physiological arousal when receiving a cancer diagnosis (Sep et al. 2014), and affecting clinically meaningful endpoints such as disease morbidity (Ward et al. 2003), postoperative pain (Gittell et al. 2000), and pain-related disability (Ferreira et al. 2013). Although patients' trust in their doctor has been shown to impact health outcomes, the neurobiological mechanisms driving trust effects on patient health remain unclear.

Understanding the effects of trustworthiness on pain in medical contexts may be especially worthwhile. Pain is common to most medical disorders and is a leading reason for patients to seek medical treatment (Loeser and Melzack 1999). Pain is also very costly to society in terms of financial and disability burden (IOM 2011) and is particularly challenging to treat given the risk of addiction (Martell et al. 2007) and poor efficacy (Shaheed et al. 2016) of opioid analgesic medications. An effect of trustworthiness on pain is likely given the growing body of literature demonstrating the modulation of pain by social, cultural, and interpersonal factors (Krahé et al. 2013; Mogil 2015; Koban and Wager 2016; Anderson and Losin 2017; López-Solà et al. 2019; Anderson et al. 2020; Goldstein et al. 2020; Gianola et al. 2021). In a previous study, we investigated the effects of clinician–patient trust and similarity on pain using face-to-face medical simulations and surveys about participants' core beliefs and values to manipulate feelings of cultural similarity and trust (Losin et al. 2017). We found that patients' feelings of clinician trustworthiness and self-similarity predicted reported pain in response to a painful diagnostic procedure analogue.

In the present study, we used functional magnetic resonance imaging (fMRI) and virtual medical interactions to investigate the brain mechanisms underlying the effect of doctor trustworthiness on pain. We enrolled a sample of healthy participants and administered experimental thermal stimulation as an analogue of a painful diagnostic procedure. To manipulate perceived trustworthiness, we used face stimuli previously developed using

Received: January 7, 2022. **Revised:** June 24, 2022. **Accepted:** June 25, 2022

a data-driven computational model approach that differ on perceived trustworthiness (Oosterhof and Todorov 2008; Todorov 2011; Todorov and Oh 2021). To simulate virtual medical interactions, we superimposed these face stimuli on bodies with doctors' attire. As our primary outcome measures, we focused on the effect of doctor trustworthiness on participant pain report and activity in brain regions that are implicated in clinical and experimentally induced pain (Melzack 1999, 2001; Davis 2000; Apkarian et al. 2005). These brain regions include those associated with the sensory-discriminative dimension of pain, including portions of the thalamus and somatosensory cortex, as well as those involved in the affective-motivational dimension of pain, including the anterior cingulate cortex (ACC) and anterior insula (aINS).

In order to provide a more sensitive and specific test of whether trust-related reductions in pain were related to pain-related brain circuitry, we also examined how doctor trustworthiness influenced responses to a previously developed multivariate pain-predictive signature, the Neurologic Pain Signature (NPS; Wager et al. 2013). The NPS includes patterns of activity that positively predict pain in the ACC, insula, secondary somatosensory cortex (S2), and thalamus and that negatively predict pain in the ventromedial prefrontal cortex (vmPFC) and precuneus. The NPS has demonstrated sensitivity and specificity to nociceptive pain across multiple pain modalities (Reddan and Wager 2018) and is not activated by nonnociceptive processes including viewing aversive images (Chang et al. 2015), observing pain in others (Krishnan et al. 2016), experiencing social rejection (Woo et al. 2014), or anticipating pain (Wager et al. 2011). Finally, as a follow-up to our primary analysis, we examined whether brain responses to pain during the virtual medical interaction were correlated with individual differences in participant ratings of doctor trustworthiness and participant mistrust in healthcare organizations more generally.

We predicted that participants playing the role of patients would report increased pain and exhibit increased pain-related brain activity in response to the painful diagnostic procedure analogue with virtual doctors who were low in facial trustworthiness based on the trustworthiness algorithm used to create the stimuli and whom participants rated as low in trustworthiness. We also predicted that increased ratings of the untrustworthiness of the doctor stimuli, and mistrust in medical organizations generally, would correlate with increased brain activity in pain-related brain regions during the painful diagnostic procedure analogue.

Materials and methods

Participants

This fMRI study was conducted with 42 healthy adults (55% female) age 18–35 ($M = 21.90$, $SD = 4.39$). Demographic characteristics for participants can be viewed in Table 1. Participants were recruited from the University of Miami and surrounding community in Miami, FL. Participants were excluded if they were: left-handed or ambidextrous, unable to tolerate the scanning procedures, had metal in their body or any other contraindication for MRI, had a prior abnormal MRI, had heavy alcohol intake (>3 drinks for women, >5 drinks for men) within 12 h prior to MRI scan, had a neurologic or systemic disorder causing cognitive impairment, had current presence of pain, were currently taking any pain medication, had a chronic pain or chronic fatigue syndrome, or had a history of psychiatric or substance use disorder in the past year (full exclusion criteria available in the Supplementary Material). All participants had normal vision or vision that was normal after correction. The study protocol was approved by the Institutional Review Board at the University of Miami. Written informed consent was obtained according to the Declaration of Helsinki (1991) for all research participants.

Table 1. Participant characteristics.

	$N = 42$
Age	
Mean (SD)	21.90 (4.39)
Median [Min, Max]	21.0 [18.0, 35.0]
Race	
White/Caucasian	24 (57.1%)
Black/African American	7 (16.7%)
Asian	7 (16.7%)
Other	4 (9.5%)
White & Asian	1 (2.4%)
Afro-Caribbean	1 (2.4%)
Not Indicated	2 (4.8%)
Ethnicity	
Non-Hispanic/Latino	27 (64.3%)
Hispanic/Latino	15 (35.7%)
Gender	
Male	19 (45.2%)
Female	23 (54.8%)

Self-report measures

Individual difference measures and stimulus ratings

Prior to arriving in the lab, participants made a series of ratings of the doctor stimuli at home via Qualtrics. These included ratings of doctors' trustworthiness, competence, warmth, attractiveness, and likelihood of choosing as one's own doctor (1 = "not at all" to 9 = "extremely" for all scales). Participants were instructed to use their "gut instinct" in making each rating. Participants completed the Medical Mistrust Index (MMI; LaVeist et al. 2009), a 17-item measure of mistrust in healthcare organizations (e.g. "Patients have sometimes been deceived or misled by health care organizations"). Immediately following the fMRI scan, participants completed questions assessing the realism of the virtual medical interaction ("How realistic did the simulated clinical interactions in the scanner feel to you?" 0 = Not at all, 100 = Completely) and belief in the stated aim of the study ("How much

did you believe in the stated purpose of the study (i.e., that we were studying factors that affect pain perception during medical care?)" 0 = Not at all, 100 = Completely). In addition, participants completed a battery of questionnaires at various points in the study not used in the current analyses to measure potential psychological and sociocultural contributors to doctor trust level effects on pain (see Supplementary Material for further details). Participants were instructed to consider their questionnaire responses in relation to their prior real-world medical encounters.

fMRI screening session

Participants completed a screening session on a separate visit prior to the fMRI scan to ensure that they could tolerate the painful heat stimulations during the scan. During the screening session, participants first completed informed consent with an experimenter. Next, the experimenter administered a series of painful heat stimulations to the participant's left forearm to simulate a painful diagnostic procedure commonly performed in an outpatient setting, such as a biopsy, musculoskeletal exam of an injury, or mammogram. Participants were instructed to stay as still as possible during the heat stimulations as practice for the fMRI scan. Based on thermal heat stimulation protocols published in previous studies (Atlas et al. 2010; Wager et al. 2013; Losin et al. 2017), heat stimulations were delivered to evenly spaced locations on the volar surface of the participant's left forearm using a 16 mm × 16 mm contact Peltier thermode from a Medoc Pathway Pain & Sensory Evaluation System (Medoc, Inc.).

In order to assess patient pain ratings in response to fixed levels of heat intensity, which have been used in prior studies (Edwards and Fillingim 1999), participants received a series of suprathreshold stimulations to prepare them for the scanning session. Each suprathreshold heat stimulation lasted 8 s in total (4.6 s at the target temperature), after which participants rated their pain intensity and pain unpleasantness on a 0–10 numeric rating scale (NRS). Participants were instructed on the difference between pain intensity and unpleasantness using previously developed language (Price et al. 1983), which describes how pain intensity measures how strong a painful stimulus feels, while pain unpleasantness measures how unpleasant or disturbing a painful stimulus is for the individual. Pain unpleasantness is thought to reflect the affective-motivational aspect of pain (Price et al. 1983).

In total, participants received 10 heat stimulations during the screening session. Trials assessed pain threshold (the temperature at which a heat stimulation is first perceived as painful), pain tolerance (the temperature at which a heat stimulation is no longer tolerable), and suprathreshold responses (fixed duration heat stimulations at target stimulus intensity levels above the pain threshold). Pain threshold (2 trials) and tolerance (2 trials) always preceded the suprathreshold (6 trials) assessment. Trial order for the threshold and tolerance trials was fixed to ensure that the amount of heat delivered to any individual skin site did not exceed previously determined safety limits (Buhle and Wager 2010) over the course of the first 2 blocks. Trial order for the suprathreshold trials was pseudorandomized such that no skin site received more than one heat stimulation in a row. The temperature of the suprathreshold stimulations was randomized.

After the heat stimulations, the experimenters reviewed the participant's pain ratings in a separate room. In addition to the study exclusion criteria, participants were ineligible for the fMRI scanning session if they: (i) asked to have the thermode removed more than once during the heat stimulations; (ii) had pain ratings that did not generally increase with stimulus intensity level (temperature); (iii) were unable to stay still during the heat stimulations; (iv) were unable to consistently complete the pain ratings as instructed; (v) were in substantial distress due to the heat stimulations; or (vi) had on their left forearm a previously unreported tattoo, metallic implant, scar tissue, or other feature that could influence their safety and/or pain sensitivity in the scanner. A total of 23 participants were excluded from the study after completing the fMRI screening session (4 participants excluded due to inability to tolerate the heat stimulations, 19 participants excluded due to scheduling issues, computer error, or study exclusion criteria). Demographic information for the participants excluded from the fMRI scan during the screening session can be viewed in Supplemental Table S1.

Trust face stimuli

Facial appearances are used to infer underlying personality traits (Hassin and Trope 2000; Zebrowitz et al. 2002) rapidly and spontaneously (Willis and Todorov 2006; Rule et al. 2013). Given the complexities of defining and measuring patient trust in medical settings (Pearson and Raeke 2000), and especially the challenges of manipulating clinician behavior in a fMRI environment, we manipulated doctor trust level and examined its effects on pain with face stimuli developed to vary in perceived trustworthiness developed using a data-driven computational modeling approach (Oosterhof and Todorov 2008; Todorov 2011; Todorov and Oh 2021). We contextualized these face stimuli by superimposing them on an image of a body dressed in typical doctor's attire. Face stimuli were +3 (high-trust) and −3 (low-trust) standard deviations from neutral on the trustworthiness dimension (Oosterhof and Todorov 2008). All task stimuli viewed by participants were delivered via an experiment using Presentation software (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA, USA). Doctor trust level was the primary experimental manipulation. A total of 25 different face identities were shown to participants over the course of the study, with an equal number of low- and high-trust doctor identities. Although each participant

saw the same face identities, the trust level assigned to the face identities was counterbalanced across participants to control for any interaction between a specific face identity and trust level. In total, 4 counterbalancing orders were used. Each participant was assigned a counterbalancing order prior to seeing any of the doctor face stimuli and saw the same combination of face identities and trust levels throughout their participation in the study.

fMRI task

Participants completed an fMRI task in which they underwent a series of virtual medical interactions and received painful thermal stimulations as an analogue to a painful diagnostic procedure. Each virtual medical interaction consisted of the following parts: doctor introduction, pain anticipation, painful heat stimulation, and pain rating. Painful heat stimulations were delivered using the same equipment described in the fMRI Screening Session. An fMRI-compatible Medoc thermode was placed on a single skin site (in randomized order) for the duration of each fMRI run (4 runs total). As prior evidence suggests that patient trust in their doctor may be more important as perceived harm increases (Mascarenhas et al. 2006; Lyu et al. 2017), participants received both a low and high stimulus intensity heat stimulation from each doctor (in randomized order) to examine the interaction between stimulus intensity and trust effects on pain.

At the beginning of the fMRI task, participants viewed a fixation cross (20 s) for an implicit baseline. Next, participants received a single heat stimulation (8 s) at a medium stimulus intensity level on the skin site chosen for that run to habituate the skin site to heat stimulation. This habituation heat stimulation was not delivered by a virtual doctor. A jittered fixation preceded (5–9 s) and followed (3.1–7.1 s) the habituation heat stimulation. Next, during the doctor introduction (Fig. 1a), participants viewed a static picture of a doctor (3 s) who introduced themselves with a text bubble saying "Hi, I'm Dr. [Surname]. Let me take a moment to look over your chart." Surnames were randomly assigned to each doctor (see Supplementary Material for full list of surnames used in the study). After a jittered fixation cross (1–4 s), participants viewed the doctor holding a chart (4 s). Next, participants viewed a heat cue consisting of the same doctor now holding a Medoc thermode and a text bubble with "The heat will begin shortly" (4 s) (Fig. 1b). Participants then viewed a jittered prestimulus period with the doctor holding the thermode (5–9 s) to separate participants' pain anticipatory brain activity from heat pain-related activity. Participants then received a single suprathrehold heat stimulation (8 s) while viewing the doctor (Fig. 1c). The suprathreshold temperatures were in the range identified as above the pain threshold for most individuals (Campbell et al. 2005) and were 47 and 49 °C for the first 11 participants, and 46 and 48 °C for subsequent participants. Heat stimulation temperatures were lowered partway through data collection to increase the number of participants who were able to complete all heat stimulation trials during scanning. This did not compromise the study goals, as the primary comparisons (low- vs. high-trust doctor, stimulus intensity level) are within-person. However, to control for between-person differences in stimulus intensity levels, stimulus intensity level group was included as a covariate in linear mixed effects models predicting multivariate pain-predictive signatures and in univariate general linear model (GLM) analyses.

A jittered poststimulus period (3.1–7.1 s) followed the heat stimulation. Next, participants rated their pain intensity (0 = no pain to 10 = most intense pain imaginable) and pain unpleasantness (0 = no unpleasant pain to 10 = most unpleasant pain imaginable) to the preceding heat stimulation, with the order of the pain ratings randomized to control for order effects (Fig. 1d). Participants then received a second heat stimulation from the same doctor and completed pain ratings. An additional baseline fixation (20 s) ended each run.

Participants saw 6 different doctors for each of the 4 functional runs of the scan (24 doctors total viewed during the scan; 1 doctor was used only to train participants immediately prior to the fMRI scan). Half of the doctors in each run were high-trust, and half were low-trust. Each virtual doctor administered 2 heat stimulations: 1 at high stimulus intensity level and 1 at low stimulus intensity level. This resulted in participants receiving a total of 48 heat stimulations from the virtual doctors over 4 separate skin sites on the forearm, 12 heat stimulations at each combination of trust level and stimulus intensity level.

fMRI task training

On the day of the fMRI scan, participants first completed fMRI safety screening forms which were reviewed by an MRI technician to confirm scanning safety. The experimenter then drew 4 evenly spaced skin site boxes (1″ from wrist and inner elbow to avoid areas with maximum sensitivity) on the participant's forearm using a washable marker and trained the participant on how to complete the virtual medical interaction fMRI task. Participants were trained on how to make their intensity and unpleasantness ratings in response to the heat stimulations. Participants were carefully instructed on the difference between pain intensity and unpleasantness using previously developed language (Price et al. 1983). The virtual doctor presented to participants during fMRI task training was not shown during the scan. Pain ratings were completed using an fMRI-compatible trackball mouse. Participants were instructed to rate their pain within the pain rating period by moving the trackball mouse cursor to their rating on a vertical NRS and clicking the left mouse button to commit their rating. The mouse cursor started at 0 at the beginning of each rating.

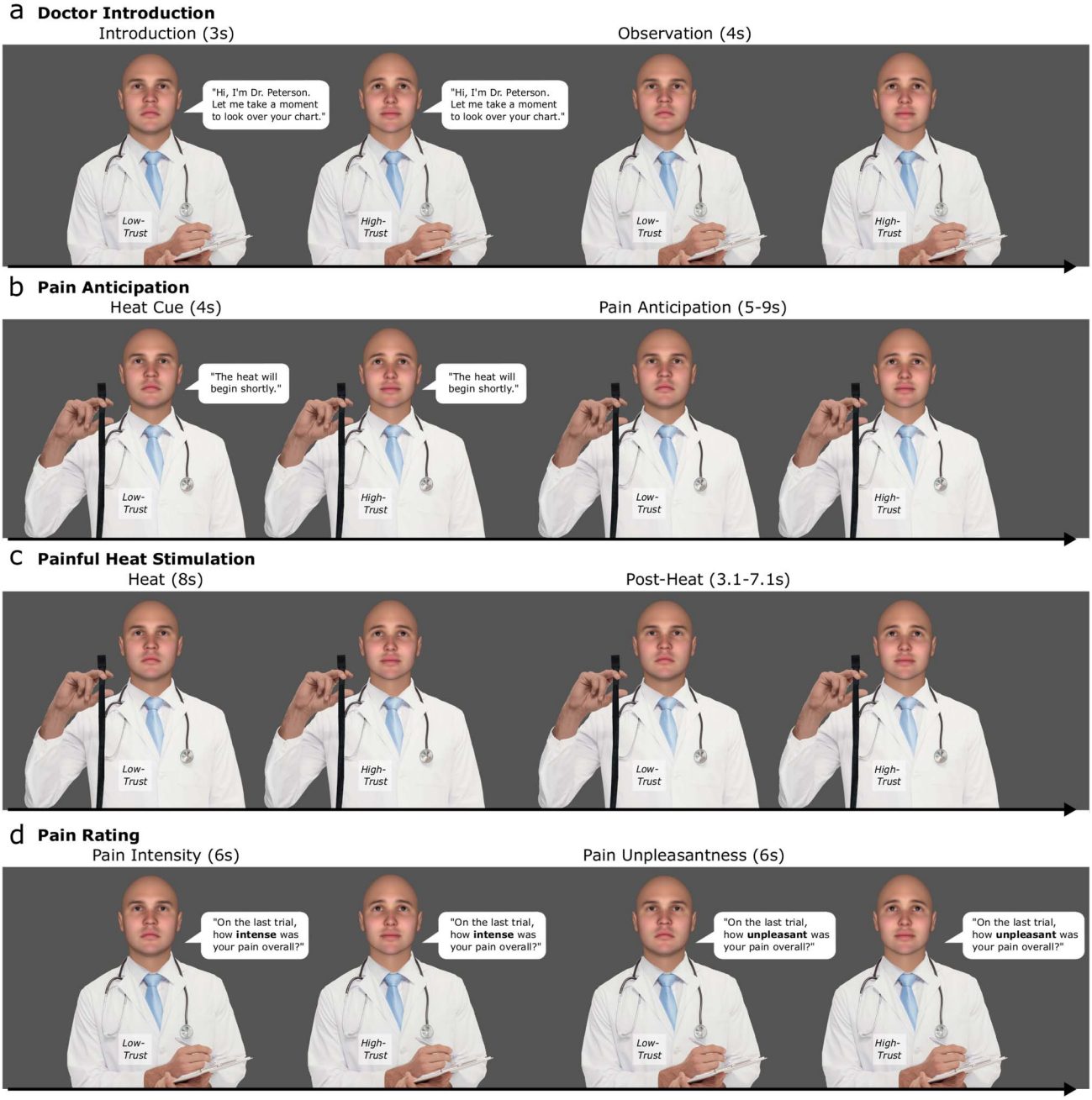

Fig. 1. Virtual medical interaction fMRI task design. Doctors depicted in the figure are of the same identity, although in the study, counterbalancing ensured that each participant saw each doctor identity at only one trust level. Note: Doctor trust levels were not communicated to participants at any point during the study (i.e. the experimental manipulation was implicit) and are included in the figure for illustration purposes only.

fMRI acquisition and preprocessing

Data acquisition

Data were collected on a GE Discovery MR750 3.0 T MR scanner at the Neuroimaging Facility at the University of Miami (Miami, FL, USA). Thirty-two axial slices covering the whole brain volume (cutting off high parietal when not possible to cover whole brain) were acquired, with a total of 485 volumes collected per functional run (TR = 1300 ms, TE = 25 ms, flip angle = 50°, FOV = 22 mm, slice thickness = 3.4 mm). A T_1-weighted anatomical image was acquired from each participant using a 3D BRAVO sequence (TI = 650 ms, sagittal orientation, flip angle = 12°, field of view = 256 mm, slice thickness = 1 mm).

Preprocessing

Data preprocessing and 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 (Jenkinson and Smith 2001; Jenkinson et al. 2002;

Greve and Fischl 2009), and brain-extraction. Preprocessing steps applied to functional data included brain extraction, motion correction using MCFLIRT (Jenkinson et al. 2002), smoothing using a full-width at half maximum (FWHM) 6-mm Gaussian smoothing kernel, removal of low frequency drift using a 90 s high-pass filter estimated from the data in FSL (Smith et al. 2004), and convolution with a double-gamma hemodynamic response function (HRF). 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 et al. 2001).

FSL's MCFLIRT (Jenkinson et al. 2002) was used for robust motion correction, which corrects for excessive head motion using rigid-body transformation across 6 standard motion parameters (rotations and translations along x, y, and z axes). In addition, a confound matrix was specified for each subject and functional run which regressed out timepoints corrupted by significant head motion. The DVARS metric was used for motion regression, which calculates the root mean square intensity difference between volume N and volume $N+1$. To further quantify excessive head motion, the mean framewise displacement (FD), or the average of rotation and translation parameter differences (using weighted scaling), was calculated for each subject (Power et al. 2012). No participants had a mean FD > 0.50 mm (Siegel et al. 2014). As a result, no participants were excluded from subsequent analyses due to excessive head motion.

Missing data

Partial fMRI task data were collected from participants for the following reasons: thermode error (2/4 runs lost for one participant), found heat intolerable (1/4 runs lost for one participant, 3/4 runs lost for another), chose not to continue scan (1/4 runs lost for one participant), thermode slipped off arm during scan (1/4 runs lost for one participant), wrong counterbalance order used (1/4 runs lost for one participant), and felt claustrophobic (1/4 runs lost for one participant, 2/4 runs lost for another). In total, 12/168 (7.14%) runs were dropped from 8/42 (19.05%) participants.

Behavioral data analysis

Manipulation checks

As a check of the doctor trust level manipulation, the effects of doctor trust level on participants' own ratings of doctor attributes were estimated using linear mixed effects models in R Version 3.3.2 (R Core Team 2019). We calculated P-values for each linear mixed effects model using Satterthwaite's degrees of freedom method in the *lmerTest* R package (Kuznetsova et al. 2015). Effect size (Cohen's d) was estimated for fixed effects in linear mixed effects models using the *EMAtools* R package (Kleiman 2017). We specified separate models for each rating type: trustworthiness, competence, warmth, attractiveness, and likelihood of choosing as own doctor. Each model was specified with crossed random effects for participant and doctor ID, which accounted for random variation between participants and between different doctor face identities (Mumford and Nichols 2006; Costafreda 2009). Additionally, each model included participant gender as a fixed effect due to evidence of gender differences in the perception of facial trustworthiness (Wincenciak et al. 2013). As a final manipulation check, we analyzed participant responses to questions assessing the realism of the virtual medical interactions and belief in the stated aims of the study ("How much did you believe in the stated purpose of the study").

Pain rating analysis

We used linear mixed effects models to estimate the effects of doctor trustworthiness on participants' pain ratings during the virtual medical interaction fMRI task. We specified 2 main models: 1 with pain intensity and 1 with pain unpleasantness as the dependent variable. Main models included crossed random effects for participant and doctor ID, as well as the following fixed effects included for statistical control: heat stimulus intensity level (high, low), the functional run of the scan (1–4), participant gender given known gender differences in pain rating (Greenspan et al. 2007) and perceptions of trustworthiness in others (Mattarozzi et al. 2015), and trial number to control for changes in participant pain ratings over the course of the task. We additionally tested whether participants' ratings of doctor trustworthiness predicted their pain ratings made during the fMRI virtual medical interaction task. To test this, we specified linear mixed effects models with participants' ratings of doctor trustworthiness made at home prior to arriving in the lab predicting their pain intensity and unpleasantness ratings (specified as separate dependent variables) during the virtual medical interaction. Fixed effects of no interest in our models were stimulus intensity level, functional run, participant gender, and trial number, while crossed random effects were participant ID and doctor ID.

fMRI data analysis

First level analysis

We conducted a first level univariate GLM analysis to identify brain activity related to pain received from a low- and high-trust doctor. We specified the first level GLM with the onsets and durations of the following events used to generate regressors: the cue period preceding the habituation heat stimulation, the preheat stimulus jitter period for the habituation heat stimulation, the habituation heat stimulation period (8 s), the doctor observation period in each condition, the cue period beginning each heat stimulation trial for each condition and stimulus intensity level, the jittered preheat stimulus period in which patients viewed the doctor (pain anticipation) for each condition and each stimulus intensity level, the jittered postheat stimulus period for each condition and each stimulus intensity level, the heat stimulation period

for each condition and each stimulus intensity level, and the pain intensity and pain unpleasantness rating periods (6 s each) for each condition and each stimulus intensity level. A fixation cross (20 s) at the beginning and end of each functional run was combined with the interstimulus interval fixation periods and used as the implicit baseline (unmodeled). In addition, we added a temporal derivative for each regressor in the GLM to correct for differences between the actual and modeled HRF.

Univariate GLM analysis

We conducted a whole-brain univariate GLM analysis to investigate trust-related differences in activity in brain regions previously implicated in pain, emotion, and face processing. Data were combined across subjects using a mixed-effects analysis (FLAME1) (Beckmann et al. 2003). Stimulus intensity level group was included as an additional covariate (orthogonalized with respect to the group mean) to statistically control for the subset of participants who received heat stimulations that were 1 °C higher than other participants. We subjected the resulting z statistic images to false discovery rate (FDR) correction with a critical threshold of $q < 0.05$, which controls for the expected proportion of false positives among suprathreshold voxels (Benjamini and Hochberg 1995; Genovese et al. 2002). The FDR threshold is determined from the observed *P*-value distribution, with the FDR *q* value representing the FDR-corrected *P*-value for the image. FDR correction at $q < 0.05$ means that, on average, 5% of the observed results will be false positives. Standard 2 × 2 ANOVA contrasts were calculated for low- versus high-trust doctors (pain anticipation, pain intensity rating, pain unpleasantness rating periods) and high versus low painful stimulus intensity level heat.

Multivariate pain-predictive signature analysis

To compare nociception-specific neural responses to painful heat administered by high-trust and low-trust doctors, we calculated the expression of a previously developed pain-predictive signature, the NPS (Wager et al. 2013) within our sample. The NPS consists of a pattern of positive and negative pain-predictive weights across distributed brain systems that are predictive of pain intensity ratings. The NPS was developed using a machine-learning–based regression technique, LASSO-PCR (least absolute shrinkage and selection operator-regularized principal components regression) (Wager et al. 2011). We calculated the degree to which each subject's whole-brain response pattern resembled the NPS pattern (pattern expression) in response to each doctor trust level (low-trust, high-trust) and during the relevant parts of the virtual medical interaction fMRI task (pain anticipation, painful heat, and pain rating) versus an implicit baseline.

NPS responses were calculated using Matlab by taking the dot product of the vectorized activation contrast beta map ($\vec{\beta}_{\text{map}}$) with the NPS pattern of voxel weights ($\vec{W}_{\text{map}}$), $\vec{\beta}^{T}_{\text{map}}\vec{W}_{\text{map}}$, yielding a continuous, scalar value representing each participant's NPS pattern response. We specified linear mixed effects models in R to test for differences in average NPS response due to our experimental conditions during each part of the fMRI task (pain anticipation, painful heat stimulation, and pain rating). NPS response during each part of the virtual medical interaction was the dependent variable and doctor trust level was the primary predictor. Stimulus intensity level group was included as a fixed effect to statistically control for the fact that a subset of participants (11/42) received low and high intensity heat stimulations that were 1 °C higher than other participants. Participant was included as a random effect with a random intercept to control for random variation between participants. Doctor ID was not included as a random effect in these models because NPS responses were averaged across all doctor identities in each doctor trust level.

Whole-brain correlation analysis

Finally, we conducted a whole-brain correlation analysis to investigate whether the response of any brain regions during different parts of the virtual medical interaction was correlated with participants' ratings of doctor trustworthiness and mistrust in healthcare organizations more generally. For participants' ratings of doctor trustworthiness, we calculated each participant's average rating of the trustworthiness of low-trust doctors, reverse-scored them such that higher values equaled more untrustworthiness, and then demeaned the values. Scores were then entered into a third-level mixed-effects GLM analysis in FSL. To examine whether mistrust of healthcare organizations in general was associated with whole-brain activity in similar brain regions involved in pain processing, we included participants' scores on the Medical Mistrust Index (MMI; LaVeist et al. 2009) in a separate whole-brain correlation analysis. Each participant's total MMI score was mean-centered and then entered in a third-level mixed-effects GLM analysis in FSL.

Bayes factor estimation of key null effects

We followed up key null findings using Bayes Factor estimation for one-sample t-tests of computed difference scores between doctor trust levels using the *ttestBF* function in the *BayesFactor* R package (v0.9.12). The *ttestBF* function utilizes the Jeffreys–Zellner–Siow prior with a scale factor = 0.707, which combines the Cauchy distribution on the standardized effect size and a noninformative Jeffreys prior on the variance of the normal population (Rouder et al. 2009).

Results

Manipulation checks

Consistent with prior findings (Todorov 2008; Todorov et al. 2008), and confirming the effectiveness of our trustworthiness manipulation, participants rated doctors lower on the perceived trustworthiness dimension

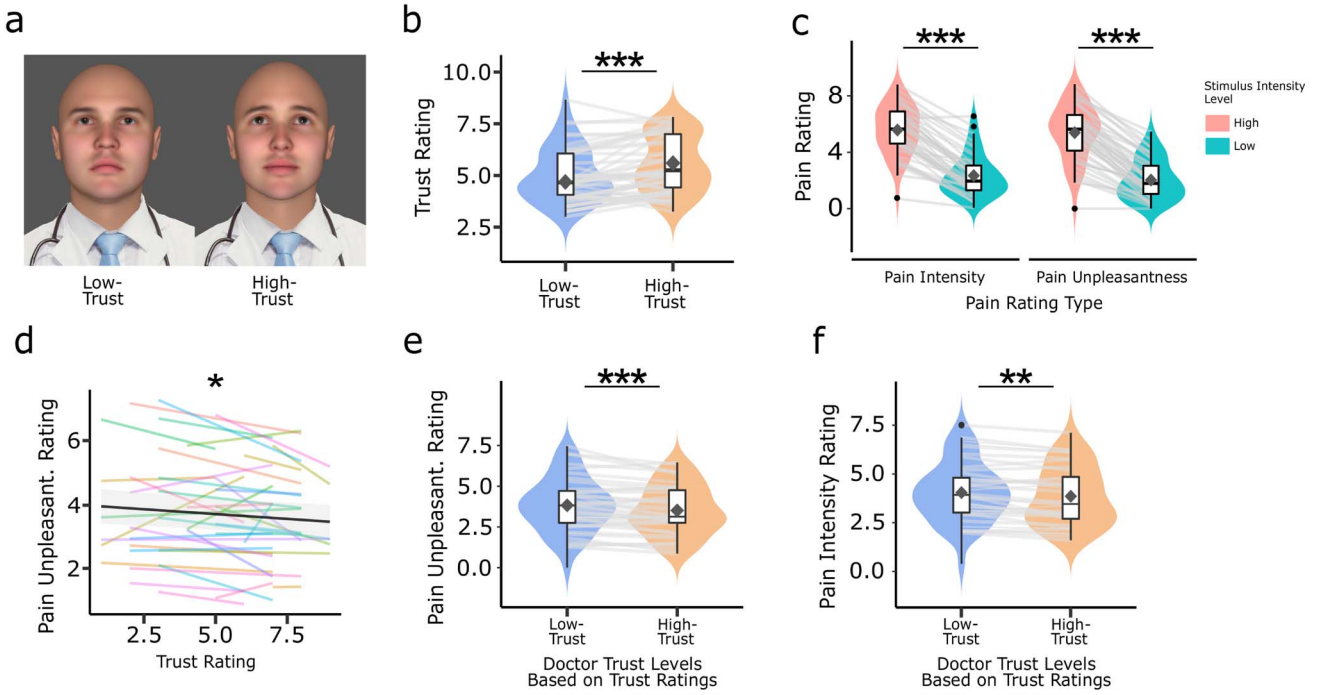

Fig. 2. Behavioral results. a) Doctor stimuli manipulated to be either high or low in facial trustworthiness (same identity); b) in manipulation checks, low-trust doctors rated as less trustworthy than high-trust doctors; c) pain intensity and pain unpleasantness ratings increased with the stimulus intensity level (temperature) of the painful heat stimulations during the virtual medical interaction; d) the less trustworthy participants perceived doctors to be (in ratings made prior to the scanning session), the more unpleasant they rated the pain associated with the painful medical procedure analogue from them during the scan; e) using doctor trust levels based on participants' own ratings of doctor trustworthiness, receiving painful heat stimulations from a low-trust versus high-trust doctor was associated with increased pain intensity and f) pain unpleasantness during the painful diagnostic procedure analogue. Results presented as combined box and violin plots of raw data, with black diamonds representing means, bars representing medians, lines representing within-subject changes in ratings due to doctor trustworthiness, and asterisks representing the results of significance tests from linear mixed effects models. Note: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

(low-trust; Fig. 2a left) as significantly less trustworthy than doctors higher on the perceived trustworthiness dimension (high-trust; Fig. 2a right), $B = -0.56$, SE = 0.07, $t_{942.66} = -7.63$, $P < 0.001$, $d = -0.50$ (Fig. 2b). Participants also rated low-trust doctors as significantly less competent ($B = -0.38$, SE = 0.08, $t_{942.76} = -4.69$, $P < 0.001$, $d = -0.31$), warm ($B = -2.54$, SE = 0.09, $t_{942.54} = -29.23$, $P < 0.001$, $d = -1.90$), and attractive ($B = -0.90$, SE = 0.08, $t_{942.47} = -11.20$, $P < 0.001$, $d = -0.73$) compared with high-trust doctors. Participants reported that they would be significantly less likely to choose low-trust doctors as their own doctor compared with high-trust doctors, $B = -1.15$, SE = 0.09, $t_{921.52} = -12.29$, $P < 0.001$, $d = -0.81$.

We examined whether participant gender influenced participant ratings given known gender differences in perceptions of trustworthiness in others (Mattarozzi et al. 2015). Participant gender did not influence participant ratings of doctor trustworthiness ($P = 0.411$, $d = -0.26$), competence ($P = 0.879$, $d = 0.05$), warmth ($P = 0.108$, $d = -0.52$), attractiveness ($P = 0.065$, $d = -0.60$), or preference as own doctor ($P = 0.077$, $d = -0.57$). Poststudy surveys indicated that participants found the study rationale believable ($M = 71.05$, $SD = 28.47$; 0 = not at all to 100 = completely) and the virtual medical interactions moderately realistic ($M = 43.21$, $SD = 25.52$).

Finally, as expected, participants' pain intensity (*low stimulus intensity:* $M = 2.30$, $SD = 1.80$; *high stimulus intensity:* $M = 5.58$, $SD = 2.12$; $B = 3.28$, SE = 0.06, $t_{1803.19} = 51.45$, $P < 0.001$, $d = 2.42$) and pain unpleasantness (*low stimulus intensity:* $M = 2.00$, $SD = 1.77$; *high stimulus intensity:* $M = 5.41$, $SD = 2.32$; $B = 3.40$, SE = 0.06, $t_{1803.19} = 53.43$, $P < 0.001$, $d = 2.52$) ratings significantly increased with the stimulus intensity level (temperature) of the heat stimulations (Fig. 2c).

Behavioral results

Increased pain ratings during painful diagnostic procedure analogue administered by low-trust doctors

The less trustworthy participants perceived doctors to be (in ratings made prior to the scanning session), the more unpleasant they rated the pain associated with the painful medical procedure analogue from them during the scan, $B = 0.09$, $SE = 0.03$, $t_{1494.19} = -2.32$, $P = 0.021$, $d = -0.12$; Fig. 2c. There was a nonsignificant trend in the same direction as the unpleasantness effect for pain intensity ratings, $B = -0.04$, $SE = 0.03$, $t_{1165.02} = -1.67$, $P = 0.096$, $d = -0.10$. Examining the relationship between doctor trust level (low-trust, high-trust) and pain ratings, doctor trust level did not influence participants reported pain intensity ($B = 0.04$, SE = 0.06, $t_{1811.81} = 0.58$, $P = 0.564$, $d = 0.03$) or pain unpleasantness ($B = 0.09$, $SE = 0.06$, $t_{1808.55} = 1.39$, $P = 0.165$, $d = 0.07$) during the scan. However, the mean pain unpleasantness rating

was higher with a low-trust ($M = 3.76$, $SD = 1.66$) versus a high-trust doctor ($M = 3.66$, $SD = 1.60$), consistent with the trust-related reductions in pain unpleasantness seen for participants' own evaluations of doctor trustworthiness. Follow-up Bayes Factor estimation for this test provided moderate evidence in support of the null hypothesis of no doctor trust level effect on pain intensity ($BF_{10} = 0.19$), while the Bayes Factor for unpleasantness ratings ($BF_{10} = 0.95$) did not provide strong support for the null or alternative hypothesis.

To account for individual differences in participants' perceptions of the trustworthiness of the doctor stimuli, we next conducted a post-hoc analysis in which we based the doctor trust levels on participants' own ratings of doctor trustworthiness. We recoded each doctor face stimulus as low-trust if the participant's own trustworthiness rating for that doctor was lower than the within-participant median rating, and high-trust if the participant's own trustworthiness rating for that doctor was higher than the within-participant median rating. We then tested the recoded doctor trust levels in our linear mixed effects models predicting reported pain. When using the doctor trust levels based on participants' own ratings of doctor trustworthiness, receiving painful heat stimulations from a low-trust ($M = 4.05$, $SD = 1.55$) versus high-trust ($M = 3.86$, $SD = 1.52$) doctor was associated with increased pain intensity, $B = 0.23$, $SE = 0.08$, $t_{1274.24} = 3.06$, $P = 0.002$, $d = 0.17$ (Fig. 2d, left). Similarly, receiving painful heat stimulations from a low-trust ($M = 3.83$, $SD = 1.69$) versus high-trust ($M = 3.51$, $SD = 1.55$) doctor was associated with increased pain unpleasantness during the painful diagnostic procedure analogue, $B = 0.32$, $SE = 0.08$, $t_{1560.62} = 4.09$, $P < 0.001$, $d = 0.21$ (Fig. 2d, right). These results provide a test of the impacts of trustworthiness on reported pain that takes into account each participant's actual perceived trustworthiness of the doctor face stimuli. In order to preserve the originally intended balanced fMRI task design, neuroimaging analyses were conducted with the original trust levels in the face stimulus set (Oosterhof and Todorov 2008). Given the implicit nature of perceived trustworthiness in the face stimuli, we believe that conducting the fMRI analyses with their original trust level coding is still a valid test of our hypotheses on the impacts of perceived trustworthiness on the neural correlates of pain.

Finally, we examined whether participant gender influenced ratings given known gender differences in pain rating (Greenspan et al. 2007). Pain ratings did not differ by gender for any models predicting participant pain ratings (all $P > 0.28$, Cohen's $d < 0.35$). Altogether, these findings suggest that lower perceived doctor trustworthiness was associated with increased reported pain during the fMRI painful diagnostic procedure analogue, an effect that was strongest when taking into consideration participants' own ratings of the trustworthiness of the doctor stimuli.

Increased activation in pain-related brain regions during painful diagnostic procedure analogue administered by low-trust doctors

As expected, we found that activation in pain-related brain regions (FDR $q < 0.05$) increased with the stimulus intensity level (temperature) of the heat stimulations (HighHeat > LowHeat). Increasing painful stimulus intensity increased activity in the primary somatosensory cortex (S1), dorsal posterior insula (dpINS), anterior insula (aINS), thalamus, dorsal anterior cingulate cortex (dACC), and supplementary motor area (SMA), and decreased activity in the ventromedial prefrontal cortex (vmPFC), consistent with prior studies using thermal heat stimulations (Wager et al. 2013; Atlas et al. 2014).

Next, we examined the effects of doctor trust level on brain activity during the painful diagnostic procedure analogue. First, we found that anticipating the painful procedure from a low-trust compared with a high-trust doctor ($\text{Low-Trust}_{\text{Anticipation}} > \text{High-Trust}_{\text{Anticipation}}$) was associated with increased activity in the posterior cingulate cortex, precuneus, right insular cortex, and lingual gyrus, consistent with prior findings of expectations of increased pain and nocebo effects (Watson et al. 2009; Schmid et al. 2013; Brown et al. 2014).

Receiving high stimulus intensity heat (48 or 49 °C) from a low- versus high-trust doctor ($\text{Low-Trust}_{\text{HighHeat}} > \text{High-Trust}_{\text{HighHeat}}$) was associated with increased activity in several brain regions previously implicated in somatic pain from thermal stimulations. These included increases in activity in several lateral somatosensory and medial regions (SMA, aINS, ventral insula) and regions associated with emotion and motivated action (SMA, aINS) (Yarkoni et al. 2011; Krishnan et al. 2016) (Fig. 3a). Receiving low stimulus intensity heat (46 or 47 °C) from a low- versus high-trust doctor ($\text{Low-Trust}_{\text{LowHeat}} > \text{High-Trust}_{\text{LowHeat}}$) was associated with increased activity in the lateral occipital cortex and middle temporal gyrus.

We next examined whole-brain activity during the pain rating periods following each heat stimulation in order to assess brain activity related to pain report. We found increased activity in several regions associated with somatic pain. There was increased activity in the right mid-insula and superior temporal gyrus when participants rated pain intensity from a low- versus high-trust doctor ($\text{Low-Trust}_{\text{IntenRate}} > \text{High-Trust}_{\text{IntenRate}}$) (Fig. 3b). There was increased activity in the dpINS, central operculum, fusiform gyrus, and primary somatosensory cortex when participants rated pain unpleasantness from a low- versus high-trust doctor ($\text{Low-Trust}_{\text{UnplRate}} > \text{High-Trust}_{\text{UnplRate}}$) (Fig. 3c). Altogether, these findings indicate that lower doctor trustworthiness was associated with increased activity within multiple brain systems implicated in nociception and other aspects of pain perception during the painful diagnostic procedure analogue.

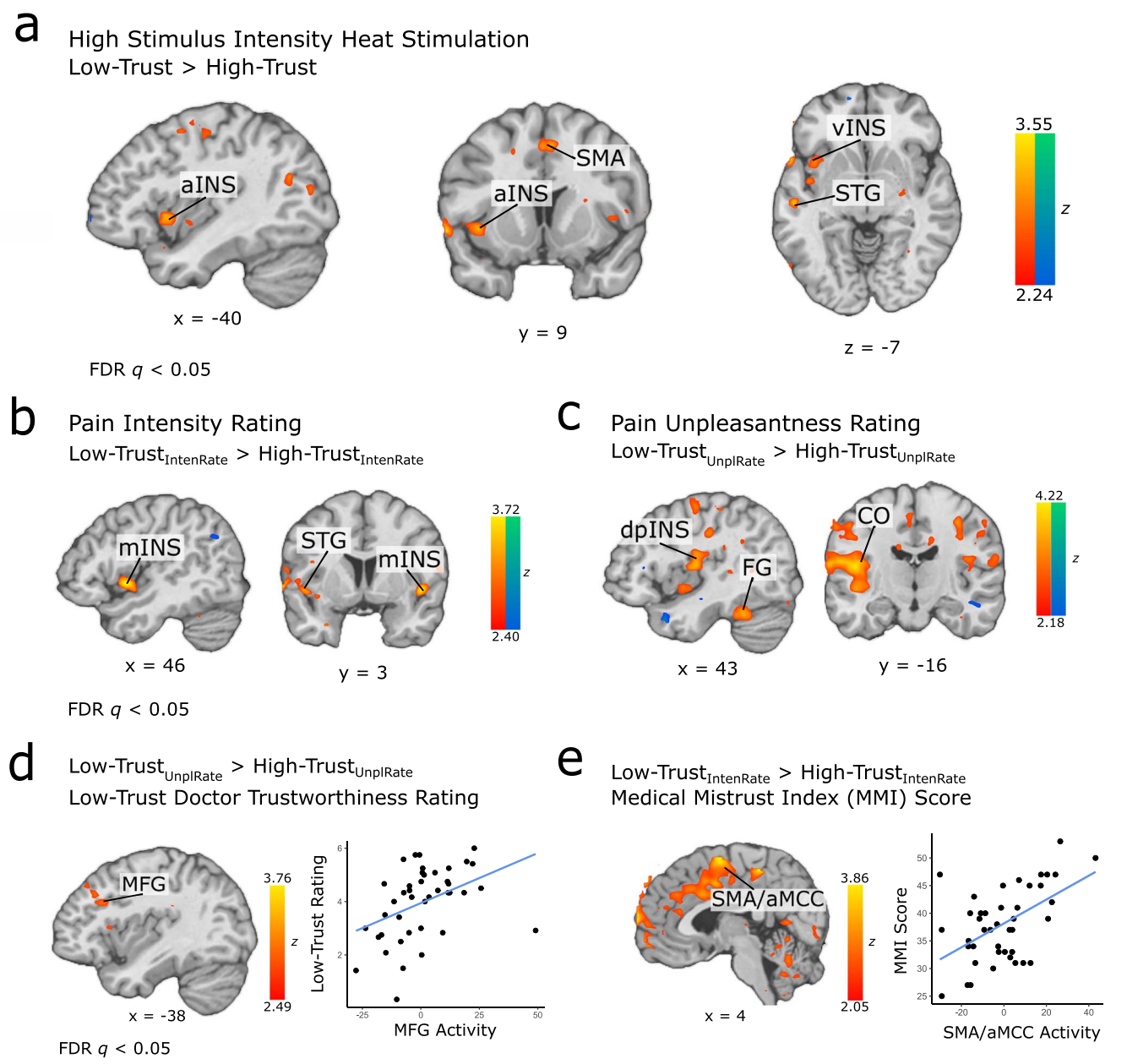

Fig. 3. Univariate GLM results. a) Brain activity during high stimulus intensity heat stimulation from a low- versus high-trust doctor; b) brain activity during pain intensity rating from a low- versus high-trust doctor; c) brain activity during pain unpleasantness rating from a low- versus high-trust doctor; d) the more untrustworthy participants rated low-trust doctors on average, the more they demonstrated increased activity in the MFG (extracted time series plotted against trustworthiness ratings for display purposes) when rating their pain unpleasantness from low- vs. high-trust doctors; e) the more mistrust in healthcare organizations participants reported (MMI score), the more participants demonstrated activity in the SMA extending into the aMCC (extracted time series plotted against MMI scores for display purposes) when rating their pain intensity from low- versus high-trust doctors. Abbreviations: dpINS = dorsal posterior insula, aINS = anterior insula, vINS = ventral insula, mINS = mid-insula, MFG = middle frontal gyrus, aMCC = anterior mid-cingulate cortex, CO = central operculum, STG = superior temporal gyrus, FG = fusiform gyrus.

Increased NPS responses during painful diagnostic procedure analogue administered by low-trust doctors

We next conducted a more sensitive test (compared with the whole-brain analysis) of whether pain-specific brain activity is impacted by doctor trustworthiness during a painful diagnostic procedure analogue. We did this by testing whether a previously developed pain-predictive neural signature; the NPS (Fig. 4a) was modulated by doctor trustworthiness during fMRI. As expected, the NPS robustly responded during low and high stimulus intensity heat stimulation (high stimulus intensity: high-trust condition: $t = 15.36$; low-trust condition: $t = 17.44$, P-values < 0.001; low stimulus intensity: high-trust condition: $t = 9.03$; low-trust condition: $t = 8.27$, P-values < 0.001). NPS responses significantly increased with the stimulus intensity level of the heat stimulation ($B = 49.45$, $SE = 3.10$, $t_{125} = 15.97$, $P < 0.001$, $d = 2.86$; Fig. 4b) and increased along with participants' reported pain intensity ($B = 12.77$, $SE = 0.89$, $t_{141.37} = 14.27$, $P < 0.001$, $d = 2.40$; Fig. 4c) and unpleasantness ($B = 11.97$, $SE = 0.90$,

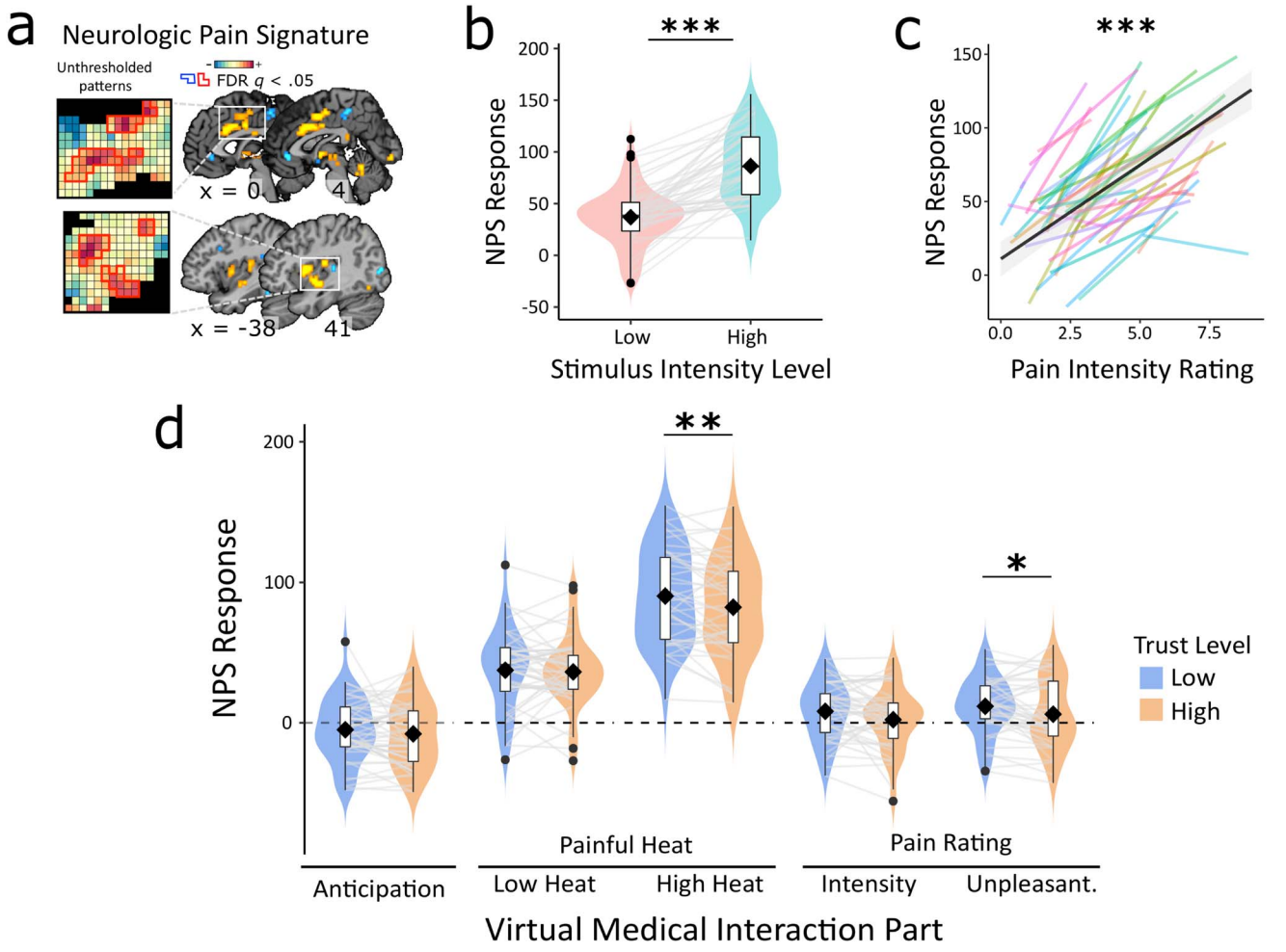

Fig. 4. Multivariate pain-predictive signature results. a) NPS, image modified and reproduced from Wager et al. (2013); b) NPS responses to low and high stimulus intensity heat stimulation during the painful diagnostic procedure analogue; c) NPS responses increased with participant pain intensity rating; d) NPS responses to low- and high- trust doctors during the virtual medical interaction fMRI task; results presented as combined box and violin plots, with black diamonds representing means, bars representing medians, lines representing within-subject changes in NPS responses due to doctor trustworthiness, and asterisks representing the results of significance tests from linear mixed effects models. Note: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

$t_{143.98} = 13.31$, $P < 0.001$, $d = 2.22$), consistent with prior studies (Wager et al. 2013; Han et al. 2022).

Next, we examined NPS responses during the virtual medical interaction fMRI task. Consistent with our hypothesis, we found that receiving the painful diagnostic procedure analogue from low- versus high-trust doctors significantly increased NPS responses during high stimulus intensity heat stimulation, $B = 7.95$, $SE = 2.85$, $t_{41} = 2.79$, $P = 0.008$, $d = 0.87$ (Fig. 4d, Table 2). In addition, NPS responses were significantly higher during pain unpleasantness rating of low- compared with high-trust doctors, $B = 5.60$, $SE = 2.63$, $t_{41} = 2.13$, $P = 0.039$, $d = 0.67$. In contrast, doctor trustworthiness marginally influenced NPS responses during pain intensity rating [$B = 5.97$, $SE = 3.02$, $t_{41} = 1.98$, $P = 0.055$, $d = 0.62$] and did not influence NPS responses during pain anticipation [$B = 2.94$, $SE = 3.12$, $t_{41} = 0.94$, $P = 0.352$, $d = 0.29$] or low stimulus intensity heat stimulation [$B = 1.19$, $SE = 3.50$, $t_{41} = 0.34$, $P = 0.736$, $d = 0.11$]. Follow-up Bayes Factor estimation provided moderate to anecdotal evidence in support of no effect of doctor trustworthiness on NPS responses during pain anticipation ($BF_{10} = 0.25$), low stimulus intensity heat stimulation ($BF_{10} = 0.18$), and pain intensity rating ($BF_{10} = 0.98$).

Altogether, these findings suggest that the NPS robustly increased with the stimulus intensity level of the heat stimulation and participants' pain intensity ratings. Receiving painful heat stimulation from low- versus high-trust doctors was associated with increased NPS responses during high stimulus intensity pain experience and pain unpleasantness recall, consistent with the results of the whole-brain analyses.

Increased brain responses during pain experience and pain reporting associated with lower doctor trustworthiness and higher medical mistrust

To better understand the psychosocial variables contributing to the increased pain-related brain responses we observed with low-trust doctors, we examined whether participants' trustworthiness ratings of low-trust doctors and mistrust in healthcare organizations, more generally, was also associated with increased

Table 2. Results of linear mixed effects model predicting NPS responses to doctor trust levels during high stimulus intensity heat stimulation.

	NPS Response		
Predictors	*Estimate*	*Standard Error*	*P*
(Intercept)	77.18	10.09	**<0.001**
Doctor Trust Level [Low-Trust]	7.95	2.85	**0.008**
Stimulus Intensity Level Group [Low]	6.97	11.63	0.553
Random Effects			
σ^2	170.55		
$\tau_{00\ Subject}$	1012.69		
ICC	0.86		
$N_{Subject}$	42		
Observations	84		
Marginal R^2/Conditional R^2	0.021/0.859		

Note. Brackets denote reference level of categorical predictors. The variable "Stimulus Intensity Level Group" is a categorical variable referring to whether participants received 49 and 47 °C heat stimulations or 48 and 46 °C heat stimulations.

neural responses to painful heat using a whole-brain analysis. We found that the more untrustworthy participants rated low-trust doctors on average, the more they demonstrated increased activity in the left middle frontal gyrus (MFG), left frontal pole (FP), and right inferior frontal gyrus when rating pain unpleasantness from low- versus high-trust doctors (Fig. 3d). In addition, the more untrustworthy participants rated low-trust doctors on average, the more they demonstrated increased activity in the left frontal opercular cortex when rating pain intensity from low- versus high-trust doctors.

Next, examining mistrust in healthcare organizations, we found that the more mistrust in healthcare organizations participants reported, the more they demonstrated increased activity in the SMA extending into the anterior midcingulate cortex (aMCC), right postcentral gyrus, right lateral occipital cortex, ACC, bilateral parietal opercular cortex, precuneus, cerebellum, and left dorsal posterior/anterior insula when rating pain intensity from low- versus high-trust doctors (Fig. 3e). Similarly, the more mistrust in healthcare organizations participants reported, the more participants demonstrated increased activity in the FP, frontal orbital cortex, MFG, superior frontal gyrus, cerebellum, and fusiform gyrus when rating pain unpleasantness from low- versus high-trust doctors. Examining brain activity during painful heat stimulation, mistrust in healthcare organizations was associated with increased activity in the lateral occipital cortex and fusiform gyrus during high stimulus intensity heat delivered from low- versus high-trust doctors. Altogether, these results suggest that less trust in doctors and healthcare organizations in general is associated with increased activity in brain regions involved in pain, attention, and emotion both when experiencing and evaluating pain.

Discussion

Summary of findings

Patient trust in their doctor is an important component of the doctor–patient relationship and has been shown to positively influence patient health outcomes. Despite the known benefits of trust in the doctor–patient relationship, the neurobiological mechanisms underlying the effects of trust on patient health outcomes remain unclear. We previously found that patient feelings of trust toward their clinician predicted reductions in reported pain (Losin et al. 2017) using face-to-face medical simulations and surveys about participants' core beliefs and values to manipulate feelings of cultural similarity and trust. In the present study, we replicated trust-related reductions in pain using a different trustworthiness manipulation: changes in the facial features of virtual doctors. We found that receiving painful heat stimulation from low- versus high-trust doctors was associated with increases in activity in brain regions involved with the sensory discriminative and affective-motivational aspect of pain and increased expression of a multivariate pain-predictive neural signature, the NPS (Wager et al. 2013), during both pain experience and pain reporting. We further found that lower levels of trust in the virtual doctors and mistrust in healthcare organizations, in general, were associated with increased brain activity in regions involved in pain, emotion, and attention during pain experience and reporting.

Doctor trustworthiness influenced sensory and affective correlates of pain

Across our analyses, we found evidence that doctor trustworthiness influenced the sensory-discriminative and affective-motivational dimensions of pain during both pain experience and reporting. Evidence of doctor trustworthiness influencing the sensory-discriminative aspect of pain includes our behavioral finding that participants rated their pain intensity higher in response to painful heat stimulations from low-trust doctors when the doctor trust levels were based on participants' own ratings of doctor trustworthiness. In our whole-brain results, we found that receiving high stimulus intensity heat stimulation from a low-trust doctor increased activity in several brain regions implicated in nociceptive pain, including the anterior insula (aINS), ventral insula (vINS),

and SMA. We also found that receiving painful heat stimulation from a low-trust doctor increased responses of the multivariate NPS pattern, which is sensitive and specific to nociceptive pain (Wager et al. 2013). In a whole-brain correlation analysis, we found that participants' mistrust in healthcare organizations was associated with increased activity in regions involved in pain, attention, and emotion when participants rated their pain intensity from low- versus high-trust doctors.

Evidence of our doctor trustworthiness manipulation modulating the affective-motivational aspects of pain include our behavioral finding that participants' ratings of the trustworthiness of the virtual doctors made prior to the scan predicted their reported pain unpleasantness, which is thought to reflect the affective-motivational aspect of pain (Price et al. 1983). Participants also rated their pain unpleasantness higher in response to painful heat stimulations from low-trust doctors when the doctor trust levels were based on participants' own ratings of doctor trustworthiness. In our whole-brain analysis, we found increases in activity in several brain regions associated with pain during both pain intensity and pain unpleasantness rating periods, as well as a significant increase in NPS responses to low-trust doctors during pain unpleasantness rating. These findings are consistent with other clinical and experimental studies finding that social, contextual, and interpersonal factors can modulate reported pain (Master et al. 2009; Eisenberger et al. 2011; López-Solà et al. 2019; Ashton-James et al. 2021) and suggest that interventions aimed at increasing doctor trustworthiness may have broad impacts on both the sensory-discriminative and affective-motivational dimensions of pain experience and reporting.

Doctor facial trustworthiness sufficient to influence pain

A novel aspect of our experimental manipulation, which was based solely on facial features, is that it was never made explicit to participants and did not involve any modification of doctor speech or behavior, in contrast to most prior studies of trustworthiness in medical contexts (Baker et al. 2003; Andreassen et al. 2006; Coran et al. 2013). Supporting the implicit nature of our manipulation, ratings made by participants at the conclusion of their participation in the study suggested that although they had some idea trustworthiness might be under study, it was clear that they did not perceive trustworthiness to be the study's primary focus. Our results add to our understanding of which characteristics of medical providers and medical interactions may have meaningful impacts on clinical outcomes such as pain. For example, we have previously demonstrated that factors relevant to the clinician, such as perceived similarity and trustworthiness (Losin et al. 2017), race and ethnicity (Anderson et al. 2020), and movement synchrony (Goldstein et al. 2020) can influence patient reported pain and its physiological correlates. Consistent with psychological findings on the rapidity of observer judgments based on facial appearances alone (Willis and Todorov 2006; Rule et al. 2013), our findings suggest that patient judgments of the trustworthiness of their doctor may be made spontaneously and potentially independent of behavioral markers of trustworthiness. As competing demands on doctor time restrict the length of medical visits (Sinsky et al. 2016), and the continuity of care with a single doctor decreases (Levene et al. 2018), patients' impressions formed on the basis of superficial characteristics of their doctor may have important consequences for the quality of the medical interaction and the doctor–patient relationship.

Limitations

Several limitations of the present study are worth noting. First, our use of healthy participants playing the role of patients in the virtual medical interactions limits the generalizability of our findings to clinical pain populations. Given known differences between patients with chronic pain and healthy individuals in central nociceptive processing (Baliki et al. 2008; Vachon-Presseau et al. 2016; Martucci and Mackey 2018), further study is needed to investigate whether the effects of trustworthiness on pain that we report here generalize to patients with chronic pain. However, although our study did not involve actual patients or doctors, our use of virtual medical interactions provided a high degree of experimental control and allowed us to examine trust-related changes in pain using functional neuroimaging. Our use of computer-generated face stimuli that varied in perceived facial trustworthiness, rather than actual doctors, enabled us to experimentally manipulate perceived trustworthiness and examine more implicit effects of trust-related changes in pain which are independent of doctor speech or behavior, providing a novel experimental contribution to the literature on doctor–patient trust. However, the relationship between the results of the present study and semantic or behavioral medical interventions, which are more commonly employed in clinical research settings, is a topic that should be examined in future studies. Second, because only white-appearing face stimuli were available for use in the present study, our findings on doctor trustworthiness may not generalize to more diverse contexts. The faces used for our doctor stimuli were also male and thus may not generalize to perceived trustworthiness in female doctors. Our findings using only male faces must also be considered within the context of gender discordance, which prior evidence suggests may influence reported pain (Aslaksen et al. 2007). Finally, as prior studies have found a strong influence of the evaluative context on subsequent face judgments (DeBruine 2005; Todorov et al. 2005), it is likely that the medical context presentation in our study strongly influenced participant brain responses to the stimuli. We did not assess brain activity to the face stimuli separate from the context of the doctor body, limiting our ability to speak to the prior literature on

the neural correlates of facial trustworthiness. However, our choice to present the trustworthiness face stimuli to participants only within the medical context is likely to have increased the overall realism and generalizability of our study. Our use of fMRI also allowed us to examine the neurobiological mechanisms of our previous finding of trust-related reductions in reported pain (Losin et al. 2017).

Conclusion

In conclusion, our findings suggest that patient trust in their doctor may have tangible impacts in reducing both sensory-discriminative and affective-motivational aspects of pain experience and reporting and point to a potential brain basis for trust-related reductions in pain via modulation of central nociceptive mechanisms. Future studies should aim to investigate how other psychological and neurobiological factors known to influence pain (e.g. anxiety, placebo analgesia) influence the present study's findings of trust-related changes in pain. For example, our finding that perceived doctor trustworthiness influenced NPS responses should be considered in the context of findings that placebo treatments generally have only small effects on the NPS (Zunhammer et al. 2018). This suggests that doctor trustworthiness as manipulated in our study may be acting through different brain mechanisms than placebo analgesia.

The findings in the present study have several implications and potential applications for real-world medical care. First, although we used computationally derived face stimuli (Oosterhof and Todorov 2008; Todorov 2011) in our study to manipulate the perceived trustworthiness of virtual doctors, we do not suggest that altering doctor facial expressions associated with trustworthiness should be an intervention for improving real-world doctor–patient relationships. Rather, our results suggest that even small and implicitly perceived changes in trustworthiness in the doctor–patient interaction may improve patient health. The results of the present study, when combined with our prior work in which we manipulated doctor trustworthiness via perceived doctor–patient similarity (Losin et al. 2017), demonstrates that perceived trustworthiness in doctor–patient interactions may be manipulated via multiple pathways, with effects on both reported pain and its neural correlates, and may point to novel interventions for improving patient care in the future.

Open practices statement

The data used for the present study are available upon reasonable request. The virtual doctor stimulus set, along with an R Markdown file of analyses conducted, is available on OSF: https://osf.io/gckhf/?view_only=3754723f69ec48d3a712da6d5384f14b.

Acknowledgments

The authors wish to thank Mary Christodoulou for her assistance with data analysis.

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

Elizabeth A. Reynolds Losin was supported by a Mentored Research Scientist Development award from National Institute On Drug Abuse of the National Institutes of Health (K01DA045735). This work was also supported by the University of Miami College of Arts and Sciences institutional startup funds.

Conflict of interest statement: None declared.

References

Anderson SR, Losin EAR. A sociocultural neuroscience approach to pain. *Cult Brain*. 2017:5(1):14–35.

Anderson SR, Gianola M, Perry JM, Losin EAR. Clinician-patient racial/ethnic concordance influences racial/ethnic minority pain: evidence from simulated clinical interactions. *Pain Med*. 2020:21(11):3109–3125.

Andreassen HK, Trondsen M, Kummervold PE, Gammon D, Hjortdahl P. Patients who use e-mediated communication with their doctor: new constructions of trust in the patient-doctor relationship. *Qual Health Res*. 2006:16(2):238–248.

Apkarian AV, Bushnell MC, Treede R-D, Zubieta J-K. Human brain mechanisms of pain perception and regulation in health and disease. *Eur J Pain*. 2005:9(4):463–484.

Ashton-James CE, Anderson SR, Mackey SC, Darnall BD. Beyond pain, distress, and disability: the importance of social outcomes in pain management research and practice. *Pain*. 2021:163(3):e426–e431.

Aslaksen PM, Myrbakk IN, Hoifodt RS, Flaten MA. The effect of experimenter gender on autonomic and subjective responses to pain stimuli. *Pain*. 2007:129(3):260–268.

Atlas LY, Bolger N, Lindquist MA, Wager TD. Brain mediators of predictive cue effects on perceived pain. *J Neurosci*. 2010:30(39):12964–12977.

Atlas LY, Lindquist MA, Bolger N, Wager TD. Brain mediators of the effects of noxious heat on pain. *Pain*. 2014:155(8):1632–1648.

Baker R, Mainous Iii AG, Gray DP, Love MM. Exploration of the relationship between continuity, trust in regular doctors and patient satisfaction with consultations with family doctors. *Scand J Prim Health Care*. 2003:21(1):27–32.

Baliki MN, Geha PY, Apkarian AV, Chialvo DR. Beyond feeling: chronic pain hurts the brain, disrupting the default-mode network dynamics. *J Neurosci*. 2008:28(6):1398–1403.

Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in FMRI. *NeuroImage*. 2003:20(2):1052–1063.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Methodol*. 1995:57(1):289–300.

Brown CA, El-Deredy W, Jones AK. When the brain expects pain: common neural responses to pain anticipation are related to

clinical pain and distress in fibromyalgia and osteoarthritis. *Eur J Neurosci*. 2014:39(4):663–672.

Buhle J, Wager TD. Does meditation training lead to enduring changes in the anticipation and experience of pain? *Pain*. 2010:150(3):382–383.

Campbell CM, Edwards RR, Fillingim RB. Ethnic differences in responses to multiple experimental pain stimuli. *Pain*. 2005: 113(1):20–26.

Chang LJ, Gianaros PJ, Manuck SB, Krishnan A, Wager TD. A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biol*. 2015:13(6):e1002180.

Coran JJ, Koropeckyj-Cox T, Arnold CL. Are physicians and patients in agreement? Exploring dyadic concordance. *Health Educ Behav*. 2013:40(5):603–611.

Costafreda SG. Pooling FMRI data: meta-analysis, mega-analysis and multi-center studies. *Front Neuroinform*. 2009:3(33):1–8.

Davis KD. The neural circuitry of pain as explored with functional MRI. *Neurol Res*. 2000:22(3):313–317.

DeBruine LM. Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *Proc R Soc B Biol Sci*. 2005:272(1566): 919–922.

Edwards RR, Fillingim RB. Ethnic differences in thermal pain responses. *Psychosom Med*. 1999:61(3):346–354.

Eisenberger NI, Master SL, Inagaki TK, Taylor SE, Shirinyan D, Lieberman MD, Naliboff BD. Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proc Natl Acad Sci*. 2011:108(28):11721–11726.

Ferreira PH, Ferreira ML, Maher CG, Refshauge KM, Latimer J, Adams RD. The therapeutic alliance between clinicians and patients predicts outcome in chronic low back pain. *Phys Ther*. 2013:93(4): 470–478.

Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*. 2002:15(4):870–878.

Gianola M, Llabre MM, Losin EAR. Effects of language context and cultural identity on the pain experience of Spanish–English bilinguals. *Affect Sci*. 2021:2(2):112–127.

Gittell JH, Fairfield KM, Bierbaum B, Head W, Jackson R, Kelly M, Laskin R, Lipson S, Siliski J, Thornhill T. Impact of relational coordination on quality of care, postoperative pain and functioning, and length of stay: a nine-hospital study of surgical patients. *Med Care*. 2000:38(8):807–819.

Goldstein P, Losin EAR, Anderson SR, Schelkun VR, Wager TD. Clinician-patient movement synchrony mediates social group effects on interpersonal trust and perceived pain. *J Pain*. 2020: 21(11–12):1160–1174.

Greenspan JD, Craft RM, LeResche L, Arendt-Nielsen L, Berkley KJ, Fillingim RB, Gold MS, Holdcroft A, Lautenbacher S, Mayer EA. Studying sex and gender differences in pain and analgesia: a consensus report. *Pain*. 2007:132(Supplement 1):S26–S45.

Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*. 2009:48(1):1095–9572.

Hall MA, Dugan E, Zheng B, Mishra AK. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? *Milbank Q*. 2001:79(4):613–639.

Han X, Ashar YK, Kragel P, Petre B, Schelkun V, Atlas L, Chang LJ, Jepma M, Koban L, Losin ER, Roy M, Woo CW, Wager TD. Effect sizes and test-retest reliability of the fMRI-based neurologic pain signature. *NeuroImage* 2022:247(118844):1–14.

Hassin R, Trope Y. Facing faces: studies on the cognitive aspects of physiognomy. *J Pers Soc Psychol*. 2000:78(5):837–852.

Howe LC, Goyer JP, Crum AJ. Harnessing the placebo effect: exploring the influence of physician characteristics on placebo response. *Health Psychol*. 2017:36(11):1074–1082.

IOM. *Relieving pain in America: a blueprint for transforming prevention, care, education and research*. Washington (DC): The National Academies Press; 2011.

Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*. 2001:5(2):143–156.

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*. 2002:17(2):825–841.

Kaptchuk TJ, Kelley JM, Conboy LA, Davis RB, Kerr CE, Jacobson EE, Kirsch I, Schyner RN, Nam BH, Nguyen LT, et al. Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ*. 2008:336(7651):999–1003.

Kleiman E. 2017. *EMAtools: data management tools for real-time monitoring/ecological momentary assessment data*. In: R package version 0.1.

Koban L, Wager TD. Beyond conformity: social influences on pain reports and physiology. *Emotion*. 2016:16(1):24–32.

Krahé C, Springer A, Weinman JA, Fotopoulou AK. The social modulation of pain: others as predictive signals of salience–a systematic review. *Front Hum Neurosci*. 2013:7(386):1–21.

Krishnan A, Woo CW, Chang LJ, Ruzic L, Gu X, Lopez-Sola M, Jackson PL, Pujol J, Fan J, Wager TD. Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *elife*. 2016:5(e15166):1–42.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2015. *lmerTest: tests in linear mixed effects models*. Version R package version 2.0–29.

LaVeist TA, Isaac LA, Williams KP. Mistrust of health care organizations is associated with underutilization of health services. *Health Serv Res*. 2009:44(6):2093–2105.

Levene LS, Baker R, Walker N, Williams C, Wilson A, Bankart J. Predicting declines in perceived relationship continuity using practice deprivation scores: a longitudinal study in primary care. *Br J Gen Pract*. 2018:68(671):e420–e426.

Loeser JD, Melzack R. Pain: an overview. *Lancet*. 1999:353(9164): 1607–1609.

López-Solà M, Geuter S, Koban L, Coan JA, Wager TD. Brain mechanisms of social touch-induced analgesia in females. *Pain*. 2019:160(9):2072–2085.

Losin EAR, Anderson S, Wager TD. Feelings of clinician-patient similarity and trust influence pain report: evidence from simulated clinical interactions. *J Pain*. 2017:18(7):787–799.

Lyu HG, Cooper MA, Mayer-Blackwell B, Jiam N, Hechenbleikner EM, Wick EC, Berenholtz SM, Makary MA. Medical harm: patient perceptions and follow-up actions. *J Patient Saf*. 2017:13(4):199–201.

Martell BA, O'Connor PG, Kerns RD, Becker WC, Morales KH, Kosten TR, Fiellin DA. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. *Ann Intern Med*. 2007:146(2):116–127.

Martucci KT, Mackey SC. Neuroimaging of pain: human evidence and clinical relevance of central nervous system processes and modulation. *Anesthesiology*. 2018:128(6):1241–1254.

Mascarenhas OA, Cardozo LJ, Afonso NM, Siddique M, Steinberg J, Lepczyk M, Aranha AN. Hypothesized predictors of patient–physician trust and distrust in the elderly: implications for health and disease management. *Clin Interv Aging*. 2006:1(2):175–188.

Master SL, Eisenberger NI, Taylor SE, Naliboff BD, Shirinyan D, Lieberman MD. A picture's worth: partner photographs reduce experimentally induced pain. *Psychol Sci*. 2009:20(11):1316–1318.

Mattarozzi K, Todorov A, Marzocchi M, Vicari A, Russo PM. Effects of gender and personality on first impression. *PLoS One*. 2015:10(9):e0135529.

Melzack R. From the gate to the neuromatrix. *Pain*. 1999:82(Supplement 1):S121–S126.

Melzack R. Pain and the neuromatrix in the brain. *J Dent Educ*. 2001:65(12):1378–1382.

Mogil JS. Social modulation of and by pain in humans and rodents. *Pain*. 2015:156(Supplement 1):S35–S41.

Mumford JA, Nichols T. Modeling and inference of multisubject fMRI data. *IEEE Eng Med Biol Mag*. 2006:25(2):42–51.

Oosterhof NN, Todorov A. The functional basis of face evaluation. *Proc Natl Acad Sci*. 2008:105(32):11087–11092.

Pearson SD, Raeke LH. Patients' trust in physicians: many theories, few measures, and little data. *J Gen Intern Med*. 2000:15(7):509–513.

Platonova EA, Kennedy KN, Shewchuk RM. Understanding patient satisfaction, trust, and loyalty to primary care physicians. *Med Care Res Rev*. 2008:65(6):696–712.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 2012:59(3):2142–2154.

Price DD, Mcgrath PA, Rafii A, Buckingham B. The validation of visual analog scales as ratio scale measures for chronic and experimental pain. *Pain*. 1983:17(1):45–56.

R Core Team. *A language and environment for statistical computing*. Vienna (Austria): Foundation for Statistical Computing; 2019.

Reddan MC, Wager TD. Modeling pain using fMRI: from regions to biomarkers. *Neurosci Bull*. 2018:34(1):208–215.

Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev*. 2009:16(2):225–237.

Rule NO, Krendl AC, Ivcevic Z, Ambady N. Accuracy and consensus in judgments of trustworthiness from faces: behavioral and neural correlates. *J Pers Soc Psychol*. 2013:104(3):409–426.

Schmid J, Theysohn N, Ga F, Benson S, Gramsch C, Forsting M, Gizewski ER, Elsenbruch S. Neural mechanisms mediating positive and negative treatment expectations in visceral pain: a functional magnetic resonance imaging study on placebo and nocebo effects in healthy volunteers. *Pain*. 2013:154(11):2372–2380.

Sep MS, Van Osch M, Van Vliet LM, Smets EM, Bensing JM. The power of clinicians' affective communication: how reassurance about non-abandonment can reduce patients' physiological arousal and increase information recall in bad news consultations. An experimental study using analogue patients. *Patient Educ Couns*. 2014:95(1):45–52.

Shaheed CA, Maher CG, Williams KA, Day R, McLachlan AJ. Efficacy, tolerability, and dose-dependent effects of opioid analgesics for low back pain: a systematic review and meta-analysis. *JAMA Intern Med*. 2016:176(7):958–968.

Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE. Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum Brain Mapp*. 2014:35(5):1981–1996.

Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, Westbrook J, Tutty M, Blike G. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med*. 2016:165(11):753–760.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004:23(Suppl 1):S208–S219.

Thom DH, Campbell B. Patient-physician trust: an exploratory study. *J Fam Pract*. 1997:44(2):169–176.

Todorov A. Evaluating faces on trustworthiness. *Ann N Y Acad Sci*. 2008:1124(1):208–224.

Todorov A. Evaluating faces on social dimensions. In: Todorov A, Fiske ST, Prentice DA, editors. *Social neuroscience: toward understanding the underpinnings of the social mind*. New York (NY): Oxford University Press; 2011. pp. 54–76.

Todorov A, Oh D. The structure and perceptual basis of social judgments from faces. In: *Advances in experimental social psychology*. Vol. 63. Cambridge (MA): Academic Press; 2021. pp. 189–245.

Todorov A, Mandisodza AN, Goren A, Hall CC. Inferences of competence from faces predict election outcomes. *Science*. 2005:308(5728):1623–1626.

Todorov A, Baron SG, Oosterhof NN. Evaluating face trustworthiness: a model based approach. *Soc Cogn Affect Neurosci*. 2008:3(2):119–127.

Vachon-Presseau E, Roy M, Woo CW, Kunz M, Martel MO, Sullivan MJ, Jackson PL, Wager TD, Rainville P. Multiple faces of pain: effects of chronic pain on the brain regulation of facial expression. *Pain*. 2016:157(8):1819–1830.

Wager TD, Atlas LY, Leotti LA, Rilling JK. Predicting individual differences in placebo analgesia: contributions of brain activity during anticipation and pain experience. *J Neurosci*. 2011:31(2):439–452.

Wager TD, Atlas LY, Lindquist MA, Roy M, Woo C-W, Kross E. An fMRI-based neurologic signature of physical pain. *N Engl J Med*. 2013:368(15):1388–1397.

Ward MM, Sundaramurthy S, Lotstein D, Bush TM, Neuwelt CM, Street RL Jr. Participatory patient–physician communication and morbidity in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2003:49(6):810–818.

Watson A, el-Deredy W, Iannetti GD, Lloyd D, Tracey I, Vogt BA, Nadeau V, Jones AK. Placebo conditioning and placebo analgesia modulate a common brain network during pain anticipation and perception. *Pain*. 2009:145(1):24–30.

Willis J, Todorov A. First impressions making up your mind after a 100-ms exposure to a face. *Psychol Sci*. 2006:17(7):592–598.

Wincenciak J, Dzhelyova M, Perrett DI, Barraclough NE. Adaptation to facial trustworthiness is different in female and male observers. *Vis Res*. 2013:87(1):30–34.

Woo CW, Koban L, Kross E, Lindquist MA, Banich MT, Ruzic L, Andrews-Hanna JR, Wager TD. Separate neural representations for physical pain and social rejection. *Nat Commun*. 2014:5(1):5380.

Woolrich M, Brady M, Smith SM. Hierarchical fully Bayesian spatio-temporal analysis of FMRI data. *NeuroImage*. 2001:13(6):287–8119.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 2011:8(8):665–670.

Zebrowitz LA, Hall JA, Murphy NA, Rhodes G. Looking smart and looking good: facial cues to intelligence and their origins. *Personal Soc Psychol Bull*. 2002:28(2):238–249.

Zunhammer M, Bingel U, Wager TD. Placebo effects on the neurologic pain signature: a meta-analysis of individual participant functional magnetic resonance imaging data. *JAMA Neurol*. 2018:75(11):1321–1330.