Neural and sociocultural mediators of ethnic differences in pain

Elizabeth A. Reynolds Losin [1]*, Choong-Wan Woo [2,3], Natalia A. Medina[1],
Jessica R. Andrews-Hanna[4], Hedwig Eisenbarth [5] and Tor D. Wager [6]*

Understanding ethnic differences in pain is important for addressing disparities in pain care. A common belief is that African Americans are hyposensitive to pain compared to Whites, but African Americans show increased pain sensitivity in clinical and laboratory settings. The neurobiological mechanisms underlying these differences are unknown. We studied an ethnicity- and gender-balanced sample of African Americans, Hispanics and non-Hispanic Whites using functional magnetic resonance imaging during thermal pain. Higher pain report in African Americans was mediated by discrimination and increased frontostriatal circuit activations associated with pain rating, discrimination, experimenter trust and extranociceptive aspects of pain elsewhere. In contrast, the neurologic pain signature, a neuromarker sensitive and specific to nociceptive pain, mediated painful heat effects on pain report largely similarly in African American and other groups. Findings identify a brain basis for higher pain in African Americans related to interpersonal context and extranociceptive central pain mechanisms and suggest that nociceptive pain processing may be similar across ethnicities.

A common belief since the time of slavery in the United States is that African Americans (AA) feel less pain than White Americans (WA)[1]. This belief has been related to undertreatment of pain in AA[2], which contributes to widespread and persistent racial and ethnic health disparities[3]. Paradoxically, AA, and in some cases Hispanic Americans (HA), actually report more pain than non-Hispanic WA in both clinical[4] and laboratory[5,6] settings. To reduce pain assessment and treatment disparities, we need to better understand the mechanisms contributing to higher pain sensitivity amongst AA.

Ethnic differences in pain sensitivity may be related to multiple factors[7–9]. Higher pain reported by AA may be due in part to variation in the basic sensory and affective processes specific to pain. Findings of less effective descending pain modulation[4,10,11], and a lower prevalence of antinociceptive genetic variants related to the endogenous opioid system in AA compared to non-Hispanic WA[12], suggest potential differences in nociceptive sensitivity. Higher pain reported by AA may also be due in part to extranociceptive aspects of pain, including sociocultural variation in life experiences that affect how people value, explain and respond to (for example, avoid or cope with) pain[7], and how they respond to the contexts in which pain sensitivity is assessed (that is, laboratory tests). African Americans experience higher incidences of discrimination[13,14] and stressful and traumatic life events[15] compared to non-Hispanic WA and engage in increased hypervigilance[16], pain catastrophizing[16] and religious pain coping[17]. In particular, increased hypervigilance[16] and discrimination[18–21] have been associated with higher reported pain amongst AA.

The relative contributions of these nociceptive and extranociceptive factors to ethnic differences in pain sensitivity are still unclear. Brain imaging can help resolve this confusion by providing measures of the multiple central nociceptive and extranociceptive systems that contribute to pain processing[22]. Here we use functional magnetic resonance imaging (fMRI) during experimental thermal pain induction (Fig. 1) in conjunction with a battery of sociocultural measures to test whether potential nociceptive and extranociceptive mechanisms differ across ethnic groups and relate to ethnic group differences in pain. We recruited a diverse sample of $n = 88$ participants consisting of 28 AA, 30 HA and 30 non-Hispanic WA (see Table 1 for sample characteristics). This sample size provides adequate power to detect racial/ethnic differences in pain report and neural responses in the range of published effect sizes in the behavioural and neuroimaging literature[6,23] (see Methods for power calculations).

To examine sociocultural contributors to pain processing, we tested whether a range of sociocultural factors previously found to influence pain would mediate group differences in pain ratings. We also searched across the brain for regions that (1) responded differently to painful heat across ethnic groups and (2) exhibited relationships with pain ratings and sociocultural factors such as perceived discrimination. If ethnic differences in pain are due in part to enhanced pain valuation and avoidance motivation elicited in response to a more adverse sociocultural context, we would expect to find heightened activity and relationships with pain ratings and sociocultural measures in brain systems related to extranociceptive aspects of pain. One such system is the pathway connecting the ventromedial prefrontal cortex (vmPFC) and nucleus accumbens (NAc), which has been shown to be involved in emotion regulation[24,25], pain valuation[26,27] and pain chronification[28], and can exhibit changes in response to chronic stress[29].

To examine nociceptive sensitivity, we looked at activity in brain regions previously linked to nociception (for example, the secondary somatosensory cortex and dorsal posterior insula[22,30]) in whole-brain analyses and tested responses in a multivariate fMRI activity pattern that closely tracks the intensity and affect of evoked nociceptive pain[23], termed the neurologic pain signature (NPS). The

[1]Department of Psychology, University of Miami, Coral Gables, FL, USA. [2]Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, Republic of Korea. [3]Department of Biomedical Engineering, Sungkyunkwan University, Suwon, Republic of Korea. [4]Department of Psychology; Cognitive Science, University of Arizona, Tucson, AZ, USA. [5]Victoria University of Wellington, Wellington, New Zealand. [6]Dartmouth College, Department of Psychology and Brain Sciences, Hanover, NH, USA. *e-mail: e.losin@miami.edu; tor.d.wager@dartmouth.edu
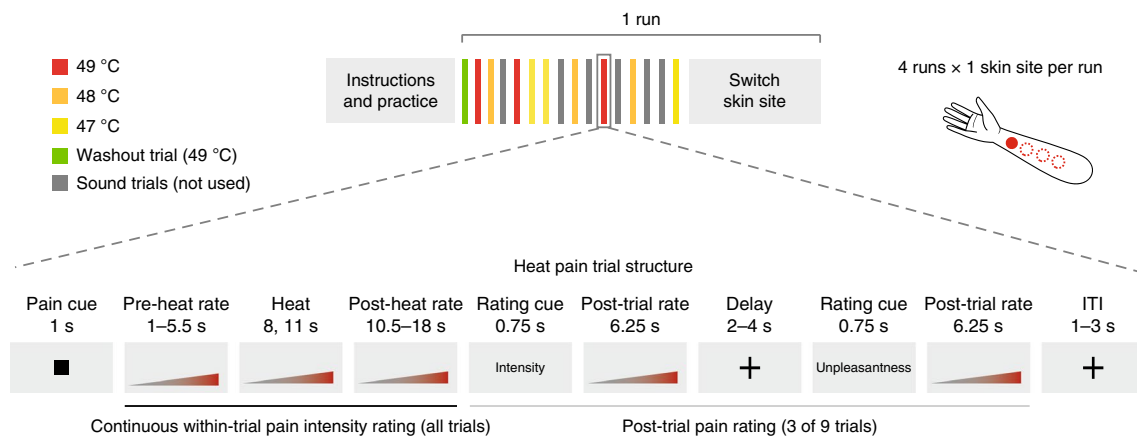
Fig. 1 | fMRI thermal stimulation task design. Experimental details can be found in the "Thermal stimulation" section. ITI, inter-trial interval.

Table 1 | Sample characteristics

	WA[a]	AA	HA		
Measure	**N or M (s.d.)**	**N or M (s.d.)**	**N or M (s.d.)**	**χ^2/F**	**P**
Analysed sample (n)	30	28	30	–	–
Gender (f[b])	15	14	15	0	1.0
Age in years	27.98 (3.97)	30.34 (7.74)	28.25 (4.64)	1.5	0.23
Recruitment (IBG[c])	16	11	15	1.24	0.54
fMRI sequence (MB[d])	8	6	11	1.72	0.42

χ^2 values are from Pearson's chi-squared tests comparing actual participant counts in each group for each measure to counts for perfectly balanced groups. F values are from linear models in R (command lm) for the ethnicity effect (coded as a three-level factor) on each measure. N, number of participants; M (s.d.), mean with standard deviation in parenthesis. [a]WA, non-Hispanic WA. [b]f, female; other participants in analysed samples identified as male. [c]IBG, University of Colorado Boulder Institute for Behavioral Genetics; other recruitment source was Denver metro area Craigslist. [d]MB, multi-band (eight simultaneous slice acquisitions, TR = 0.46 s); other fMRI sequence was standard (single slice acquisition, TR = 1.3 s).

NPS does not explain all aspects of pain, but is sensitive and specific to pain in the 90–100% range across multiple fMRI studies (for a review see ref. [31]). It provides an objective measure that responds strongly to pain evoked by noxious input in particular, but not to several types of psychological 'pain' and negative emotion[32–34]. If AA are more sensitive to pain than non-Hispanic WA due in part to enhanced nociceptive input (for example, receptor genetics or reduced descending inhibition), then we would expect brain systems related to nociception and pain affect, including the NPS, to (1) exhibit heightened activity in AA and (2) mediate ethnicity-related increases in pain ratings.

We found that higher pain in AA was paralleled by a steeper dose–response relationship between noxious stimulus intensity and activity in brain regions related to emotion regulation and valuation, including the vmPFC and NAc, which correlated with pain ratings, perceived discrimination and reduced trust in the experimenter. Furthermore, the vmPFC–NAc pathway mediated the relationship between painful stimulus intensity and pain ratings in AA participants, but not in non-AA participants (HA or non-Hispanic WA). In contrast, the NPS tracked increases in noxious stimulus intensity in all three ethnic groups, with no evidence for elevated responses in AA participants and moderate evidence in support of the null hypothesis of equivalence among ethnic groups. Furthermore, we found moderate evidence that the NPS mediated the dose–response relationship between painful stimulus intensity and pain rating equivalently between AA and non-AA participants. These NPS findings suggested that nociceptive pain processing is probably similar across ethnic groups. Taken together, our findings identify a brain basis for higher pain in AA related to sociocultural context and extra-nociceptive central pain mechanisms, suggesting

that interventions geared towards reducing discrimination and increasing clinician trust may be promising ways to mitigate ethnic disparities in pain.

Results

African American participants report pain as more intense and unpleasant than non-Hispanic WA and HA participants. Participants continuously rated moment-to-moment pain intensity during thermal stimulation at three intensity levels (low (L) = 47 °C, medium (M) = 48 °C and high (H) = 49 °C) (see Methods for details about within-trial pain intensity rating). All stimulus intensities were above the median temperature associated with reported pain in previous studies[23] and the activation of specific nociceptors[35] (>45 °C). On average, participants rated the maximum pain intensity for each temperature between 17 (moderate) and 53 (very strong) on the 0–100 generalized labelled magnitude scale (L: mean (M) = 32.10, s.d. = 20.22; M: M = 41.94, s.d. = 20.36; H: M = 50.34, s.d. = 21.20). As expected, the area under the curve (AUC) of within-trial pain intensity rating (hereafter referred to as 'pain rating') increased with increasing temperature, showing a dose–response relationship ($t(84)$ = 17.02, P < 0.001, b = 6.68, 95% confidence interval (CI) = [5.92, 7.44]; Fig. 2a), and this effect was seen in each ethnic group when analysed separately (all P < 0.001; Fig. 2a and Supplementary Table 1; see Extended Data Fig. 1 for graphs of raw continuous pain-rating data). Consistent with previous studies of experimental and clinical pain[4–6], we found that AA participants rated their pain as more intense than HA and WA participants ($t(84)$ = 2.79, P = 0.01, b = 9.73, CI = [2.97, 16.49]; Fig. 2a, left), and exhibited a steeper dose–response relationship between noxious stimulus intensity and pain rating ($t(84)$ = 2.50, P = 0.01, b = 2.12,
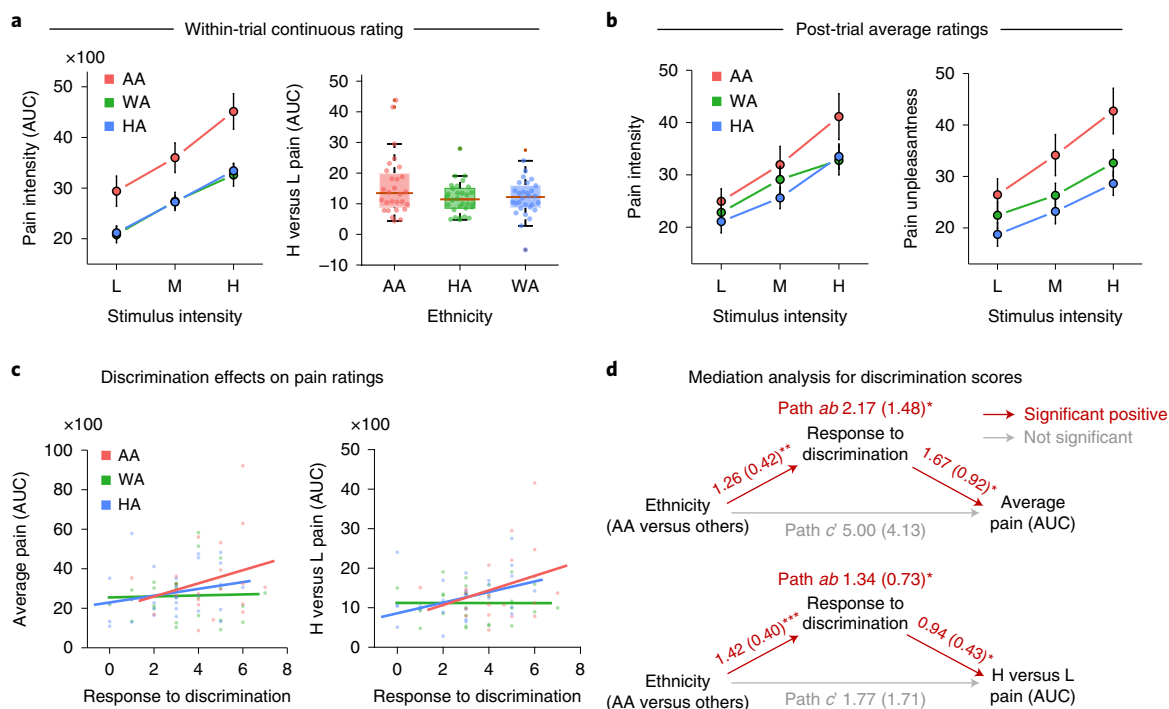
Fig. 2 | Differences in pain rating across ethnic groups and mediation by discrimination. a, Mean AUC of within-trial pain intensity rating (present in all trials) at each temperature of heat stimulation (left, line plot) and the difference between high and low heat stimulation (right, box plot) in each ethnic group. Box plot elements: red centre line, median; box limits, upper and lower quartiles; whiskers, limits of non-outlier data; points matched to box shading, participant means; dark red points, outliers. **b**, Mean post-trial pain intensity (left) and pain unpleasantness (right) rating (present in one of four trials) at each temperature of heat stimulation in each ethnic group. **a,b**, AA participants rated pain as more intense and unpleasant than HA and WA participants, and exhibited a steeper dose–response relationship with painful stimulus intensity. Error bars represent within-participant s.e.m. Data are from 88 participants. **c**, Relationship between participants' frequency of responding to discrimination (for example, by filing a complaint) and the AUC of their within-trial pain intensity rating. Left: average pain rating across three stimulation temperatures; right: average pain rating for high minus low temperatures. **a–c**, AUC values for continuous within-trial pain intensity rating depicted on the y axes have been divided by 100. **d**, Path diagrams and statistics for mediation analyses of participant ethnicity effects (AA versus (HA + WA)) on pain rating by participants' frequency of responding to discrimination. Participants' frequency of responding to discrimination significantly mediated higher pain reports and a steeper dose–response relationship with painful stimulus intensity in AA compared to HA and WA participants. Path coefficients are listed for each path with standard errors in parentheses. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. **c,d**, Data are from the 81 participants who completed the response to discrimination measure.

CI = [0.47, 3.78]; Fig. 2a, right). Comparison of HA and WA participants yielded no statistically significant difference in pain rating or the dose–response relationship between stimulus intensity and pain (all $P > 0.8$; Fig. 2a and Supplementary Table 1). A similar pattern of results was found for post-trial pain intensity and unpleasantness ratings (Fig. 2b and Supplementary Table 1; see Methods for post-trial rating details). Some of the post-trial rating results, however, were only marginally significant ($0.05 < P < 0.1$), probably due to the fact that we had only 25% the number of post-trial ratings compared to within-trial continuous pain intensity ratings.

African American participants report higher discrimination and lower trust in experimenter. Among 19 sociocultural measures hypothesized to potentially help explain observed group differences in pain rating, only three significantly differed between AA and non-AA (HA + WA) participants after correcting for 19 statistical tests (Bonferroni, $P < 0.003$). Compared to WA and HA participants, AA participants reported having experienced more incidences of daily and major discrimination ($t(80) = 5.58$, $P < 0.001$ (corrected), $b = 8.74$, CI = [5.62, 11.86]), having more frequently responded to discrimination (for example, by filing a complaint; referred to hereafter as 'response to discrimination') ($t(78) = 3.60$, $P = 0.01$ (corrected), $b = 1.45$, CI = [0.65, 2.26]) and feeling less trust in the experimenter (the same White male in his mid-30s for all participants)

($t(83) = -3.49$, $P = 0.01$ (corrected), $b = -3.89$, CI = [−6.11, −1.67]). AA participants did not differ statistically significantly from WA and HA participants in other hypothesized contributors to ethnic differences in pain report, including socioeconomic status, stressful life events, hypervigilance and pain catastrophizing (see Supplementary Table 2 for group means and statistics for all measures). Thus, discrimination history and trust in the experimenter were the sociocultural candidate mediators most likely to explain the higher pain ratings of the AA group.

History of discrimination mediates higher pain intensity ratings by AA participants. Among the three candidate mediators, only participants' history of responding to discrimination ('Response to discrimination' in Fig. 2c,d) significantly mediated the relationship between their ethnicity and pain ratings both on average (path ab: $z = 2.12$, $P = 0.03$, $b = 2.17$, CI = [1.34, 3.42]) and their dose–response relationship with painful stimulus intensity (path ab: $z = 2.55$, $P = 0.01$, $b = 1.34$, CI = [0.90, 1.89]; Fig. 2d and Supplementary Table 3). Analyses decomposing these mediation effects into their component parts showed that frequency of responding to discrimination was higher in AA participants than in WA and HA participants (path a, average pain model: $z = 2.92$, $P = 0.004$, $b = 1.26$, CI = [0.98, 1.54]; dose–response model: $z = 3.46$, $P < .001$, $b = 1.42$, CI = [1.16, 1.69]). Participants' frequency of
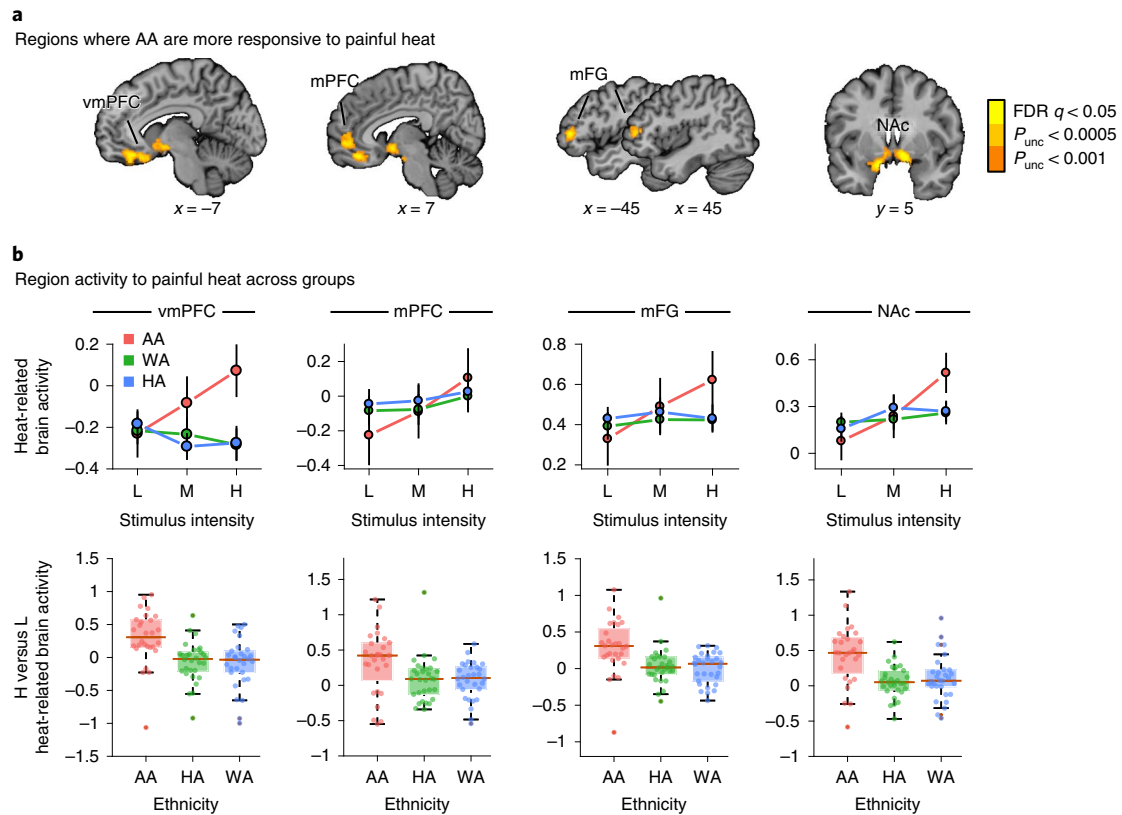
Fig. 3 | Results of whole-brain voxel-wise GLM analysis showing brain regions exhibiting a stronger dose–response effect of painful heat in AA participants. a, vmPFC, mPFC, bilateral mFG and bilateral NAc exhibited a steeper dose–response effect of painful stimulus intensity in AA compared to HA and WA participants (FDR-corrected $q < 0.05$ ($P < 0.000047$)). For the purposes of display we included voxels meeting two additional, more liberal, uncorrected voxel-wise thresholds ($P < 0.0005$ and $P < 0.001$) that were in contact with voxels meeting the more stringent FDR-corrected threshold. **b**, Data from the four regions (defined at $P < 0.001$, uncorrected and combining across hemispheres for bilaterally activated NAc and mFG). Top: line plots depicting the mean parameter estimate for each level of stimulus intensity in each ethnic group plotted with error bars representing within-participant s.e.m. Bottom: box plots of the mean parameter estimate difference between high and low stimulus intensity in each ethnic group. Box plot elements: red centre line, median; box limits, upper and lower quartiles; whiskers, limits of non-outlier data; points matched to box shading, participant means; dark red points, outliers. **a,b**, Data are from 88 participants. P_{unc} = uncorrected P value threshold; x and y indicate Montreal Neurological Institute (MNI) coordinates of brain slices shown.

responding to discrimination in turn predicted higher pain ratings (path b; $z = 2.00$, $P = 0.05$, $b = 1.67$, CI = [1.07, 2.03]) and a steeper dose–response relationship with stimulus intensity (path b; $z = 2.35$, $P = 0.02$, $b = 0.94$, CI = [0.61, 1.19]), controlling for participant ethnicity. Together, these findings suggest that a history of responding to discrimination may predispose individuals to react more strongly to physically painful stimuli.

Frontostriatal regions are more responsive to increases in painful heat in AA participants. Next, we used a whole-brain voxel-wise general linear model (GLM) analysis to test whether the higher levels of pain intensity reported by AA participants were accompanied by brain systems that responded differently to painful heat or exhibited a different dose–response relationship with stimulus intensity in AA versus HA and WA participants. A set of brain regions associated with the extra-nociceptive rather than nociceptive aspects of pain exhibited a steeper dose–response relationship with painful stimulus intensity in AA participants compared to HA and WA participants. These regions included frontostriatal regions previously associated with pain valuation[26,27], modulation[24,25] and chronification[28]: the vmPFC, medial prefrontal cortex (mPFC) and NAc, as well as bilateral portions of the middle frontal gyrus (mFG) (Fig. 3a and Supplementary Table 4; all voxel-wise results reported are significant at a false discovery rate (FDR)-corrected threshold of $q < 0.05$

($P < 0.000047$)). Data from each of these regions are shown in Fig. 3b. Tests of average activity across stimulus intensity levels did not reveal any regions that responded statistically significantly differently in AA compared to HA and WA participants at FDR $q < 0.05$.

During painful heat, higher activity in frontostriatal regions mediates higher pain ratings in AA participants. Next, we used multilevel mediation analyses to test whether activity within the regions showing stronger dose–response effects of painful heat in the AA group in the whole-brain analysis (vmPFC, mPFC, NAc and mFG) mediated the relationship between painful stimulus intensity and trial-by-trial pain rating. These models included moderated mediation effects to test whether this stimulus–brain–pain relationship differed between AA and other groups (HA and WA), paralleling behavioural findings on pain reports. We found that activity within the mPFC, mFG and NAc clusters each partially mediated the relationship between painful stimulus intensity and pain ratings (path ab; mPFC: $z = 3.31$, $P < 0.001$, $b = 10.47$, CI = [3.93, 17.55], Fig. 4a; NAc: $z = 2.20$, $P = 0.03$, $b = 7.67$, CI = [0.61, 14.29]; mFG: $z = 2.87$, $P = 0.004$, $b = 9.12$, CI = [2.62, 15.39]; all, Supplementary Table 5). Furthermore, we found that the mPFC mediated painful heat effects on pain rating to a greater degree in the AA group than in the HA and WA groups ($z = 2.00$, $P = 0.046$, $b = 17.88$, CI = [1.22, 43.33]—that is, ethnicity was a significant moderator (Fig. 4a and Supplementary Table 5).
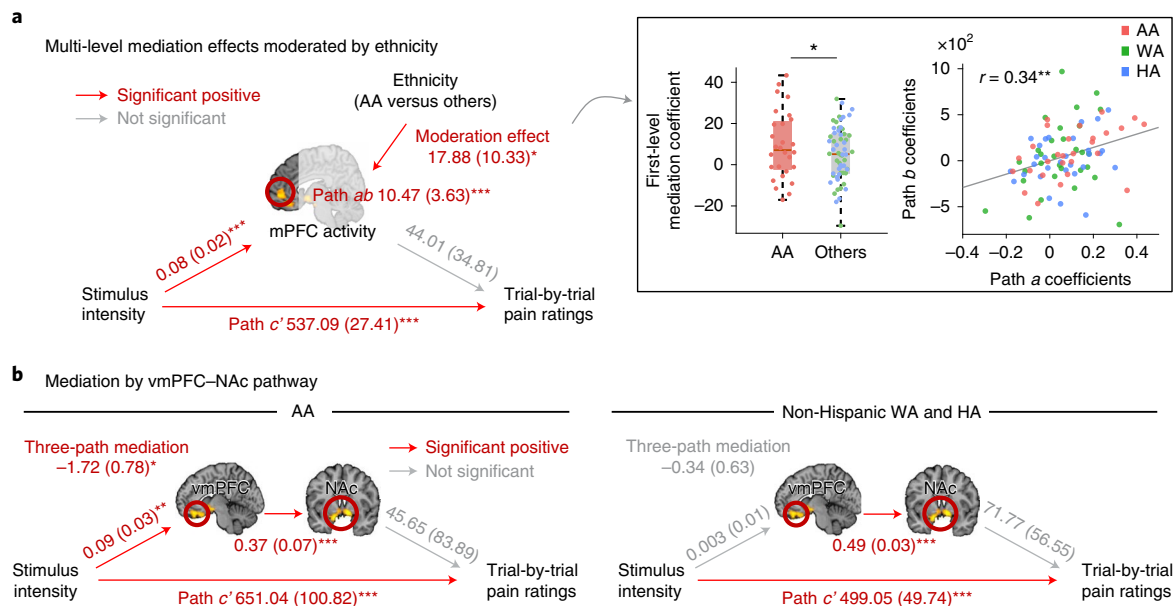
Fig. 4 | Mediation analyses showing that brain regions exhibiting a stronger dose–response effect of painful heat in AA participants in the whole-brain analysis mediate the relationship between painful stimulus intensity and pain rating differently in the AA group compared to the HA and WA groups. **a**, Path diagram and statistics for moderated, multi-level mediation analysis between painful stimulus intensity, activity in the the mPFC region from whole-brain analysis and trial-by-trial pain rating, moderated by participant ethnicity (AA – (HA + WA)). The mPFC mediated the relationship between painful stimulus intensity and pain rating to a greater degree in the AA group than in the HA and WA groups. **b**, Path diagrams and statistics for three-path, multi-level mediation analyses between painful stimulus intensity, connectivity between the vmPFC and NAc regions from whole-brain analysis and trial-by-trial pain rating in the AA and the non-AA (WA + HA) groups separately. Connectivity between the vmPFC and NAc mediated the relationship between painful stimulus intensity and pain rating in the AA group only. Path coefficients are listed for each path with standard errors in parentheses. Data are from 88 participants. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Analyses decomposing the mediation effects of the mPFC, mFG and NAc into their component parts showed that, as expected from the whole-brain analysis, activity within each of these regions increased with increasing stimulus intensity across all groups on average (path a; mPFC: $z = 3.49$, $P < 0.001$, $b = 0.08$, CI = [0.04, 0.11]; NAc: $z = 3.36$, $P < .001$, $b = 0.09$, CI = [0.05, 0.11]; mFG: $z = 3.67$, $P < 0.001$, $b = 0.05$, CI = [0.02, 0.08]). This stimulus intensity–brain (path a) relationship was stronger for the AA group than for the HA and WA groups in the mFG ($z = 2.59$, $P = 0.01$, $b = 0.08$, CI = [0.02, 0.15]) and marginally stronger in the mPFC ($z = 1.67$, $P = 0.10$, $b = 0.07$, CI = [−0.02, 0.15]) (Supplementary Table 5). The relationships between brain responses in the mPFC, mFG and NAc and trial-by-trial pain rating controlling for painful stimulus intensity (path b) were not statistically significant in the group overall (marginal in NAc: $z = 1.55$, $P = 0.12$, $b = 64.61$, CI = [−15.98, 144.82]). However, when not controlling for stimulus intensity in a separate linear mixed-effects regression analysis, activity in the mFG ($t(84) = 2.58$, $P = 0.01$, $b = 0.002$, CI = [0.0005, 0.004]), NAc ($t(84) = 3.84$, $P < 0.001$, $b = 0.004$, CI = [0.002, 0.006]) and mPFC (marginal, $t(84) = 1.92$, $P = 0.06$, $b = 0.002$, CI = [−0.00007, 0.005]) had positive relationships with pain rating, which was stronger in the mFG for AA compared to non-AA participants ($t(84) = 2.33$, $P = 0.02$, $b = 0.004$, CI = [0.0005, 0.008]). Additionally, there were significant correlations between paths a and b in the mediation analysis across individuals (mPFC: $r(86) = 0.34$, $P = 0.001$, Fig. 4a; NAc: $r(86) = 0.24$, $P = 0.03$; mFG: $r(86) = 0.33$, $P = 0.002$), revealing substantial interindividual heterogeneity in their relationships with pain. This means that individuals who showed particularly strong stimulus intensity-dependent increases in the frontostriatal brain mediators also showed stronger positive effects of frontostriatal activity on pain. The significant multi-level mediations are driven by a combination of average path a and b effects and their correlation across individuals. The moderated mediation in the mPFC indicates that individual differences in the role of the mPFC in pain are explained in part by participants' ethnicities, with AA participants having the strongest pro-pain function of this frontal region.

Finally, previous studies have identified vmPFC–NAc connectivity as being important for pain valuation[26,27] and chronification[28], but the two-path mediation results above do not address whether vmPFC and NAc form a functional pathway here. Using a three-path mediation analysis[36,37], we tested (1) whether there is significant connectivity between the vmPFC and NAc clusters and (2) whether this putative functional pathway mediated the relationship between painful stimulus intensity and trial-by-trial pain ratings for the AA and non-AA (HA + WA) groups. Across all participants and in both the AA and non-AA groups separately, the vmPFC and NAc were significantly positively correlated (path $b2$; Fig. 4b and Supplementary Table 6, all $P < 0.001$), suggesting that they form a functional pathway. The vmPFC–NAc pathway mediated the relationship between painful stimulus intensity and pain rating in the AA group ($z = −2.11$, $P = 0.035$, $b = −1.72$, CI = [−2.28, −1.23]) but not in the non-AA group ($z = −0.51$, $P = 0.61$, $b = −0.34$, CI = [−0.77, 0.10]); Fig. 4b). Unfortunately, because moderated three-path mediation is not yet implemented in the toolbox, a direct comparison of these effects between the AA and non-AA groups was not possible. Analysis across all participants showed a mixture of the two effects, with a trend-level mediation effect ($z = −1.66$, $P = 0.10$, $b = −0.83$, CI = [−1.17, −0.50]; Extended Data Fig. 2). Together, these findings suggest that heightened responses to pain in a frontostriatal pathway connecting vmPFC and NAc represent a brain mechanism underlying the higher pain rating observed in the AA group.

During painful heat, higher activity in frontostriatal regions is related to more discrimination and less trust. We also conducted
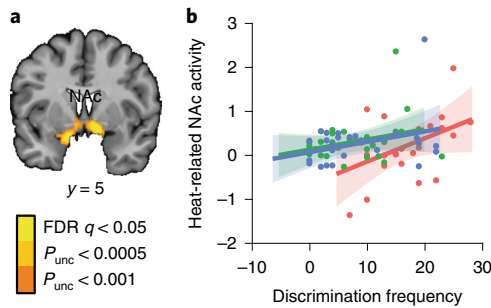
Fig. 5 | Relationship between NAc activity and discrimination frequency. **a**, NAc region from the whole-brain analysis identified as exhibiting a steeper dose–response effect of stimulus intensity on brain activity in AA participants. **b**, Painful-heat-related activity (parameter estimates) from this NAc region plotted as a function of discrimination frequency, with regression lines and standard error bands for each ethnic group. Average activity during painful heat within this NAc region increased with increasing discrimination frequency, and this relationship was stronger for the AA than for the non-AA (HA + WA) group. Dots indicate individual participants. Red dots, AA participants; green dots, WA participants; blue dots, HA participants. Data are from the 83 participants who completed the discrimination frequency measure. P_{unc}, uncorrected P value threshold; y, MNI y coordinate of brain slice.

follow-up exploratory analyses of the relationship between the candidate sociocultural mediators described above (discrimination frequency, response to discrimination and trust in the experimenter) within each region showing stronger dose–response effects of painful heat in the AA group in the whole-brain analysis (vmPFC, mPFC, NAc and mFG). Within these regions we looked for relationships with both average activity during painful heat and the dose–response effect of painful stimulus intensity. We considered main effects or interactions that survived correction for multiple comparisons across sociocultural mediators, regions of interest (ROIs) and pain metrics (Bonferroni, $P < 0.002$, 30 tests). We found that average activity during painful heat within the NAc cluster increased with increasing discrimination frequency ($t(75) = 3.49$, $P = 0.02$ (corrected), $b = 0.03$, CI = [0.01, 0.04]), and that this relationship was stronger for the AA group ($t(75) = 2.19$, $P = 0.03$, $b = 0.04$, CI = [0.003, 0.07]; Fig. 5). Tests of each group separately showed a positive relationship in the AA group ($t(21) = 2.72$, $P = 0.01$, $b = 0.05$, CI = [0.01, 0.09]) and a marginally positive relationship in the HA group ($t(25) = 1.72$, $P = 0.1$, $b = 0.02$, CI = [−0.003, 0.04]; Fig. 5). Together, these findings suggest that the NAc may become sensitized to painful stimuli in those with a history of negative social treatment.

There was also a trust-by-group interaction within the NAc and mPFC clusters, such that average activity during painful heat within the NAc ($t(78) = −4.00$, $P = 0.004$ (corrected), $b = −0.08$, CI = [−0.12, −0.04]) and mPFC ($t(78) = −3.33$, $P = 0.04$ (corrected), $b = −0.1$, CI = [−0.15, −0.04]) was stronger for those with lower trust in the experimenter, particularly for the AA compared to the WA and HA groups. This finding suggests that trust may have buffered against pain-related activation of these regions in AA participants who trusted the experimenter more. Tests in each group separately revealed that NAc activity was (marginally) strongest in AA participants with the least trust in the experimenter $t(24) −1.95$, $P = 0.06$, $b = −0.04$, CI = [−0.09, 0.002]), whereas it was strongest in HA ($t(26) = 2.73$, $P = 0.01$, $b = 0.04$, CI = [0.009, 0.07]) and WA($t(24) = 2.13$, $P = 0.04$, $b = 0.03$, CI = [0.001, 0.07]) participants with the most trust. mPFC activity was also (marginally) strongest in AA participants with the least trust in the experimenter ($t(24) = −1.99$, $P = 0.06$, $b = −0.06$, CI = [−0.11, 0.002]. We did not find any statistically significant relationships between the third

sociocultural mediator, response to discrimination, and average brain responses to painful heat. We also did not find any statistically significant relationships between any of the sociocultural mediators and the dose–response effect of stimulus intensity when correcting for multiple comparisons (Bonferroni, $P < 0.002$, 30 tests).

NPS responses to painful heat are probably similar across ethnic groups. Next we conducted a more sensitive test (compared to the whole-brain analysis) for ethnic group differences in nociception sensitive brain systems using the NPS[23] (Fig. 6a). Replicating our previous findings[31], the NPS responded more strongly with increasing painful stimulus intensity ($t(84) = 8.11$, $P < 0.001$, $b = 4.01$, CI = [3.10, 5.01]; Fig. 5b). The linear dose–response effect of stimulus intensity on NPS response was seen in each ethnic group when analysed separately (all $P < 0.01$; Supplementary Table 7). Furthermore, there were no statistically significant differences among ethnic groups in the response of the NPS to painful heat on average ((AA − (HA + WA)): $t(83) = −0.92$, $P = 0.36$, $b = −2.93$, CI = [−9.54, 3.40]; (HA − WA): $t(83) = 0.70$, $P = 0.49$, $b = 1.25$, CI = [−2.22, 4.78]) or in the dose–response relationship of the NPS with painful stimulus intensity ((AA − (HA+WA)): $t(84) = −0.07$, $P = 0.94$, $b = −0.08$, CI = [−2.32, 2.04]; (HA − WA): $t(84) = 0.09$, $P = 0.93$, $b = 0.05$, CI = [−1.13, 1.24]). Follow-up estimation of Bayes factors for these tests provided moderate evidence (ranging 5.62–8.47/1 in favour of the null hypotheses) in support of equivalence across ethnic groups in the NPS response ((AA − (HA + WA)): Bayes factor (BF_{01}) = 5.62; (HA − WA): BF_{01} = 6.70) and the dose–response effect of painful stimulus intensity ((AA − (HA + WA)): BF_{01} = 8.47; (HA − WA): BF_{01} = 8.46). The combination of positive findings for NPS temperature effects and moderate evidence in favour of the null hypothesis of equivalence of NPS responses across ethnic groups suggests that nociceptive pain systems probably operate similarly across ethnicities, and demonstrates that (1) our fMRI measures are roughly equally sensitive and (2) our haemodynamic model fits equally well across ethnic groups.

NPS mediates dose–response relationship between painful stimulus intensity and pain rating similarly between AA and non-AA groups. Finally, we tested whether the NPS mediated the relationship between painful stimulus intensity and trial-by-trial pain rating, and whether it did so differently in the AA and non-AA (HA + WA) groups. Consistent with our previous work, we found that the NPS positively mediated the relationship between painful stimulus intensity and pain rating ($z = 3.38$, $P < 0.001$, $b = 20.03$, CI = [9.31, 31.42]; Fig. 6c and Supplementary Table 8). Importantly, there was no statistically significant difference in this effect between AA and non-AA (HA + WA) groups ($z = −0.14$, $P = 0.89$, $b = −1.76$, CI = [−24.61, 23.04]. Follow-up Bayes factor estimation provided moderate evidence ($BF_{01} = 8.41$) in support of the null hypothesis of equivalence between AA and non-AA participants in NPS mediation of the relationship between painful stimulus intensity and pain rating. Interestingly, NPS mediation of the relationship between painful stimulus intensity and pain rating was weaker in HA compared to WA ($z = −2.09$, $P = 0.04$, $b = −11.41$, CI = [−23.86–0.75]).

Decomposing the mediation into its component parts showed that, as expected from the regression analysis on NPS responses, increases in painful stimulus intensity produced significant increases in NPS responses (path a: $z = 3.85$, $P < 0.001$, $b = 3.69$, CI = [2.94, 4.49]), which did not statistically significantly differ between AA and non-AA groups ($P > 0.4$; Supplementary Table 5). There was also a significant positive relationship between NPS responses and trial-by-trial pain rating controlling for painful stimulus intensity (path b) such that as NPS responses increased, pain intensity ratings increased ($z = 3.59$, $P < 0.001$, $b = 6.55$, CI = [3.04, 10.46]. This effect also did not statistically significantly differ between the AA and non-AA groups ($P > 0.3$; Supplementary Table 5). The NPS also had
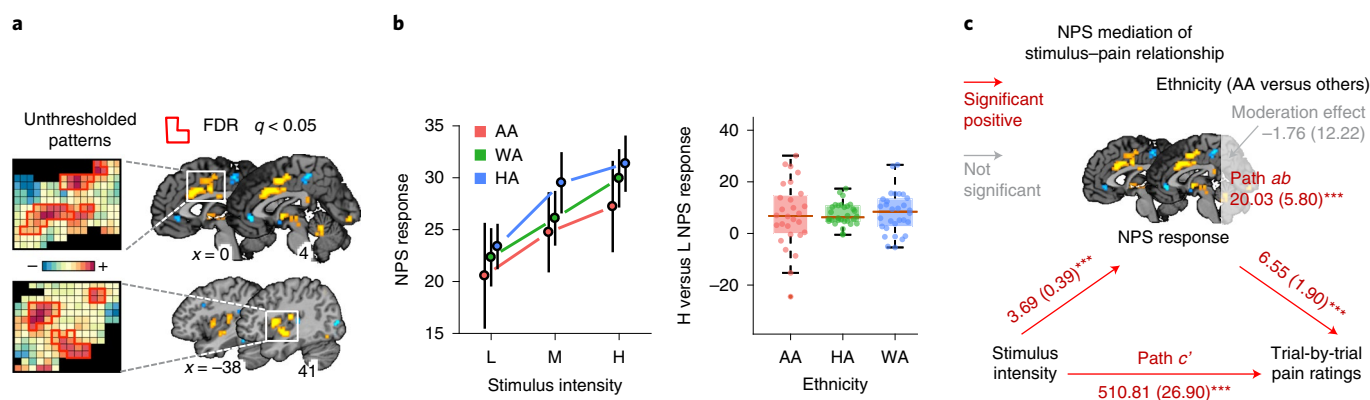
Fig. 6 | NPS responses to painful heat and relationship with pain rating across ethnic groups. a, The NPS pattern, an a priori multivariate pattern of fMRI signal that is sensitive and specific to nociceptive physical pain[23]. **b**, Mean NPS response to each temperature of heat stimulation (left, line plot) and the difference between high and low heat stimulation (right, box plot) in each ethnic group. Error bars represent within-participant s.e.m. Box plot elements: red centre line, median; box limits, upper and lower quartiles; whiskers, limits of non-outlier data; points matched to box shading, participant means; dark red points, outliers. NPS response values indicate the strength of the signature response pattern expression and are calculated by taking a dot-product of the NPS pattern weights and activation maps for each single heat trial. NPS responses increased with painful stimulus intensity across ethnic groups; neither the average NPS response nor the dose–response relationship with painful stimulus intensity statistically significantly differed between ethnic groups, and we found moderately strong evidence for their equivalence. **c**, Path diagram from multi-level mediation analysis between painful stimulus intensity, NPS pattern expression during painful heat and trial-by-trial pain rating moderated by participant ethnic group ((AA – (HA + WA)). The NPS mediated the relationship between painful stimulus intensity and pain rating, and this effect did not statistically significantly differ between AA and non-AA (HA and WA) participants; we found moderately strong evidence of its equivalence. Path coefficients are listed for each path with standard errors in parentheses. Data are from 88 participants. ***$P < 0.001$; x, MNI x coordinate of brain slice.

a strong positive relationship with pain rating ($t(84) = 5.94$, $P < 0.001$, $b = 0.14$, CI = [0.10, 0.19]) when not controlling for stimulus intensity in a separate linear mixed-effects regression analysis. However, when not controlling for painful stimulus intensity, the NPS–pain relationship was weaker in the AA group than in non-AA groups ($t(84) -3.02$, $P = 0.003$, $b = -0.14$, CI = [-0.23, -0.05]). This interaction with participant ethnicity suggests that extranociceptive brain systems may have contributed to the stimulus intensity-dependent aspects of pain rating to a greater degree in AA than in non-AA participants. This result is consistent with our finding of stronger dose–response effects of painful heat in frontostriatal regions outside the NPS in the AA group in the whole-brain analysis. Together, these findings suggest that the brain system represented by the NPS, which receives nociceptive input from the spinothalamic and spinoreticular pathways and is particularly strongly associated with evoked pain, is unlikely to underlie the higher pain reported by AA participants. Additionally, we did not find statistically significant relationships between any of the candidate sociocultural mediators and NPS pattern expression or the dose–NPS pattern expression relationship when correcting for multiple comparisons (Bonferroni, $P < 0.002$, 30 tests), providing no positive evidence that the NPS is related to the sociocultural factors that mediated ethnic differences in pain rating.

Discussion

African Americans consistently exhibit increased pain sensitivity compared to non-Hispanic WA in clinical and laboratory settings[4–6]. The neurobiological mechanisms underlying these differences are unknown, though they probably contribute to the persistence of ethnic disparities in pain diagnosis and treatment. We replicated findings of higher pain report amongst AA compared to HA and non-Hispanic WA[4–6], and found that the frequency of responding to discrimination mediated ethnic differences in pain report. Higher pain report amongst AA participants was accompanied by differences in brain activation during pain that were not shared with other ethnic groups, some of which mediated ethnic group differences in pain report. Specifically, frontostriatal regions previously

associated with pain valuation[26,27], regulation[24,25,38] and chronification[28] (vmPFC, mPFC and NAc) exhibited a stronger dose–response effect of painful stimulus intensity in AA participants compared to HA and WA participants. Furthermore, activity within the vmPFC–NAc pathway mediated the relationship between painful stimulus intensity and pain ratings in AA participants but not in non-AA (HA + WA) participants. NAc activity correlated with individual differences in discrimination frequency across groups, and NAc and mPFC activity correlated more positively with low experimenter trust in AA participants than in non-AA participants. In contrast, the NPS, a multivariate signature associated with nociceptive pain, tracked painful stimulus intensity in each ethnic group separately (with roughly equivalent effect sizes) and mediated stimulus intensity effects on pain, as in previous studies[23]. However, NPS responses to painful stimuli and mediation effects did not statistically significantly differ between AA and non-AA groups and we found moderate evidence of their equivalence in Bayes factor analyses. We also found no positive evidence that NPS responses were related to discrimination or trust. Together these findings suggest that the nociceptive pathways comprising the NPS (most notably specific portions of the posterior insula, medial and lateral thalamus, anterior cingulate and anterior and mid-insula) show largely similar sensitivity across ethnic groups, and that the drivers of ethnic differences in pain sensitivity probably arise elsewhere.

There is some evidence that higher pain reports in AA may be a learned behaviour in response to a history of inadequate pain treatment, which has been found to characterize minority medical interactions[39–41]. Thus, the increased pain reported by AA compared to HA and non-Hispanic WA could be directly related to pain report as an interpersonal, communicative behaviour. However, our findings that frontostriatal brain regions play an important role in increased pain report in AA participants suggest that ethnic differences in pain report are unlikely to arise solely at the level of communicative decision making. Rather these differences may rely, at least in part, on evaluative mechanisms that are linked to pain independent of stimulus intensity[31,37] and play a role in psychological

responses to, and modulation of, pain, particularly as it becomes chronic. Specifically, the vmPFC and NAc regions involved in ethnic differences in pain report have not typically been found to track experimentally induced pain intensity[37,42] (as replicated here in WA and HA participants) but do track intensity of pain once it has become chronic[33,42,43]. Additionally, activity in the vmPFC is higher and vmPFC–insula connectivity is greater in patients with fibromyalgia[44,45] and failed back surgery syndrome[46], suggesting a shift from an 'anti-pain' role for the vmPFC in healthy individuals to a 'pro-pain' role in pain patients. Thus, the pattern of frontostriatal responses to pain seen in AA participants in the present study resembles the pattern of frontostriatal responses to pain seen in chronic pain patients.

Our findings suggest that the link between chronic pain and ethnic differences in pain sensitivity may lie in the chronic stress associated with discrimination. Discrimination has been consistently associated with chronic stress and other adverse health outcomes in AA and other minority groups in both theoretical models of minority health disparities[47–51] and empirical studies[47,48,52–55]. The mPFC/vmPFC–NAc pathway has also been found to undergo plasticity as a result of chronic stress[29], and activity within this pathway predicts the transition to chronic pain[28], a disorder also associated with previous stress, trauma and early life adversity[56]. Furthermore, activity in the vmPFC and nearby mPFC shows increased connectivity with nociceptive regions (insula) when pain is uncontrollable and thus more stressful[57], and mPFC/vmPFC activity has been shown to be related to pain catastrophizing[58] and anticipatory anxiety[59] and negatively correlated with placebo analgesia[59]. Thus, the higher pain-related mPFC, vmPFC and NAc activity and vmPFC–NAc mediation of the stimulus–pain relationship we observed in AA participants may be related to the chronic stress associated with discrimination and resulting altered appraisals of painful stimuli as more threatening, more potentially damaging and less controllable. This hypothesis is further supported by our findings of higher levels of discrimination in the AA group and a positive relationship between activity in the NAc during pain and perceived discrimination.

It is also plausible that the higher pain sensitivity we and others have observed in AA compared to WA participants may be related to previous negative experiences with medical care in particular, which are more common in AA compared to non-Hispanic WA populations[60]. Although we did not directly measure participants' previous experiences with medical care, this hypothesis is supported by the relationship between average NAc and mPFC activity during painful heat and low trust in the experimenter that we observed in AA participants to a greater degree than in HA and WA participants, who showed the opposite pattern when tested separately. Together, these findings support the hypothesis that increased exposure to stressful life experiences such as discrimination, and accompanying changes in brain systems related to pain valuation, modulation and chronification, may contribute to heightened pain report in AA compared to non-Hispanic WA. These findings set the stage for future work specifically aimed at understanding the role played by discrimination in pain disparities, how this might be related to experiences with medical care and what interventions might affect both brain and behavioural manifestations of higher pain sensitivity.

Our findings also contribute to an understanding of other potential sources of ethnic variations in pain sensitivity. Previous studies have suggested physiological variation in peripheral and central mechanisms of nociception as a likely cause of higher pain sensitivity in AA[4,10–12,61,62]. However, our finding of moderate evidence in support of equivalent expression of the NPS pattern across the AA, HA and WA groups, as well as moderate evidence of equivalent NPS mediation of the stimulus intensity–pain relationship between the AA and non-AA groups, challenges that view and suggests that nociceptive contributions to pain may be largely similar across these different ethnic groups. The only NPS effects related to ethnicity

were a weaker mediation of the stimulus–pain relationship in HA compared to WA participants and a weaker relationship between NPS pattern expression and pain ratings in AA compared to non-AA participants when not controlling for painful stimulus intensity, though the NPS–pain relationship did not statistically significantly differ across ethnic groups when controlling for stimulus intensity. These findings converge with our whole-brain findings of stronger dose–response effects of painful heat in frontostriatal regions outside the NPS in the AA group, suggesting that frontostriatal systems that are not nociceptive and outside the NPS contribute more to pain ratings in AA compared to non-Hispanic WA.

One reason the NPS may not capture all of the neural features contributing to pain ratings in AA is that, although the training sample for the NPS reflected roughly nationally representative ethnic diversity[63] (79% Caucasian, 5% HA and 16% AA), it still contained a majority of WA[23]. Thus, although recent studies in our laboratories and others have found that the NPS generalizes across ethnically and geographically diverse samples from North America, Europe and Asia[64,65] (see also Van Oudenhove, L., Kragel, P. A., Kano, M., Ly, G. & Wager, T. D., manuscript in preparation), the present study adds to our understanding of the aspects of the NPS that may and may not generalize to non-WA samples. Future studies should focus on models of activity in frontostriatal systems that contribute to pain independently of the NPS and which may account for variability in pain sensitivity across ethnic groups.

Our findings of neural differences related to sociocultural factors are in line with a growing literature in cultural neuroscience. This literature has demonstrated differences in brain function underlying cultural variation in a variety of social and cognitive domains including emotion processing, perception of the self and others, sensory perception and attention (for reviews see refs. [66,67]). Our findings extend these cultural neuroscience findings by demonstrating that sociocultural variability can also be seen in brain systems connected to health outcomes[68]. Our findings also provide important evidence against the counterfactual and damaging view held by both clinicians and lay people that AA are less sensitive to pain than WA[1]. In addition to providing another replication of higher pain report among AA compared to WA, our findings also reveal a potential neurobiological mechanism underlying these differences.

Findings in the present study should be interpreted in light of several limitations. First, there was greater head movement in AA participants than in WA and HA participants. However, we believe it is extremely unlikely that head movement explains the observed group differences in brain activity or brain–pain correlations, for three reasons. (1) We took extensive measures to minimize the influence of head movement on our results, using both standard movement corrections (for example, image realignment) as well as more stringent movement controls (for example, removal of image intensity outliers and trials suggested to have high multicollinearity with movement regressors). (2) When we repeated analyses additionally controlling for average head motion during each trial, results were qualitatively unchanged (Supplementary Tables 9–11). (3) The pattern of ethnic group differences in head movement was not consistent with that of ethnic group differences in neural responses to pain. Specifically, we found some evidence of a weaker relationship between the NPS and pain ratings in both the AA and HA groups compared to the WA group, whereas we found a group difference in head motion only between the AA and WA groups.

Second, ethnic group differences in the effects of the experimental context, including racial/ethnic concordance with the experimenter and familiarity with research participation and the MRI environment, may have contributed to our findings. However, although we did find ethnic differences in feelings of trust and racial/ethnic similarity towards the experimenter, neither of these measures was related to ethnic group differences in pain rating or neural responses. Future studies will be needed to better characterize

the contributions made by ethnic group differences in the effects of the experimental/clinical context to ethnic differences in pain and its neural correlates.

Third, relationships between sociocultural variables hypothesized as contributing to higher pain ratings in AA participants (discrimination and experimenter trust) were not consistent across pain rating and fMRI results. Frequency of responding to discrimination mediated ethnic group differences in pain rating. In contrast, frequency of responding to discrimination did not exhibit a statistically significant relationship with activity in regions exhibiting a stronger dose–response effect of painful heat in AA participants or any other regions in whole-brain analyses. Instead, average activity within the NAc and mPFC clusters showing stronger dose–response effects of painful heat in AA participants was positively related to frequency of experiencing discrimination and negatively related to experimenter trust (more strongly so in AA participants). Although these results suggest that sociocultural factors related to negative interpersonal experiences may contribute to heightened pain responses, larger studies may be required to definitively identify the strongest sociocultural predictors of pain sensitivity and their brain correlates.

Fourth, because participants were healthy young adults with similar socioeconomic status across ethnic groups, our findings may not reflect all of the factors that contribute to ethnic differences in pain in the general population. However, because we observed group differences in pain report consistent with the literature, our findings suggest potential brain mechanisms of these differences unconfounded by socioeconomic and health-related factors.

Finally, we did not find heightened pain report by HA participants compared to WA participants as has been found in some previous studies[69,70], and results from HA and WA participants were similar in most of our fMRI analyses. Hispanic American individuals vary widely in culture and background, and pain sensitivity in HA may vary substantially with the particular groups studied. Further exploration of sociocultural and neural factors contributing to pain report, particularly among HA, is needed in future studies.

Together, our findings support the hypothesis that higher levels of reported pain amongst AA compared to non-Hispanic WA may arise in part from differences in extra-nociceptive brain systems implicated in pain modulation, valuation and chronification, which may in turn result from the long-term effects of negative social treatment. Ethnic minorities, particularly AA, bear a disproportionate burden of pain and its negative health and financial consequences[4–6,71]. Our findings suggest that the higher levels of pain reported by AA in experimental and clinical settings probably reflect differences in the internal valuation of pain and its consequences for behaviour, and that interventions aimed at decreasing racial discrimination and increasing clinician trust amongst AA may help to alleviate these pain disparities.

Methods

Participants. Participants were 97 individuals (47 male, age 19–54 years; $M = 28.98$, s.d. $= 5.56$), 33 AA individuals (15 male), 32 non-Hispanic WA individuals (16 male) and 32 HA individuals (16 males), based on self-reported ethnicity. Participants were recruited from the greater Denver area through Craigslist or one of three different participant pools from the University of Colorado Boulder Institute for Behavioral Genetics (IBG) to capitalize on existing genetic data on these participants in future analyses. Participants reported no current or recent (past 6 months) neurological or psychiatric diagnosis and reported no current use of psychoactive or pain medications. Participants also reported no pain-related medical conditions, no reason to believe they would be especially sensitive or insensitive to contact heat and did not report currently experiencing an unusual amount of pain.

Nine participants were excluded from the present analyses for the following reasons: thermal stimulator error (4), claustrophobia (1), didn't fit in head coil (1), metallic thread in hair extensions (1), found pain intolerable (1), didn't meet demographic criteria (1). These exclusions resulted in a final sample of 88 participants (28 AA, 30 WA, 30 HA), age 19–54 years ($M = 28.82$, s.d. $= 5.67$) (see Table 1 for additional demographic details on final sample). This sample size is sufficient to detect ethnic/racial differences in the range of effects reported in

the neuroimaging and behavioural literature[6,23]. Based on Study 2 of ref. [23], the effect size of NPS responses (one of our primary neuroimaging outcome measures) for linear effects of increasing noxious stimulus intensity is Cohen's $d = 1.90$. This study is thus powered to detect racial/ethnic differences in NPS responses between AA participants and HA and WA participants, the racial contrast used in our analyses, of less than half that size ($d = 0.65$ or larger; $P < 0.05$ two-tailed; 80% power). In units of stimulus intensity, that effect size corresponds to 80% power to detect racial differences equivalent to about 0.50 °C. In whole-brain analyses at $P < 0.001$ (our approximate FDR-corrected threshold for $q < 0.05$), the study is powered to detect racial/ethnic differences that are larger ($d = 0.93$, 80% power) but still within reasonable range—that is, half the size of the NPS effect for linear effects of stimulus intensity. For reference, this effect size corresponds to the effects of a 0.71 °C change in stimulus intensity (benchmarking against the effect size of stimulus intensity on NPS responses). Finally, in terms of pain-rating effects, our sample size allows us to detect effect sizes of $d = 0.65$ or larger with 80% power, which is in the range of pain-rating differences between AA and non-Hispanic WA reported in a recent meta-analysis of 41 studies which averaged $d = 0.64$ for pain tolerance and $d = 0.46$ for pain intensity ratings[6].

The final sample of 88 participants was utilized in all analyses except those involving the self-report measures, in which case the sample was limited to participants with data for the self-report measure in question. See Supplementary Table 2 for number of participants with responses for each self-report measure. Ethnicity groups were matched on age, gender, recruitment source (Craigslist versus IBG) and fMRI sequence type (see Table 1 for statistics). The study was approved by the University of Colorado Boulder Institutional Review Board and we complied with all relevant ethical regulations when carrying out the study. Written informed consent was obtained from all participants, who were financially compensated for their participation.

Unused study components. Data reported here were collected as part of a larger study. The primary aim of the larger study was the same as the study reported here—that is, to understand sociocultural and neural mechanisms underlying ethnic disparities in pain. However, two other investigators included additional fMRI tasks, self-report measures and behavioural measures in the study to capitalize on the large and diverse sample being recruited for the primary study aims. One secondary aim was to explore the contribution of task-unrelated thoughts to pain experience, and to investigate how brain networks linked to experimentally induced pain are functionally organized at rest. Therefore, questions assessing task-unrelated thoughts were asked following each run of the pain task, in addition to a resting state run, and questionnaires assessing trait mind wandering, mindfulness, rumination and anxiety. Another secondary aim was to understand modulation of pain perception and its neural correlates by different psychological factors. To accomplish this aim, a pain-modulation fMRI task was included with thermal stimulations given before, during and after two of four tasks intended to modulate the pain experience: positive mood induction, negative mood induction, a stress induction or an n-back task. A last secondary aim was related to emotion processing independent of pain stimulation, for which participants watched a series of short emotional video clips. The task and analyses reported here were administered before the fMRI tasks related to the additional secondary aims in all participants and thus should not have influenced the results of the main task.

Self-report measures. To test potential psychological and sociocultural contributors to ethnic group differences in pain report related to our hypothesized mechanisms, all participants completed the following questionnaires online prior to their laboratory visit via Qualtrics: Barratt Simplified Measure of Social Status[72], Life Events Checklist[73], The Williams Major and Everyday Discrimination questions (WQ)[74], Brief History of Pain Questionnaire (an in-house measure with questions adapted from the McGill Pain Questionnaire[75] and painDETECT[76]), Trait Positive and Negative Affect Schedule[77], State and Trait Anxiety Inventory form X (STAI) state subscale[78], Penn State Worry Questionnaire[79], Pain Beliefs Questionnaire[80], Fear of Pain Questionnaire-III (ref. [81]), Kohn Reactivity Scale[82], Pain Catastrophizing Scale[83], Perceived Similarities Measure[84] and Wake Forest Physician Trust Scale (WFPTS)[85]. Note that questionnaires that were modified (WFPTS and PCS: modified to refer to the experimenter rather than a physician) or for which not all subscales were administered (STAI) may have different psychometric properties than those previously published. Questionnaires for which statistically significant group differences were found in the present analyses are described in detail below. In the same session, participants also completed additional questionnaires that pertained to other analyses not reported in the current paper. Additional questionnaires were intermixed with the questionnaires described below.

The WQ questions[74] asked whether participants had experienced nine different types of major discrimination—for example, being unfairly denied a bank loan (sum = major discrimination subscale score)—and ten different types of daily unfair treatment—for example, being treated with less courtesy than other people (sum = daily unfair treatment subscale score). We combined the major discrimination and daily unfair treatment subscales to create a total frequency of discrimination score, which we use in the present analyses. We did so by

multiplying the major discrimination subscale score by 3 and adding it to the daily unfair treatment subscale score, yielding a score ranging from 0 (no experience with discrimination) to 37 (highest experience with discrimination). Participants were also asked whether they had ever engaged in each of seven different responses to discrimination—for example, filing a complaint. We calculated a discrimination response subscale score by summing responses to yield a total score ranging from 0 (no history of responding to discrimination) to 7 (most extensive history of responding to discrimination).

The WFPTS[85] consists of ten statements about patients' trust in their physician—for example, 'Your doctor is extremely thorough and careful' and 'All in all, you have complete trust in your doctor'. We adapted these statements to apply to the experimenter, the same White male in his mid-30s for all participants for the full length of the ~4-h experimental session. Immediately after the scanning session, participants rated agreement with each statement on a scale from 1 (strongly agree) to 5 (strongly disagree). We reversed and summed responses to each measure resulting in a total score ranging from 10 (least trust in experimenter) to 50 (most trust in experimenter).

When missing data were present within these self-report measures, the missing response was replaced with the mean of that participant's responses on the given subscale if 50% or more of responses were present, otherwise the missing response was replaced with a 'not applicable' (NA) and treated as missing data (case-wise deletion). The average percentage of missing data (participants) per survey measure was 2.75% (s.d. = 2.94%, range = 0–8.0%).

fMRI task. The present task was always administered second in the scanning order (before the other pain tasks) following a high-resolution structural scan and 7-min resting state scan. During the thermal stimulation task, participants experienced painful thermal stimulations and provided ratings of the experience. The experimenter who collected the data could not be blinded to the race/ethnicity of research participants but was blind to our hypotheses regarding the effects of participant demographics. Data analysis was not performed blind to the conditions of the experiment.

Thermal stimulation. Thermal stimulation was delivered to four evenly spaced locations (one per run) on the volar surface of the left forearm using a $16 \times 16\,mm^2$ (model ATS) contact Peltier thermode (Medoc, Inc.). Thermal stimulation was delivered at three temperatures (47, 48 and 49 °C), all above the median temperature associated with reported pain in previous studies[23] and the activation of specific nociceptors[35] (>45 °C). All heat stimuli consisted of a sustained period of time (plateau) at the target temperature flanked by 1.7-s ramp periods to get to/from the target temperature to the 32 °C baseline. Heat stimuli were delivered with three different temporal profiles (heat conditions): Short, 8 s, 4.6-s plateau; Long, 11 s, 7.3-s plateau; and Offset, 11 s, 7.3-s plateau, with a 1-s, 1 °C temperature spike. For analyses in the present manuscript we either collapsed across or did not consider differences in heat conditions as they were not of interest here. Each heat trial was preceded by a cue, and all parts of the trial were separated by variable delays to allow for effective deconvolution of the BOLD signal associated with each trial element. See Fig. 1 for more details of the trial and task structure. Participants underwent a total of 36 heat trials, consisting of one trial at each temperature with each temporal profile in each of four runs. Trial order was randomized for each participant according to the following constraints: (1) trials within each temporal profile were evenly split between and randomly distributed within the first and second halves (Short and Long trials) or between thirds (Offset trials) of each run; (2) temperatures were then randomly distributed across trials within each temporal profile. At the start of each run, a single 49 °C long-duration (11-s) stimulus was delivered to allow for the initial habituation of the skin site to contact heat. This 'washout' stimulus was not used in the analyses. During pre-scan training and before each run, the participant was reminded that they could stop the task at any time if the pain became intolerable or for any other reason.

Pain rating. During each stimulation, participants were asked to continuously rate the intensity of the pain (not heat) they perceived on a 100-point generalized labelled magnitude scale[86] using an MRI-compatible trackball (Current Designs). The scale anchors were 0 (No Experience) to 100 (Strongest Imaginable Experience). Intermediate labels were placed as follows: 1.4 (Barely Detectable), 6 (Weak), 17 (Moderate), 35 (Strong) and 53 (Very Strong), though only labels and not numbers were visible to participants. The general anchors on the scale have been found to allow for effective comparison of sensory and affective experiences across modalities and people, and the label spacing has been found to provide the scale with ratio properties[86]. The AUC of the continuous within-trial pain intensity rating was used in the present analyses and is referred to throughout the manuscript as 'pain rating'. See Extended Data Fig. 1 for a graph of the raw continuous pain intensity-rating data. We have previously validated the use of continuous pain rating during the noxious stimulation period[23,33]. After a subset of stimulations (one stimulus at each temperature–duration combination), participants were also asked to rate the overall pain intensity and pain unpleasantness (order counterbalanced across trials) experienced on the previous trial using the same labelled magnitude scale as used for the continuous rating, referred to as the post-trial pain intensity and post-trial pain unpleasantness

ratings. During pre-scan task training we carefully described the distinction between intensity and unpleasantness ratings using the standard language developed by Price et al.[87], which describes the intensity ratings as 'how strong the pain feels' and the unpleasantness ratings as 'how unpleasant or disturbing the pain is'. As in Price et al.[87], we also used an analogy to the volume versus pleasantness of music, and emphasized that pain intensity and unpleasantness should be rated independently.

Additional task conditions. In addition, twenty-four 8-s trials of aversive sounds in two conditions were interspersed with the heat stimuli. The first aversive sound condition consisted of a physically aversive recording of nails on a chalkboard from a study of the psychoacoustics of aversive sounds[88] played at three different levels of intensity (5-Db steps). The second aversive sound condition consisted of a subset of emotionally aversive sounds (attacks, screaming and crying) from the International Affective Digital Sounds database[89] with the highest arousal and lowest pleasure. Intensity levels for aversive sounds were determined using the arousal–pleasure difference scores. Occurrences of the sound conditions were evenly distributed between, and randomly distributed within, the thirds of each run. These stimuli were not used in the present analysis and thus will not be described further here.

Pre-scan task training. Before the scanning session, participants were familiarized with the task by practising each condition without actual heat or sound stimulation. Instead, the experimenter asked the participant to imagine that the heat or sound was occurring as he/she practised making continuous ratings. In the scanner, participants were given an additional opportunity to practise the task (with imagined rather than real stimulation) to reinforce their understanding of the task and rating procedure within the scanner environment. Thus, it was not until the first trial of the actual MRI task that participants first experienced the heat or sound stimuli.

fMRI acquisition and preprocessing. *Data acquisition.* Data were collected on a 3 Tesla Siemens Trio MRI scanner at the University of Colorado Boulder Center for Innovation and Creativity. A high-resolution T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) structural scan ($1 \times 1 \times 1\,mm^3$ voxels, repetition time (TR): 2,530 ms, echo time (TE): 1.64 ms, flip angle: 7°, inversion time (Tl): 1,200 ms, field of view (FoV) read: 256 mm, echo spacing: 12.2 ms, bandwidth: 651 Hz Px^{-1}, time: 6:03) was performed on each participant to allow for normalization and display of functional data. During the four runs of the thermal pain task, either a multi-band (eight simultaneous slices) or standard (one slice at a time) echo-planar imaging sequence was performed. On the first 25 participants, the multi-band sequence was used ($3 \times 3 \times 3\,mm^3$ voxels, TR: 460 ms, TE: 29 ms, slices: 56, multi-band factor: 8, flip angle: 44°, FoV read: 248 mm, echo spacing: 0.51 ms, bandwidth: 2,772 Hz Px^{-1}, time: 10:15). Due to interference problems between the multi-band sequence and the thermal stimulator that arose after an update to the thermal stimulator software, the remaining 63 participants were scanned with a standard sequence ($3.4 \times 3.4 \times 3.4\,mm^3$ voxels, TR: 1,300 ms, TE: 25 ms, slices: 26, flip angle: 50°, FoV read: 220 mm, echo spacing: 0.55 ms, bandwidth: 2,170 Hz Px^{-1}, time: 10:15). To control for the potential effects of the difference in scanning sequence, scanning sequence is included as a covariate in brain imaging analyses.

Missing data. Several participants had partial fMRI data for the following reasons: thermode failed to deliver heat (1/4 runs (nine pain trials) for two participants, and a single trial for eight additional participants), scanning cessation due to finding pain intolerable (1/4 runs for one participant and 3/4 runs for another), scanning cessation due to claustrophobia (2/4 runs for one participant and 3/4 runs for another) and missing data due to scanner error (1/4 runs for one participant). These omissions resulted in a total of 11/352 (3.13%) runs being dropped from a total of 6/88 participants.

Preprocessing. The following preprocessing steps were applied to the brain imaging data before statistical analysis. The structural T1-weighted MPRAGE was co-registered to the mean functional image using an iterative mutual information algorithm in SPM8 with manual adjustment of the registration starting point. The MPRAGE was subsequently normalized to the MNI-152 template using SPM8. The initial images of every functional scan (standard sequence: 7, multi-band sequence: 20) were discarded to allow for stabilization of signal intensity. Image intensity outliers resulting from gradient and motion-related artefacts were identified and removed from the dataset before statistical analyses using the following procedure. First, we calculated the mean and standard deviation of image intensity values across all voxels within each slice. We then created a matrix of these values concatenated across slices within a volume by volumes across time. Next, the Mahalanobis distance of this matrix was calculated and images with a significant chi-square value (multiple-comparison correction using the stricter of FDR or Bonferroni) were identified as outliers and included as nuisance covariates in first-level statistical analysis (percentage motion outliers: $M = 4.80$, s.d. = 1.97, range = 1.92–12.05). Functional images were corrected for timing differences in slice acquisition (only for standard sequence) and realigned to the first image to

correct for head motion using SPM8. Functional images were warped to the MNI-152 template using warping parameters from the co-registered structural images. Finally, functional images were interpolated to $2 \times 2 \times 2\,mm^3$ and smoothed with an 8-mm FWHM Gaussian kernel.

To calculate a summary statistic of head motion we calculated average geometric displacement on each trial across the six motion parameters (X, Y, Z, pitch, roll, yaw) using the following procedure: (1) calculate the mean for each of the six motion parameters across all images included; (2) subtract the mean from each motion parameter, square each difference and sum results across all six to obtain one value per image; (3) take square root of the result from step 2 (this gives the distance from the mean for each image); and (4) calculate a mean across all images included. Average geometric displacement per heat pain trial across all participants was 0.14 mm (s.d. = 0.17). We tested for ethnic group differences in geometric displacement both on average and the dose–response relationship with stimulus intensity using a linear mixed-effects model in R, with the same parameters as the models used to test for group differences in pain rating described below. We found that AA participants (M = 0.21 mm, s.d. = 0.21 mm) moved significantly more on average ($t(84) = 2.90$, $P = 0.005$, $b = 0.1$, CI = [0.04, 0.17]) and had a steeper dose–response relationship with painful stimulus intensity ($t(84) = 4.29$, $P < 0.0001$, $b = 0.04$, CI = [0.02, 0.06]) than did WA (M = 0.11 mm, s.d. = 0.15 mm) and HA (M = 0.12 mm, s.d. = 0.12 mm) participants. No statistically significant movement differences existed between HA and WA participants (average: $t(84) = 0.15$, $P = 0.88$, $b = 0.003$, CI = [−0.04, 0.04]; temperature: $t(84) = 1.03$, $P = 0.31$, $b = 0.005$, CI = [−0.005, 0.02]).

African American participants both moved more and reported higher levels of pain on average, which was the phenomenon under study. Furthermore, in the single-trial data used for analysis (see details below) we found that participants who reported experiencing more pain also moved more ($t(84) = 10.62$, $P < 0.001$, $b = 0.002$, CI = [0.0015, 0.0022]) probably as a result of that pain. Therefore, we felt it was more appropriate to control for movement in our analyses and to exclude only the portion of data from each participant that actually showed evidence of movement contamination rather than to completely exclude participants with higher levels of movement, as the latter strategy would have resulted in excluding more AA participants than participants in the other groups, creating a confound to the interpretation of any findings of group differences.

To minimize the influence of head movement on our results, we realigned images and removed image intensity outliers as described above and included the six motion parameters as well as their mean-centred squares, derivatives and squared derivatives as nuisance regressors in the first-level fMRI analysis described below. To further rule out group differences in head motion as a confounding factor in our fMRI analyses, we repeated regression analyses reported on single-trial fMRI data controlling for single-trial average geometric displacement. This was not possible for mediation analyses due to intolerance of missing data in these models. Results, reported in Supplementary Tables 9–11, were largely unchanged when controlling for trial-by-trial geometric displacement.

Behavioural data analysis. *Pain-rating analysis.* Tests for ethnic group differences in pain rating were carried out using linear mixed-effects models in R[90] with the command lmer from the package lme4. Degrees of freedom for all effects in lmer models throughout the paper were estimated by subtracting the number of between-person model parameters from the number of participants as in ref. [91], which were then used to calculate corresponding P values (two-tailed) using the t distribution (pt function in R). Each of the three pain rating variables was used as a dependent measure in separate models. One model used post-trial ratings of pain intensity (nine per participant) as the dependent variable; another used post-trial ratings of pain unpleasantness (nine per participant); and the third model used the AUC of continuous within-trial pain intensity rating (36 per participant). The following factors were included in each model: (1) participant as a random effect; (2) participant ethnicity as two fixed-effect orthogonal contrasts: AA (coded as 0.68) compared to HA and non-Hispanic WA (each coded as −0.32) and HA (coded as 1) compared to non-Hispanic WA (coded as −1), as these contrasts represent hypothesized group differences in pain report based on previous studies[4–6]; (3) the linear effect of temperature as a fixed effect (47 °C (L), 48 °C (M) and 49 °C (H) coded as −1, 0, 1) with a random slope to account for between-participant differences in temperature response; (4) interactions of temperature with each ethnicity contrast to test for group differences in the temperature effect on pain rating; and (5) participant gender as a fixed effect to control for previously documented effects of gender on pain rating[92].

Psychological mediator analyses. We tested for potential psychological and sociocultural mediators of ethnic group differences in pain report using a two-stage process. The first stage was an exploratory analysis in which we tested for ethnic group differences in each of the 19 psychological and sociocultural self-report measures that paralleled ethnic group differences in pain rating, 'candidate mediators'. For these analyses we used linear models in R (command lm) with each participant's score on a given self-report measure as the dependent variable and the two orthogonal ethnicity contrasts used in the pain rating analyses as the predictors. We corrected for the 19 statistical tests by adjusting P values using Bonferroni correction.

In the second stage, we tested whether any of the candidate mediators identified in the first stage mediated the observed ethnic differences in pain report using a mediation analysis based on a three-variable path statistical model[93] using the Mediation Toolbox (https://github.com/canlab)[94–96]. In these analyses, participant ethnicity was the predictor (X) variable coded as 0.68 for AA participants and −0.32 for WA and HA participants, because this was the observed difference in pain rating we were trying to explain. Average participant pain rating was the outcome (Y) variable (three-temperature average or H–L temperature average) and the participant scores on the candidate psychological mediators meeting the above criteria served as the mediator (M) variable (one analysis per candidate mediator). We also controlled for participant gender and the HA (1) versus WA (−1) contrast as second-level covariates. We used bootstrapping for significance testing. We estimated distributions of participant-level path coefficients by randomly sampling with replacement 10,000 observations (rows) from the matrix of [a b c' c (a × b)] path coefficients. Bootstrapped 95% confidence intervals are presented with all mediation statistics, and two-tailed P values were calculated from the bootstrap confidence interval.

fMRI data analysis. *First-level analysis and robust regression.* We tested for group differences in average brain activity during pain (averaged across temperature levels) and brain responses to increases in painful heat (the temperature effect) using a standard GLM analysis. First-level GLM analyses were conducted using SPM8 to estimate individual participants' activation at each voxel. The four runs of the thermal stimulation task were concatenated for each participant. Boxcar functions representing the time courses of the different components of the task were convolved with a canonical haemodynamic response function and included as regressors in the first-level GLM. These task components included: (1) the period of heat or sound stimulation in each condition at each intensity level; (2) the cue period preceding each trial type; (3) the first heat 'washout trial' during each run; (4–5) the jittered pre- and post-stimulus rating periods; (6–7) the post-stimulus overall pain intensity or unpleasantness rating periods; (8) trials on which the thermode mistakenly did not deliver heat; and (9) the screen signalling the end of the task. The fixation cross epochs between trials, between trial components and at the beginning and end of the task were used as the implicit baseline. A high-pass filter of 224 s, which is well suited for pain of longer duration[33], was applied. The following regressors of non-interest (nuisance variables) were also included in the first-level model: (1) 'dummy' regressors representing each run (run intercepts); (2) regressors modelling linear drift across the duration of each run; (3) the six estimated head motion parameters (X, Y, Z, pitch, roll, yaw), their mean-centred squares, their derivatives and their squared derivatives for each run (24 columns in total); and (4) indicator vectors for signal intensity outliers (see description of outlier identification above). We entered two contrasts into the first-level analysis. The first contrast averaged across heat conditions and temperature levels compared to the resting baseline, to investigate average brain activity during pain. The second contrast compared average activity across all conditions at the highest temperature to average activity across all conditions at the lowest temperature (H–L), to investigate brain responses to increases in painful heat (the temperature effect).

Second-level (group) analyses were conducted using robust regression[97]. We compared average brain activity during pain and brain responses to increases in painful heat between ethnic groups using the same ethnicity contrasts as in pain rating and psychological mediator analyses. These contrasts were AA (coded as 0.68) compared to HA and non-Hispanic WA (each coded as −0.32) and HA (coded as 1) compared to non-Hispanic WA (coded as −1). As in other models, we also controlled for participant gender and type of scanning sequence (multi-band or standard). Results were thresholded using FDR $q < 0.05$ ($P < 0.000047$). For the purposes of display, we included voxels meeting two additional, more liberal, uncorrected voxel-wise thresholds: $P < 0.0005$ and $P < 0.001$ that were in contact with voxels meeting the more stringent FDR-corrected threshold.

Single-trial analysis. To estimate single-trial response magnitudes for use in NPS pattern expression analyses and analyses probing relationships between brain activity and pain rating, we employed a single-trial analysis approach as in refs. [98,99] by constructing a GLM design matrix that included one regressor for each trial. In this model, we included the same boxcar regressors representing each instance of the different task components as included in the standard GLM analysis above, as well as the same nuisance covariates. Additionally, we included a trial-specific regressor for the duration of the heat or sound stimulation period in each trial, also convolved with the haemodynamic response function. These single-trial beta images were used in NPS pattern expression and brain-versus-pain rating analyses. Because of the short duration of individual trials, single-trial estimates of brain activity can be strongly influenced by signal intensity artefacts caused by factors such as head motion. Therefore, we calculated trial-by-trial variance inflation factors, indices of the increase in variance of estimated regression coefficients due to multicollinearity with other predictor variables (in this case with nuisance regressors). Any trials with variance inflation factors >3.5 were excluded from the single-trial analyses.

To test the role played by group differences in brain responses to pain in mediating observed group differences in pain rating, we extracted single-trial parameter estimates from each cluster exhibiting significant group differences in

the whole-brain analysis. We then tested whether activity within these regions mediated the relationship between painful stimulus intensity and trial-by-trial pain rating, and whether this stimulus–brain–pain relationship differed between the AA and the HA and WA groups using a multi-level moderated mediation analysis. We used the same function in the Mediation Toolbox (https://github.com/canlab)[94–96] as we used to test for sociocultural mediators of group differences in pain. In these analyses, painful stimulus intensity on each trial was the predictor (X) variable, trial-by-trial pain rating was the outcome (Y) variable and brain responses in each region that responded differently to painful heat across ethnic groups in the GLM analysis served as the mediator (M) variable (one analysis per region). We tested for group differences in this mediation effect with a second-level moderator for participant ethnicity, comparing AA participants to WA and HA participants as in other analyses. We also controlled for participant gender, the HA (1) versus WA (−1) contrast and fMRI pulse sequence (standard or multi-band) as second-level covariates. We used bootstrapping for significance testing. We estimated distributions of participant-level path coefficients by randomly sampling with replacement 10,000 observations (rows) from the matrix of [a b c' c ($a \times b$)] path coefficients. Two-tailed P-values were calculated from the bootstrapped confidence interval. We also used a three-path mediation analysis (also Mediation Toolbox) to test whether the connection between the vmPFC and NAc clusters from the GLM, a pathway that has been particularly implicated in pain valuation[26,27] and chronification[28], mediated the relationship between painful stimulus intensity and trial-by-trial pain rating, and whether it did so differently in the AA and non-AA (WA and HA) groups. Because second-level moderators are not currently implemented in the three-path mediation algorithm in the Mediation Toolbox, we ran the three-path mediation analysis separately in AA and non-AA groups. This also meant we were not able to control for participant gender, the HA (1) versus WA (−1) contrast and fMRI pulse sequence in the three-path mediation models. Because of errors due to missing data, we were also not able to run a version of these mediation models controlling for geometric displacement.

We also tested the relationship between activity within a given cluster and pain rating and whether these brain–pain relationships differed by ethnic group without controlling for painful stimulus intensity. This analysis served as a complement to path b in the mediation analysis, which tested these brain–pain relationships controlling for painful stimulus intensity. We used separate linear mixed-effects models predicting activity within each region. Fixed factors in these models were fMRI sequence, gender, within-trial pain intensity rating, orthogonal ethnicity contrasts and the interaction between each ethnicity contrast and pain rating, while participant was a random factor.

Pattern expression analysis. We tested group differences in expression of the NPS pattern on average and in response to temperature. To do so, we calculated the strength of NPS pattern expression for each heat trial. We calculated a dot-product of each vectorized beta image ($\vec{\beta}_{map}$) with the NPS pattern (voxel-wise weight map; $\vec{W}_{map}$)—that is, $\vec{\beta}_{map}^{T} \vec{W}_{map}$, yielding a continuous, scalar pattern expression value. We then tested for ethnic group differences in average NPS pattern expression and NPS temperature response using a linear mixed-effects model similar to that used for pain rating, except with NPS pattern expression values as the dependent variable (fixed factors: temperature, gender, fMRI sequence, orthogonal ethnicity contrasts and their interaction with temperature; random factors: participant with random slope for temperature). As with the analyses with ROIs from the whole-brain analysis, we also tested the relationship between NPS pattern expression and pain rating without controlling for painful stimulus intensity to complement the NPS mediation analyses. We used the same model as in the brain–pain ROI analyses but with NPS pattern expression as the dependent variable.

Key null findings were followed up by estimating Bayes factors for t-tests using the BayesFactor R package (v.0.9.8), which uses the Jeffreys–Zellner–Siow prior (JZS) with a scale factor of 0.707 as in ref. [100]. The justification for the selection of the JZS prior for this mathematical derivation of Bayes factors for a one-sample t-test is detailed in ref. [100]. Briefly, the JZS prior is the combination of the Cauchy distribution on effect size and the Jeffreys prior on variance, which was selected to serve as an objective, non-informative prior for a one-sample t-test. To assess the sensitivity of our analyses to different priors, we performed our analyses using the other two priors available in the the BayesFactor R package, the Scaled-Information prior and the Unit-Information prior. Bayes factors calculated with all three priors were similar and all provided evidence in favour of the null in the same moderate range, based on categories outlined in ref. [101], suggesting that our analyses are robust to variation in priors. This Bayes factor calculation method did not allow us to calculate a credible interval or median for the posterior distribution, because it did not allow direct access to the posterior.

Self-report versus brain analyses. Finally, we tested whether the self-report measures we identified as candidate mediators of group differences in pain rating were related to (1) average activity or the dose–response relationship with painful stimulus intensity within brain ROIs exhibiting significant group differences in the GLM analysis, or (2) NPS pattern expression or the dose–NPS expression relationship. For these analyses we used linear models in R (command lm) with average values in a given ROI (vmPFC, mPFC, NAc and mFG) or NPS pattern expression, either across heat pain trials or during high-versus-low heat pain trials,

as dependent variables in separate models. We included one of the three candidate sociocultural mediators (discrimination frequency, response to discrimination or trust in experimenter) and its interaction with each of the two orthogonal ethnicity contrasts as predictors of interest, and participant gender and fMRI pulse sequence as control variables. Due to the exploratory nature of these follow-up analyses, we corrected for multiple comparisons across clusters (5: four ROIs + NPS), candidate mediators (3) and average activation versus the dose–response relationship with stimulus intensity (2) using Bonferroni correction (30 tests in total). For any main effect or interaction that survived multiple multiple-comparison correction, we conducted follow-up analyses within each ethnic group separately for which we did not correct for multiple comparisons.

All statistical tests presented in the manuscript are two-tailed unless otherwise noted. Unstandardized regression coefficients are presented as a measure of effect size. Confidence intervals for linear mixed-effects regression models were calculated using the confint() command from the MASS package in R with the profile method, which computes a likelihood profile and finds the appropriate cut-offs based on the likelihood ratio test. We visually examined the distribution of errors and statistical assumptions. In some cases, some deviation from homoscedasticity and normality of residuals was evident. Given that these deviations were relatively minor, we opted not to transform variables to maximize interpretability of findings. For voxel-wise maps, robust regression mitigated potential outliers and consequent potential violations of normality and homoscedasticity assumptions, which may vary across brain regions. For mediation analyses, the use of bootstrap tests mitigated potential issues with normality violations and homoscedasticity, as these do not rely on strong assumptions about the functional form of the error distribution.

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

Data availability

The data that support the findings of this study are available from the corresponding author on request.

Code Availability

The Mediation Toolbox used to conduct the mediation analyses can be freely accessed at https://github.com/canlab.

References

1. Trawalter, S. & Hoffman, K. M. Got pain? Racial bias in perceptions of pain. *Soc. Person. Psychol. Compass* **9**, 146–157 (2015).
2. Hoffman, K. M., Trawalter, S., Axt, J. R. & Oliver, M. N. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proc. Natl Acad. Sci. USA* **113**, 4296–4301 (2016).
3. National Center for Health Statistics (US). *Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities* (National Center for Health Statistics (US), 2016).
4. Anderson, K. O., Green, C. R. & Payne, R. Racial and ethnic disparities in pain: causes and consequences of unequal care. *J. Pain* **10**, 1187–1204 (2009).
5. Rahim-Williams, B., Riley, J. L., Williams, A. K. & Fillingim, R. B. A quantitative review of ethnic group differences in experimental pain response: do biology, psychology, and culture matter? *Pain Med.* **13**, 522–540 (2012).
6. Kim, H. J. et al. Racial and ethnic differences in experimental pain sensitivity: systematic review and meta-analysis. *PAIN* **158**, 194–211 (2017).
7. Anderson, S. R. & Losin, E. A. R. A sociocultural neuroscience approach to pain. *Culture Brain* **5**, 14–35 (2016).
8. Craig, K. D. Social communication model of pain. *PAIN* **156**, 1198–1199 (2015).
9. Bates, M. S. Ethnicity and pain: a biocultural model. *Soc. Sci. Med.* **24**, 47–50 (1987).
10. Campbell, C. M. et al. Ethnic differences in diffuse noxious inhibitory controls. *J. Pain* **9**, 759–766 (2008).
11. Campbell, C. M. et al. Ethnic differences in the nociceptive flexion reflex (NFR). *PAIN* **134**, 91–96 (2008).
12. Gelernter, J., Kranzler, H. & Cubells, J. Genetics of two mu opioid receptor gene (*OPRM1*) exon I polymorphisms: population studies, and allele frequencies in alcohol-and drug-dependent subjects. *Mol. Psychiatry* **4**, 476–483 (1999).
13. Williams, D. R. & Mohammed, S. A. Discrimination and racial disparities in health: evidence and needed research. *J. Behav. Med.* **32**, 20–47 (2009).
14. Williams, D. R. Race, socioeconomic status, and health. The added effects of racism and discrimination. *Ann. N.Y. Acad. Sci.* **896**, 173–188 (1999).

15. Hatch, S. L. & Dohrenwend, B. P. Distribution of traumatic and other stressful life events by race/ethnicity, gender, SES and age: a review of the research. *Am. J. Commun. Psychol.* **40**, 313–332 (2007).

16. Campbell, C. M., Edwards, R. R. & Fillingim, R. B. Ethnic differences in responses to multiple experimental pain stimuli. *PAIN* **113**, 20–26 (2005).

17. Cano, A., Mayo, A. & Ventimiglia, M. Coping, pain severity, interference, and disability: the potential mediating and moderating roles of race and education. *J. Pain* **7**, 459–468 (2006).

18. Burgess, D. J. et al. The effect of perceived racial discrimination on bodily pain among older African American men. *Pain Med.* **10**, 1341–1352 (2009).

19. Edwards, R. R. The association of perceived discrimination with low back pain. *J. Behav. Med.* **31**, 379–389 (2008).

20. Goodin, B. R. et al. Perceived racial discrimination, but not mistrust of medical researchers, predicts the heat pain tolerance of African Americans with symptomatic knee osteoarthritis. *Health Psychol.* **32**, 1117–1126 (2013).

21. Mathur, V. et al. Discrimination in health care settings is associated with greater clinical and laboratory pain in sickle cell disease. *J. Pain* **15**, S37 (2014).

22. Apkarian, A. V., Bushnell, M. C., Treede, R. D. & Zubieta, J. K. Human brain mechanisms of pain perception and regulation in health and disease. *Eur. J. Pain* **9**, 463–463 (2005).

23. Wager, T. D. et al. An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* **368**, 1388–1397 (2013).

24. Sokol-Hessner, P., Camerer, C. F. & Phelps, E. A. Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Soc. Cognit. Affect. Neurosci.* **8**, 341–350 (2013).

25. Hutcherson, C. A., Plassmann, H., Gross, J. J. & Rangel, A. Cognitive regulation during decision making shifts behavioral control between ventromedial and dorsolateral prefrontal value systems. *J. Neurosci.* **32**, 13543–13554 (2012).

26. Borsook, D., Upadhyay, J., Chudler, E. H. & Becerra, L. A key role of the basal ganglia in pain and analgesia – insights gained through human functional imaging. *Mol. Pain* **6**, 1–17 (2010).

27. Baliki, M. N., Geha, P. Y., Fields, H. L. & Apkarian, A. V. Predicting value of pain and analgesia: nucleus accumbens response to noxious stimuli changes in the presence of chronic pain. *Neuron* **66**, 149–160 (2010).

28. Baliki, M. N. et al. Corticostriatal functional connectivity predicts transition to chronic back pain. *Nat. Neurosci.* **15**, 1117–1119 (2012).

29. Dias-Ferreira, E. et al. Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* **325**, 621–625 (2009).

30. May, A. Neuroimaging: visualising the brain in pain. *Neurol. Sci.* **28** (Suppl. 2), S101–S107 (2007).

31. Woo, C. W. et al. Quantifying cerebral contributions to pain beyond nociception. *Nat. Commun.* **8**, 1–14 (2017).

32. Chang, L. J., Gianaros, P. J., Manuck, S. B., Krishnan, A. & Wager, T. D. A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biol.* **13**, e1002180 (2015).

33. Krishnan, A. et al. Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *eLife* **5**, 1–42 (2016).

34. Kross, E., Berman, M. G., Mischel, W., Smith, E. E. & Wager, T. D. Social rejection shares somatosensory representations with physical pain. *Proc. Natl Acad. Sci. USA* **108**, 6270–6275 (2011).

35. Craig, A. D., Krout, K. & Andrew, D. Quantitative response characteristics of thermoreceptive and nociceptive lamina I spinothalamic neurons in the cat. *J. Neurophysiol.* **86**, 1459–1480 (2001).

36. Taylor, A. B., MacKinnon, D. P. & Tein, J.-Y. Tests of the three-path mediated effect. *Organ. Res. Methods* **11**, 241–269 (2008).

37. Woo, C.-W., Roy, M., Buhle, J. T. & Wager, T. D. Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain. *PLoS Biol.* **13**, e1002036 (2015).

38. Leknes, S. et al. The importance of context: when relative relief renders pain pleasant. *PAIN* **154**, 402–410 (2013).

39. Goyal, M. K., Kuppermann, N., Cleary, S. D., Teach, S. J. & Chamberlain, J. M. Racial disparities in pain management of children with appendicitis in emergency departments. *JAMA Pediatrics* **169**, 996–1002 (2015).

40. Staton, L. J. et al. When race matters: disagreement in pain perception between patients and their physicians in primary care. *J. Nat. Med. Assoc.* **99**, 532–538 (2007).

41. Todd, K. H., Deaton, C., D'Adamo, A. P. & Goe, L. Ethnicity and analgesic practice. *Ann. Emergency Med.* **35**, 11–16 (2000).

42. Baliki, M. N. et al. Chronic pain and the emotional brain: specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain. *J. Neurosci.* **26**, 12165–12173 (2006).

43. Apkarian, A. V., Hashmi, J. A. & Baliki, M. N. Pain and the brain: specificity and plasticity of the brain in clinical chronic pain. *PAIN* **152**, S49–S64 (2011).

44. López-Solà, M. et al. Towards a neurophysiological signature for fibromyalgia. *PAIN* **158**, 34–47 (2017).

45. Kutch, J. J. et al. Brain signature and functional impact of centralized pain: a multidisciplinary approach to the study of chronic pelvic pain (MAPP) network study. *PAIN* **158**, 1979–1991 (2017).

46. Kornelsen, J. et al. Default mode network functional connectivity altered in failed back surgery syndrome. *J. Pain* **14**, 483–491 (2013).

47. Mays, V. M., Cochran, S. D. & Barnes, N. W. Race, race-based discrimination, and health outcomes among African Americans. *Annu. Rev. Psychol.* **58**, 201–225 (2007).

48. Bennett, G. G. et al. Stress, coping, and health outcomes among African-Americans: a review of the John Henryism hypothesis. *Psychol. Health* **19**, 369–383 (2004).

49. Clark, R., Anderson, N. B., Clark, V. R. & Williams, D. R. Racism as a stressor for African Americans: a biopsychosocial model. *Am. Psychol.* **54**, 808–816 (1999).

50. Williams, D. R. Race, socioeconomic status, and health the added effects of racism and discrimination. *Ann. N.Y. Acad. Sci.* **896**, 173–188 (1999).

51. McEwen, B. S. Stress adaptation, and disease: allostasis and allostatic load. *Ann. N.Y. Acad. Sci.* **840**, 33–44 (1998).

52. Zeiders, K. H., Doane, L. D. & Roosa, M. W. Perceived discrimination and diurnal cortisol: examining relations among Mexican American adolescents. *Hormones Behav.* **61**, 541–548 (2012).

53. McCallum, T., Sorocco, K. H. & Fritsch, T. Mental health and diurnal salivary cortisol patterns among African American and European American female dementia family caregivers. *Am. J. Geriatr. Psychiatry* **14**, 684–693 (2006).

54. Kapuku, G. K., Treiber, F. A. & Davis, H. C. Relationships among socioeconomic status, stress-induced changes in cortisol, and blood pressure in African American males. *Ann. Behav. Med.* **24**, 320–325 (2002).

55. Lee, D. B. et al. Psychological pathways from racial discrimination to cortisol in African American males and females. *J. Behav. Med.* **41**, 208–220 (2018).

56. Lampe, A. et al. Chronic pain syndromes and their relation to childhood abuse and stressful life events. *J. Psychosom. Res.* **54**, 361–367 (2003).

57. Brascher, A. K., Becker, S., Hoeppli, M. E. & Schweinhardt, P. Different brain circuitries mediating controllable and uncontrollable pain. *J. Neurosci.* **36**, 5013–5025 (2016).

58. Lee, J. et al. Encoding of self-referential pain catastrophizing in the posterior cingulate cortex in fibromyalgia. *Arthritis Rheumatol.* **70**, 1308–1318 (2018).

59. Wager, T. D., Atlas, L. Y., Leotti, L. A. & Rilling, J. K. Predicting individual differences in placebo analgesia: contributions of brain activity during anticipation and pain experience. *J. Neurosci.* **31**, 439–452 (2011).

60. Williams, D. R. & Rucker, T. D. Understanding and addressing racial disparities in health care. *Health Care Finan. Rev.* **21**, 75–90 (2000).

61. Woodrow, K. M., Friedman, G. D., Siegelaub, A. B. & Collen, M. F. Pain tolerance: differences according to age, sex and race. *Psychosom. Med.* **34**, 548–556 (1972).

62. Edwards, C. L., Fillingim, R. B. & Keefe, F. Race, ethnicity and pain. *PAIN* **94**, 133–137 (2001).

63. U.S. Census Bureau. *Population Profile of the United States: Dynamic Version.* 1–9 (US Census Bureau, 2005).

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

65. Ma, Y. et al. Serotonin transporter polymorphism alters citalopram effects on human pain responses to physical pain. *Neuroimage* **135**, 186–196 (2016).

66. Han, S. Understanding cultural differences in human behavior: a cultural neuroscience approach. *Curr. Opin. Behav. Sci.* **3**, 68–72 (2015).

67. Han, S. in *Brain Mapping: An Encyclopedic Reference* Vol. 3 (ed. Toga, A. W.) 217–220 (Elsevier, 2015).

68. Losin, E. A. R. Culture, brain, and health: introduction to the special issue. *Cult. Brain* **5**, 1–3 (2017).

69. Rahim-Williams, F. B. et al. Ethnic identity predicts experimental pain sensitivity in African Americans and Hispanics. *PAIN* **129**, 177–184 (2007).

70. Portenoy, R. K., Ugarte, C., Fuller, I. & Haas, G. Population-based survey of pain in the United States: differences among white, African American, and Hispanic subjects. *J. Pain* **5**, 317–328 (2004).

71. Shavers, V. L., Bakos, A. & Sheppard, V. B. Race, ethnicity, and pain among the U.S. adult population. *J. Health Care Poor Underserv.* **21**, 177–220 (2010).

72. Barratt, W. *The Barratt Simplified Measure of Social Status (BSMSS): Measuring SES* http://socialclassoncampus.blogspot.com/2012/06/barratt-simplified-measure-of-social.html (2006).

73. Gray, M. J., Litz, B. T., Hsu, J. L. & Lombardo, T. W. Psychometric properties of the life events checklist. *Assessment* **11**, 330–341 (2004).

74. Williams, D. R., Yan, Y., Jackson, J. S. & Anderson, N. B. Racial differences in physical and mental health: socio-economic status, stress and discrimination. *J. Health Psychol.* **2**, 335–351 (1997).

75. Melzack, R. The McGill Pain Questionnaire: major properties and scoring methods. *PAIN* **1**, 277–299 (1975).

76. Freynhagen, R., Baron, R., Gockel, U. & Tolle, T. R. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. *Curr. Med. Res. Opin.* **22**, 1911–1920 (2006).

77. Watson, D., Clark, L. A. & Tellegen, A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Personal. Soc. Psychol.* **54**, 1063–1070 (1988).

78. Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R. & Jacobs, G. A. *Test Manual for the State Trait Anxiety Inventory* (Consulting Psychologists Press, 1983).

79. Meyer, T. J., Miller, M. L., Metzger, R. L. & Borkovec, T. D. Development and validation of the Penn State Worry Questionnaire. *Behav. Res. Ther.* **28**, 487–495 (1990).

80. Edwards, L. C., Pearce, S. A., Turner-Stokes, L. & Jones, A. The Pain Beliefs Questionnaire: an investigation of beliefs in the causes and consequences of pain. *PAIN* **51**, 267–272 (1992).

81. McNeil, D. W. & Rainwater, A. J. 3rd Development of the Fear of Pain Questionnaire—III. *J. Behav. Med.* **21**, 389–410 (1998).

82. Kohn, P. M. Sensation-seeking, augmenting-reducing, and strength of the nervous system. *Motivation, Emotion, and Personality: Proc. XXIII International Congress of Psychology* 167–173 (Elsevier, 1985).

83. Sullivan, M. J., Bishop, S. R. & Pivik, J. The Pain Catastrophizing Scale: development and validation. *Psychol. Assess.* **7**, 524–532 (1995).

84. Street, R. L., O'Malley, K. J., Cooper, L. A. & Haidet, P. Understanding concordance in patient–physician relationships: personal and ethnic dimensions of shared identity. *Ann. Fam. Med.* **6**, 198–205 (2008).

85. Hall, M. A. et al. Measuring patients trust in their primary care providers. *Med. Care Res. Rev.* **59**, 293–318 (2002).

86. Bartoshuk, L. M., Fast, K. & Snyder, D. J. Differences in our sensory worlds: invalid comparisons with labeled scales. *Curr. Dir. Psychol. Sci.* **14**, 122–125 (2005).

87. Price, D. D., Mcgrath, P. A., Rafii, A. & Buckingham, B. The validation of visual analog scales as ratio scale measures for chronic and experimental pain. *PAIN* **17**, 45–56 (1983).

88. Reuter, C. & Oehler, M. Psychoacoustics of chalkboard squeaking. *J. Acoust. Soc. Am.* **130**, 2545–2545 (2011).

89. Bradley, M. M. & Lang, P. J. *The International Affective Digitized Sounds (IADS-2): Affective Ratings of Sounds and Instruction Manual* (Univ. of Florida, 2007).

90. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).

91. Lindquist, M. A., Spicer, J., Asllani, I. & Wager, T. D. Estimating and testing variance components in a multi-level GLM. *Neuroimage* **59**, 490–501 (2012).

92. Greenspan, J. D. et al. Studying sex and gender differences in pain and analgesia: a consensus report. *PAIN* **132**, S26–S45 (2007).

93. Baron, R. M. & Kenny, D. A. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Personal. Soc. Psychol.* **51**, 1173–1182 (1986).

94. Wager, T. D. et al. Brain mediators of cardiovascular responses to social threat: part I: reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity. *Neuroimage* **47**, 821–835 (2009).

95. Atlas, L. Y., Bolger, N., Lindquist, M. A. & Wager, T. D. Brain mediators of predictive cue effects on perceived pain. *J. Neurosci.* **30**, 12964–12977 (2010).

96. Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A. & Ochsner, K. N. Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* **59**, 1037–1050 (2008).

97. Wager, T. D., Keller, M. C., Lacey, S. C. & Jonides, J. Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage* **26**, 99–113 (2005).

98. Koyama, T., McHaffie, J. G., Laurienti, P. J. & Coghill, R. C. The single-epoch fMRI design: validation of a simplified paradigm for the collection of subjective ratings. *Neuroimage* **19**, 976–987 (2003).

99. Rissman, J., Gazzaley, A. & D'Esposito, M. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* **23**, 752–763 (2004).

100. Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D. & Iverson, G. Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychonom. Bull. Rev.* **16**, 225–237 (2009).

101. Lee, M. D. & Wagenmakers, E.-J. *Bayesian Cognitive Modeling: A Practical Course* (Cambridge Univ. Press, 2013).

Acknowledgements

Author contributions

E.A.R.L. and T.D.W. designed the study. E.A.R.L., J.R.A-H. and H.E. collected the data. E.A.R.L., C-W.W., T.D.W. and N.A.M. analysed the data. E.A.R.L. wrote the manuscript and E.A.R.L and C-W.W. created the figures, both with substantial input from the other authors.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-020-0819-8.

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-020-0819-8.

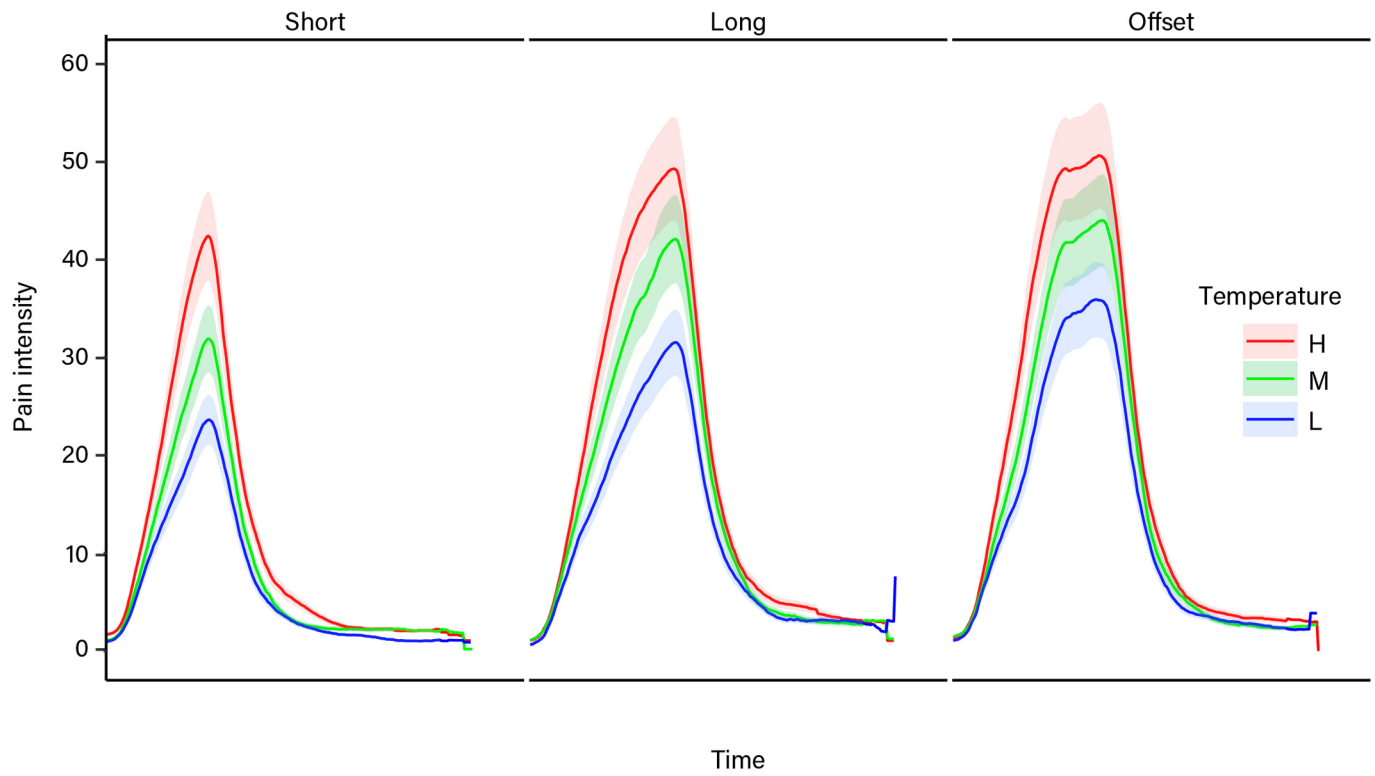**Correspondence and requests for materials** should be addressed to E.A.R.L. or T.D.W.

Peer review information Primary Handling Editors: Mary Elizabeth Sutherland and Stavroula Kousta

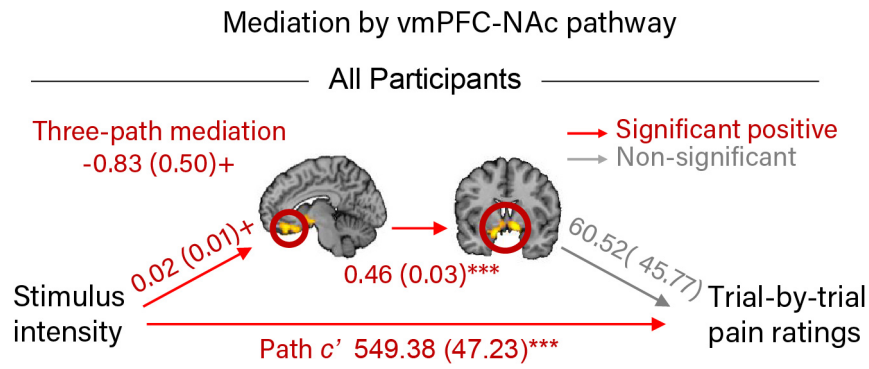**Reprints and permissions information** is available at www.nature.com/reprints.

Within-trial continuous pain rating by heat condition

Extended Data Fig. 1 | Average continuous pain intensity ratings for the three stimulus intensity levels by heat condition. Error bands are between-participant standard error.

Mediation by vmPFC-NAc pathway

Extended Data Fig. 2 | Path diagrams and statistics for three-path, multi-level mediation analysis between painful stimulus intensity, connectivity between the vmPFC and NAc clusters from the GLM, and trial-by-trial pain rating across all participants. Connectivity between the vmPFC and NAc regions from the GLM exhibited a trend (p <0.1) towards mediating the relationship between painful stimulus intensity and pain rating. Path coefficients are listed for each path with standard errors in parentheses.

nature research

Corresponding author(s):   Elizabeth R. Losin

Last updated by author(s):   Sep 17, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see <u>Authors & Referees</u> and the <u>Editorial Policy Checklist</u>.

Statistics

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

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

Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.

Software and code

Policy information about <u>availability of computer code</u>

Data collection	MATLAB was used for stimulus delivery during fMRI aquisiton and Qualtrics was used for survey data collection.
Data analysis	SPM8 was used to analyze MRI data. The Mediation Toolbox (https://github.com/canlab) was used to perform mediation analysis, and R statistical analysis software, version 3.6.1, was used to perform all other statistical analyses in the paper.

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

Data

Policy information about <u>availability of data</u>

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

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

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

Life sciences study design

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

Sample size	Our a priori sample size goal for the study was N = 90 (30 per ethnic group).
Data exclusions	Criteria for excluding participants were any reason that resulted in minimal or no neuroimaging data or other violations of inclusion criteria (e.g. demographics) discovered after study enrollment. These criteria are described in Methods: Participants, para 2. The original sample size in the study was 97 participants. Nine participants were excluded from analyses for the following reasons: thermal stimulator error (4), claustrophobia (1), didn't fit in head coil (1), metallic thread in hair extensions (1), found pain intolerable (1), didn't meet demographic criteria (1). These exclusions resulted in a final sample of 88 participants. This is described in Methods: Participants, para 2.
Replication	Our findings of higher pain reports in African American participants compared to non-Hispanic white participants replicate many previous findings in the experimental and clinical pain literature. Likewise, our findings that the NPS tracks painful stimulus intensity replicate the results of 9 other studies. However, we did not attempt replication of any of our findings within the current study.
Randomization	Comparison groups of interest in the present study were racial/ethnic groups based on participants self report, therefore, random group assignment was not used. Criteria for group assignment are described in the Methods: Participants, para 1.
Blinding	Because groups in the present study were based on participants' race/ethnicity, it is likely that the experimenter was aware of participant group membership to some extent. However, the experimenter was not informed of the participant's self-reported racial or ethnic identification. This is not currently stated in the manuscript, but can be added.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	Participants were 97 individuals (47 male, age 19-54 years old; M = 28.98, SD = 5.56), 33 African-American (AA) subjects (15 male), 32 Non-Hispanic White American (WA) subjects (16 male), and 32 Hispanic White American (HA) subjects (16 males), based on self-reported ethnicity. Participants reported no current or recent (past 6 months) neurological or psychiatric diagnosis and reported no current use of psychoactive or pain medications. Participants also reported no pain-related medical conditions, no reason to believe they would be especially sensitive or insensitive to contact heat, and did not report currently experiencing an unusual amount of pain.
Recruitment	Participants were recruited from the greater Denver area through Craigslist or one of three different subject pools from the University of Colorado Boulder Institute for Behavioral Genetics (IBG) in order to capitalize on existing genetic data on these participants in future analyses.
Ethics oversight	The study was approved by the University of Colorado Boulder IRB.

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

Magnetic resonance imaging

Experimental design

Design type	Slow epoch single-trial design.
Design specifications	Each participant experienced 36 heat trials (9 per run). Additionally, each participant had 3 trials each of two sound

Design specifications	conditions per run, not analyzed in current manuscript. Thermal stimulation in each trial was either 8 or 11 seconds, while sound stimulation was always 8 seconds. Trials also included cue and rating periods with each part of the trial and each trial separated by jittered interval fixation crosses ranging from 1-5.5 seconds in duration.
Behavioral performance measures	Participants provided several different ratings of the intensity and unpleasantness of the pain they experienced as a result of the thermal stimulation during the fMRI task. We tested for the linear effects of stimulus intensity on pain rating in order to verify participants were responding as expected.

Acquisition

Imaging type(s)	Functional and structural
Field strength	3T
Sequence & imaging parameters	(MPRAGE) structural scan (1×1×1 mm voxels, TR: 2530 ms, TE 1: 1.64 ms, Flip angle: 7°, Tl: 1200 ms, FoV Read: 256 mm, echo spacing 12.2 ms, bandwidth: 651 hz/px, time: 6:03).
	Functional scans (4 run per participant) were either multi-band or standard EPI sequences.
	Multiband (first 25 ppts): (3×3×3 mm voxels, TR: 460 ms, TE: 29 ms, slices: 56, multiband factor=8, flip angle: 44°, FoV read: 248 mm, echo spacing: 0.51 ms, bandwidth: 2772 Hz/Px, time: 10:15).
	Standard (last 63 participants) due to interference with thermal stimulator: (3.4×3.4×3.4 mm voxels, TR: 1300 ms, TE: 25 ms, slices: 26, flip angle: 50°, FoV read: 220 mm, echo spacing: 0.55 ms, bandwidth: 2170 Hz/Px, time: 10:15).
Area of acquisition	The slices in our functional (EPI) sequences covered the whole brain in many participants and most of the brain in the remainder of participants. Slices were perscribed at an oblique angle that passed under the frontal pole and base of the cerebellum. Thus, when whole brain coverage was not possible, the far upper portion of the parietal lobe was excluded. This is standard practice in our laboratory based on brain regions of interest in pain studies.
Diffusion MRI ☐ Used ☒ Not used	

Preprocessing

Preprocessing software	Most preprocessing steps were performed using algorithms in SPM8 implemented with batch processing code written in MATLAB. Functional images were corrected for timing differences in slice acquisition (only for standard sequence) and realigned to the first image to correct for head motion using SPM8. Functional images were interpolated to 2x2x2 mm, and smoothed with 8 mm FWHM Gaussian kernel, also using SPM8. See below for additional reprocessing steps.
Normalization	T1-weighted MPRAGE was co-registered to mean functional image using an iterative mutual information algorithm in SPM8 with manual adjustment of the registration starting point (linear transformation). The MPRAGE was subsequently normalized to the MNI-152 template using SPM8 (nonlinear transformation). Functional images were warped to the MNI-152 template using warping parameters from the co-registered structural images.
Normalization template	MNI-152
Noise and artifact removal	The following regressors of non-interest (nuisance variables) were also included in the first level model: 1) "dummy" regressors representing each run (run intercepts); 2) regressors modeling linear drift across the duration of each run; 3) the six estimated head motion parameters (X, Y, Z, pitch, roll, yaw), their mean-centered squares, their derivatives, and their squared derivatives for each run (24 columns total); and 4) indicator vectors for signal intensity outliers (see description of outlier identification below).
Volume censoring	The initial images of every functional scan (standard sequence: 7, multiband sequence: 20) were discarded to allow for stabilization of signal intensity. Image intensity outliers resulting from gradient and motion-related artifacts were identified and removed from the data set prior to statistical analyses using the following procedure. First, we calculated the mean and standard deviation of image intensity values across all voxels within each slice. We then created a matrix of these values concatenated across slices within a volume x volumes across time. Next, the Mahalanobis distance of this matrix was calculated and images with a significant chi-square value (multiple comparison correction using the more strict of FDR or Bonferroni) were identified as outliers and included as nuisance covariates in first level statistical analysis. % motion outliers: M = 4.80, SD = 1.97, Range = 1.92-12.05.

Statistical modeling & inference

Model type and settings	A mass univariate approach was used for whole-brain analyses with estimation of fixed effects at the first level and random effects at the second level. A time series correlation model was not used. This approach yields valid results as the population level and avoids the negative effects of selecting an inappropriate autocorrelation model or poorly estimating the autocorrelation model.
	Expression of a previously derived multivariate pattern (the neurologic pain signature, NPS) was computed by taking dot-product of each vectorized beta image for each trial from the first level model (see details of single trial analysis in next section) with the NPS pattern, yielding a continuous scalar pattern expression value.
Effect(s) tested	A stimulus type (3 heat conditions 2 sound conditions) x intensity level (low, medium, high) factorial design was used. The first level model contained regressors for the following task components 1) the period of heat or sound stimulation

in each condition at each intensity level, 2) the cue period preceding each trial type, 3) the first heat "washout trial" during each run, 4-5) the jittered pre-stimulus and post-stimulus rating periods, 6-7) the post-stimulus overall pain intensity or unpleasantness rating periods, 8) trials on which the thermode mistakenly did not deliver heat, and 9) the screen signaling the end of the task and 10) nuisance regressors described above. The fixation cross epochs in between trials, in between trial components, and at the beginning and end of the task were used as the implicit baseline. We entered two contrasts into the first level analysis. The first contrast averaged across heat conditions and temperature levels compared to the resting baseline, to investigate average brain activity during pain. The second contrast compared average activity across all conditions at the highest temperature to average activity across all conditions at the lowest temperature (H-L), to investigate the dose-response relationship between stimulus intensity and brain activation (the temperature effect).

Second-level (group) analyses were conducted using robust regression. We compared average brain activity during pain and brain responses to increases in painful heat between ethnic groups using the same ethnicity contrasts as in pain rating and psychological mediator analyses. These contrasts were African Americans (coded as .68) compared to Hispanic and non-Hispanic whites (each coded as -.32) and Hispanics (coded as 1) compared to non-Hispanic whites (coded as -1). As in other models, we also controlled for participant gender, and type of scanning sequence (multiband or standard).

We also employed a single-trial analysis approach as in by constructing a GLM design matrix that included one regressor for the duration of the heat or sound stimulation period in each trial, also convolved with the hemodynamic response function. The beta images for each trial from this analysis were used in the multivariate pattern (NPS) expression analysis described in the previous section and in 2-path and 3-path multi-level mediation analyses between stimulus intensity (X), brain regions showing dose-response effects of painful heat that differed between ethnic groups in the GLM analysis or the NPS (M), and pain rating (Y).

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s) | Follow-up analyses to whole-brain analyses were conducted using functionally derived ROIs from the whole brian analysis.

Statistic type for inference (See Eklund et al. 2016) | Voxel-wise FDR

Correction | FDR correction is used at a threshold of q < .05

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis | Expression of a previously derived multivariate pattern (the neurologic pain signature, NPS) is explored by taking dot-product of each vectorized beta image with the NPS pattern, yielding a continuous scalar pattern expression value. Derivation of the multivariate pattern is described in: Wager, T. D. et al. An fMRI-based neurologic signature of physical pain. New England Journal of Medicine 368, 1388-1397 (2013).