

ARTICLE

<https://doi.org/10.1038/s41467-019-11934-y>

OPEN

# Different brain networks mediate the effects of social and conditioned expectations on pain

Leonie Koban <sup>1,2,3,4</sup>, Marieke Jepma<sup>5</sup>, Marina López-Solà<sup>6</sup> & Tor D. Wager <sup>1,2,7</sup>

Information about others' experiences can strongly influence our own feelings and decisions. But how does such social information affect the neural generation of affective experience, and are the brain mechanisms involved distinct from those that mediate other types of expectation effects? Here, we used fMRI to dissociate the brain mediators of social influence and associative learning effects on pain. Participants viewed symbolic depictions of other participants' pain ratings (social information) and classically conditioned pain-predictive cues before experiencing painful heat. Social information and conditioned stimuli each had significant effects on pain ratings, and both effects were mediated by self-reported expectations. Yet, these effects were mediated by largely separable brain activity patterns, involving different large-scale functional networks. These results show that learned versus socially instructed expectations modulate pain via partially different mechanisms—a distinction that should be accounted for by theories of predictive coding and related top-down influences.

<sup>1</sup>Institute of Cognitive Science, University of Colorado Boulder, Muenzinger D244, 345 UCB, Boulder, CO 80302, USA. <sup>2</sup>Department of Psychology and Neuroscience, University of Colorado Boulder, Muenzinger D244, 345 UCB, Boulder, CO 80302, USA. <sup>3</sup>Brain and Spine Institute (ICM), Control-Interoception-Attention Team, 47 Boulevard de l'Hôpital, 75013 Paris, France. <sup>4</sup>Marketing Area, INSEAD, Boulevard de Constance, 77300 Fontainebleau, France. <sup>5</sup>Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129B, 1018 WS Amsterdam, The Netherlands. <sup>6</sup>Division of Behavioral Medicine and Clinical Psychology, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, University of Cincinnati, 3333 Burnet Avenue, MLC2 7031 Pain Research Center, Cincinnati, OH 45229, USA. <sup>7</sup>Department of Psychological and Brain Sciences, Dartmouth College, HB 6207, Moore Hall, Hanover, NH 03755, USA. Correspondence and requests for materials should be addressed to L.K. (email: [leonie.koban@colorado.edu](mailto:leonie.koban@colorado.edu))

Expectations and beliefs shape human experience and behavior in many important ways. They are thought to serve as priors—initial belief distributions—in Bayesian and predictive coding frameworks<sup>1,2</sup>, shaping sensory perception via feedback projections<sup>3</sup> and guiding action and decision-making<sup>2</sup>. They are core features of persistent mindsets that shape clinical symptoms, physiology, and other life outcomes<sup>4</sup>. Prominent examples illustrating the power and complexity of expectations are placebo and nocebo effects—changes in subjective experience and objective physiology due in part to expectations about medical treatment and symptom relief<sup>5–7</sup>.

While expectations and related forms of prior information are often treated as a unitary concept, they can be based on different sources. In many cases, expectations are based on what we have learned from our own prior experience, via classical conditioning or other forms of associative learning. In this case, prior expectations can help to stabilize the perception of otherwise noisy or ambiguous input<sup>8,9</sup>. However, expectations can also stem from secondary sources, such as what others tell us about their experiences. Vicarious experiences communicated via symbols or words—termed here “social information”—allow us to form priors about experiences we have never had. Several studies demonstrate that information from both vicarious and direct experiences influence feelings and decisions, and may even have synergistic effects on learning and behavior<sup>10–16</sup>.

However, it remains unclear whether similar or different neural systems mediate direct experience-driven and vicarious influences. Experimental studies often create “top-down” effects via a combination of associative learning (i.e., classical conditioning) and social information (e.g., instructions by an experimenter or doctor), without trying to separate their sources<sup>5–7,17–22</sup>. Thus, the commonalities and differences in the neurophysiological pathways involved remain largely unexplored<sup>15</sup>. Here, we create expectations about pain based on two different sources, social information and associative learning, in order to dissociate their underlying brain mechanisms.

While prediction and predictive coding are thought to be central to information representation in multiple neural circuits throughout the brain<sup>2,7,23</sup>, there is also evidence for specific systems representing expectations<sup>24</sup>. Key regions include the ventromedial prefrontal cortex (vmPFC)<sup>25</sup> and dorsolateral prefrontal cortex (dlPFC)<sup>26,27</sup>, which are thought to have central roles in decision-making<sup>28</sup>, instruction effects<sup>16,29–32</sup>, and top-down biasing of information flow based on context<sup>33,34</sup>. Further, expectation effects on perceptual decision-making have been shown to involve the hippocampus<sup>35</sup>, in line with a role for this region in both retrospective memory and prospective thought<sup>36,37</sup>. Similarly, placebo and nocebo effects, which typically combine both suggestions and reinforcement to induce expectations, are thought to operate via recruitment of dlPFC and vmPFC and their interactions with subcortical regions<sup>6,38–41</sup>. Yet, differential effects of social suggestions and learning have not been disentangled.

Here, we employed a recently developed experimental paradigm that allowed us to independently manipulate conditioned cue information and social information<sup>42</sup>. In each of 96 trials of a learning task, 38 participants were presented with two types of information (Fig. 1a). In order to investigate associative learning effects on pain, they were shown one of two conditioning stimuli (CS, drawings of an animal and vehicle), which were predictive of low-to-medium (CS<sub>LOW</sub>, 50% 48 °C and 50% 49 °C) or medium-to-high (CS<sub>HIGH</sub>, 50% 49 °C and 50% 50 °C) thermal pain stimulation. Independent of these cues, in order to investigate social instruction effects, we presented participants with the pain ratings of ten other fictive participants (Fig. 1a). This social information depicted either low (Social<sub>LOW</sub>) or high pain ratings (Social<sub>HIGH</sub>).

However, unbeknown to the participants, social cues were not predictive of actual thermal stimulation. Participants indicated how much pain they expected, then received a brief thermal stimulation, and rated experienced pain at the end of each trial (Fig. 1b).

We employed multilevel mediation to (1) characterize changes in pain-related brain activity for Social<sub>HIGH</sub> > Social<sub>LOW</sub> trials and CS<sub>HIGH</sub> > CS<sub>LOW</sub> trials and to (2) identify brain mediators of these experimental manipulations on pain ratings. We tested two competing predictions: if associative learning and social information effects have common brain mechanisms, brain effects should be found in similar regions. Alternatively, if the effects of conditioning and social information have separable underlying mechanisms, brain effects should be seen in distinct areas, such as “mentalizing” and prefrontal areas for social information effects and associative learning-related regions, such as amygdala and hippocampus for conditioned effects on pain.

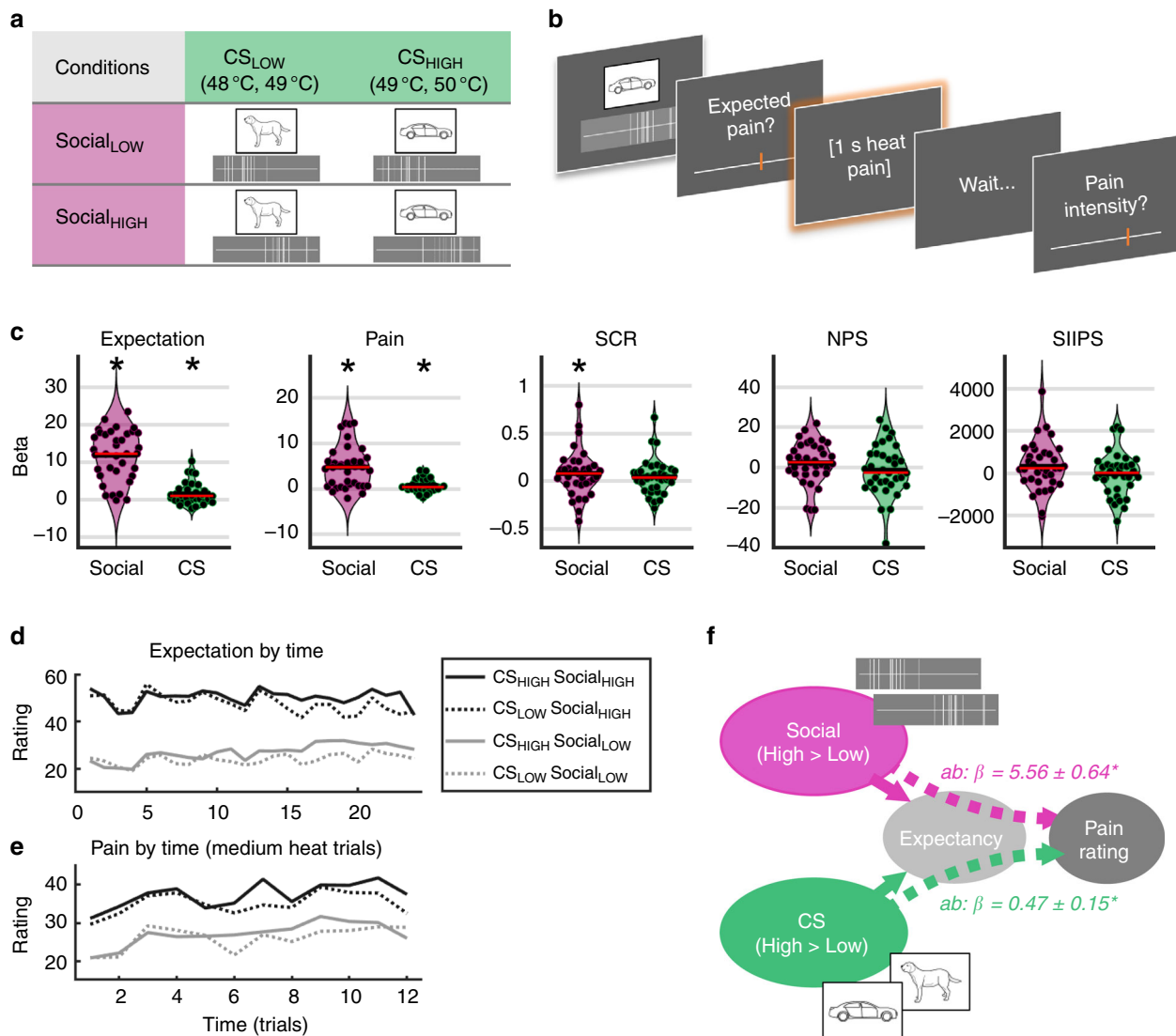
Our results show that both types of effects are mediated by consciously reported expectations, but their brain mediators are largely different. Limbic and posterior brain regions mediate conditioning effects, while prefrontal and parietal regions more strongly mediate social information effects on pain. Together, these results suggest that top-down modulation of experience can stem from distributed sources in frontoparietal, limbic, and default mode areas, depending on the source of information. Expectations induced by social influence and instructions might be especially powerful in shaping perception and experience, by bypassing learning networks and directly engaging higher-level prefrontal and parietal association cortices.

## Results

**Behavior and skin conductance.** Higher stimulus temperatures (50 °C > 49 °C > 48 °C) significantly increased pain ratings ( $\beta = 7.55$  [95% CI 5.88–9.22],  $t(35) = 8.85$ ,  $p < 0.001$ , Cohen’s  $d = 1.50$ ). Consistent with our predictions, both social information and CS indicative of high (vs. low) pain increased ratings of pain expectations (Fig. 1c; Social<sub>HIGH</sub>–Social<sub>LOW</sub>:  $\beta = 11.75$  [9.59, 13.92],  $t(35) = 10.65$ ,  $p < 0.001$ , Cohen’s  $d = 1.80$ ; CS<sub>HIGH</sub>–CS<sub>LOW</sub>:  $\beta = 1.45$  [0.51, 2.38],  $t(35) = 3.02$ ,  $p = 0.005$ , Cohen’s  $d = 0.51$ ). The effects of social information on expectations were much stronger than those of the CS ( $t(35) = 7.66$ ,  $p < 0.001$ , Cohen’s  $d = 1.28$ ). The Social effect on expectations decreased slightly over time (cue  $\times$  time interaction,  $\beta = -0.05$  [–0.08, –0.03],  $t(35) = -4.57$ ,  $p < 0.001$ , Cohen’s  $d = 0.77$ ), but was present throughout the experiment (Fig. 1d). In contrast, CS effects on expectations only emerged over time, reflecting learning (cue  $\times$  time interaction,  $\beta = 0.04$  [0.01, 0.06],  $t(35) = 3.02$ ,  $p = 0.004$ , Cohen’s  $d = 0.51$ ).

Both types of information strongly influenced pain as well (Fig. 1c): heat at the same, medium-level intensity (49 °C) was rated as more painful in the Social<sub>HIGH</sub> compared with the Social<sub>LOW</sub> condition ( $\beta = 4.78$  [3.27, 6.29],  $t(35) = 6.19$ ,  $p < 0.001$ , Cohen’s  $d = 1.05$ ), and in the CS<sub>HIGH</sub> compared with the CS<sub>LOW</sub> condition ( $\beta = 0.74$  [0.28, 1.20],  $t(35) = 3.13$ ,  $p = 0.003$ , Cohen’s  $d = 0.53$ ). In parallel to expectations, pain ratings were more strongly influenced by social information than CS cues ( $t(35) = 4.86$ ,  $p < 0.001$ , Cohen’s  $d = 0.81$ ). While CS effects on pain became stronger with time (reflecting learned pain modulation,  $\beta = 0.02$  [0.0007, 0.038],  $t(35) = 2.03$ ,  $p = 0.049$ , Cohen’s  $d = 0.34$ ), the social influence on pain ratings remained stable over time (i.e., cue  $\times$  time interaction:  $p > 0.20$ , Fig. 1e).

Expectations formally mediated both Social and CS effects on pain (path *ab* effect for Social,  $\beta = 5.56$  [4.31, 6.81],  $t(35) = 8.56$ ,  $p < 0.001$ , Cohen’s  $d = 1.47$ ; for CS,  $\beta = 0.47$  [0.18, 0.76],  $t(35) =$



**Fig. 1** Design and results. **a** Design. The social information could be either low or high on average (Social<sub>LOW</sub> or Social<sub>HIGH</sub>), but was not correlated with actual heat intensity. In contrast, the learning cues (CS<sub>LOW</sub> and CS<sub>HIGH</sub>) were followed by low-to-medium or medium-to-high heat intensity, respectively. **b** Each trial started with the simultaneous presentation of social information (“pain ratings of previous participants”, depicted as vertical lines on a VAS) and one of the CS (animal or vehicle drawing). Participants then rated their pain expectation, received a short heat pain stimulation, and rated how much pain they experienced. **c** Effects on behavior, physiology, and brain patterns. Violinplots show the effects for Social<sub>HIGH</sub> > Social<sub>LOW</sub> and CS<sub>HIGH</sub> > CS<sub>LOW</sub>. Each dot reflects the beta (effect magnitude) estimate of one participant. Both social information and CS significantly influenced expectation (social information:  $t(35) = 10.65, p < 0.001$ , CS:  $t(35) = 3.02, p = 0.005$ ) and pain ratings (social:  $t(35) = 6.19, p < 0.001$ , CS:  $t(35) = 3.13, p = 0.003$ ). Skin conductance responses (SCR) were significantly modulated by social information ( $t(35) = 2.04, p = 0.049$ ), but not by CS. Neither NPS nor SIIPS showed significant responses to social information or CS. Asterisks denote significant effects at  $p < 0.05$  (using  $t$  tests). **d, e** Time course of expectation and pain ratings. The difference between dotted and solid lines reflects the CS effect, and the difference between gray and black lines the social information effect. CS effects on expectation and pain increased over time (interaction effects CS\*Time on expectation:  $t(35) = 3.02, p = 0.004$ , on pain:  $t(35) = 2.03, p = 0.049$ ). Social information effects on expectations and pain remained significant throughout the experiment, but decreased over time for expectation ratings (Social  $\times$  Time on expectations:  $t(35) = -4.57, p < 0.001$ ). The x-axis shows trials per condition. **f** Behavioral mediation analysis. Expectation ratings significantly mediated both Social ( $t(35) = 8.56, p < 0.001$ ) and CS effects on pain ( $t(35) = 3.27, p < 0.001$ ). Source data for panels **c, d,** and **e** are provided as a Source Data file

3.27,  $p < 0.001$ , Cohen’s  $d = 0.53$ , Fig. 1f; see also Supplementary Fig. 1), replicating our previous findings<sup>42</sup>.

Paralleling the behavioral findings, skin conductance responses (SCR) to pain (Fig. 1c; Supplementary Fig. 2) were higher in Social<sub>HIGH</sub> than Social<sub>LOW</sub> trials ( $\beta = 0.08$  [0.003, 0.16],  $t(35) = 2.04, p = 0.049$ , Cohens’  $d = 0.34$ ). SCR were only numerically higher for CS<sub>HIGH</sub> compared to CS<sub>LOW</sub> trials ( $t(35) = 1.47, p = 0.15$ ). In contrast to our earlier findings, expectation ratings did

not significantly mediate social information and CS effects on SCR ( $t$  test, both  $p$ ’s > 0.20), potentially due to noisier physiological data in the MRI environment.

**Effects on established multivariate brain measures related to pain.** We next tested for effects on two previously described multivariate brain markers of pain. The neurological pain

signature (NPS<sup>43</sup>) was optimized and validated to predict pain ratings with high accuracy during evoked pain related to multiple types of noxious stimuli. It is thought to reflect primarily pain of nociceptive origin<sup>43</sup>, as it does not respond to several forms of psychological pain modulation<sup>44,45</sup>. The stimulus-intensity independent pain signature (SIIPS1)<sup>46</sup> was developed to predict pain independent of nociceptive input and mediates effects on pain of several psychological manipulations, including cued expectancy and perceived control<sup>46</sup>. Here, while single-trial fMRI responses in both NPS and SIIPS were significantly associated with trial-by-trial pain ratings (NPS:  $\beta = 0.032$  [0.015, 0.049],  $t(35) = 3.69$ ,  $p < 0.001$ , Cohens'  $d = 0.62$ ; SIIPS:  $\beta = 0.0002$  [0.00008, 0.0004],  $t(35) = 3.06$ ,  $p = 0.004$ , Cohens'  $d = 0.52$ ), neither of the two patterns was significantly modulated by either the social information or the CS ( $t$  tests, all  $p$ 's  $> 0.1$ , see Fig. 1c). This suggests that social information and learning effects on pain may not influence primary nociceptive processing, but exert their effects via other mechanisms.

**Whole brain mediation analysis.** We estimated two multilevel-mediation models<sup>17,41,47</sup> to assess the mechanisms of social information and associative learning on pain-related brain activity modeled on a single-trial level (see Fig. 2a). In brief, Social<sub>HIGH</sub> versus Social<sub>LOW</sub> and CS<sub>HIGH</sub> versus CS<sub>LOW</sub> served as predictors (coded 1 and -1 for high and low, respectively), and single-trial pain ratings served as the outcome. Trial-by-trial beta images for medium-temperature pain events (matched on stimulus intensity) were included as mediators. CS cues served as first-level covariates in the Social→Brain→Pain model, and vice versa for the CS→Brain→Pain model, so that both models controlled for both cue effects. Path *a* effects describe changes in activity based on the manipulation (similar to a classic 2nd level contrast), i.e., Social<sub>HIGH</sub> > Social<sub>LOW</sub> and CS<sub>HIGH</sub> > CS<sub>LOW</sub>. Path *b* characterized brain activity that predicts pain ratings independent of the two experimental manipulations. Path *ab* (the *a*\**b* product or “mediation effect”) characterized brain activity that significantly mediates the effect of the manipulations (social information and CS) on pain ratings. An extended overview of mediation is provided in Supplementary Fig. 1.

**Cue-independent contributions to pain.** We first identified brain areas associated with higher pain ratings independent of social information and CS cues (Path *b*,  $q < 0.05$  FDR-corrected, yellow in Fig. 2b). These included areas typically involved in pain processing, such as the dorsal anterior cingulate cortex (ACC) and premotor cortex, mid and posterior insulae, somatosensory cortex (leg area, corresponding to the stimulation site), ventrolateral thalamus, and cerebellum, as well as occipital (visual) cortex.

In order to assess how these regions map onto large-scale networks, we calculated the spatial pattern similarity (Pearson correlation coefficients) of the mediation maps to seven cortical resting-state networks<sup>48</sup>. The wedge plot in Fig. 2b shows these similarity values, with the radius of each wedge proportional to the correlation strength and area proportional to variance explained. Pain-predictive activity independent of manipulations was concentrated in somatomotor, ventral attention, and visual networks (see Supplementary Tables 1, 6).

We further calculated the similarity between the Path *b* map and association test  $z$ -score brain maps of 525 terms from the Neurosynth meta-analytic database<sup>49</sup>. While relationships with terms should be interpreted with caution and are only suggestive, they can be useful for contextualizing findings and provide a more fine-grained comparison with existing large-scale databases than the canonical seven-network parcellation. The top ten

unique terms associated with pain-related (path *b*) effects were themed around pain processing (e.g., “noxious”, “heat”, “pain”, “painrelated”) and somatomotor function (“foot”, “limb”, “sensation”, “somatosensory”, “muscle”, “voluntary”, see Supplementary Table 7). Thus, meta-analysis-based decoding here seemed related to the stimulation to the right calf muscle.

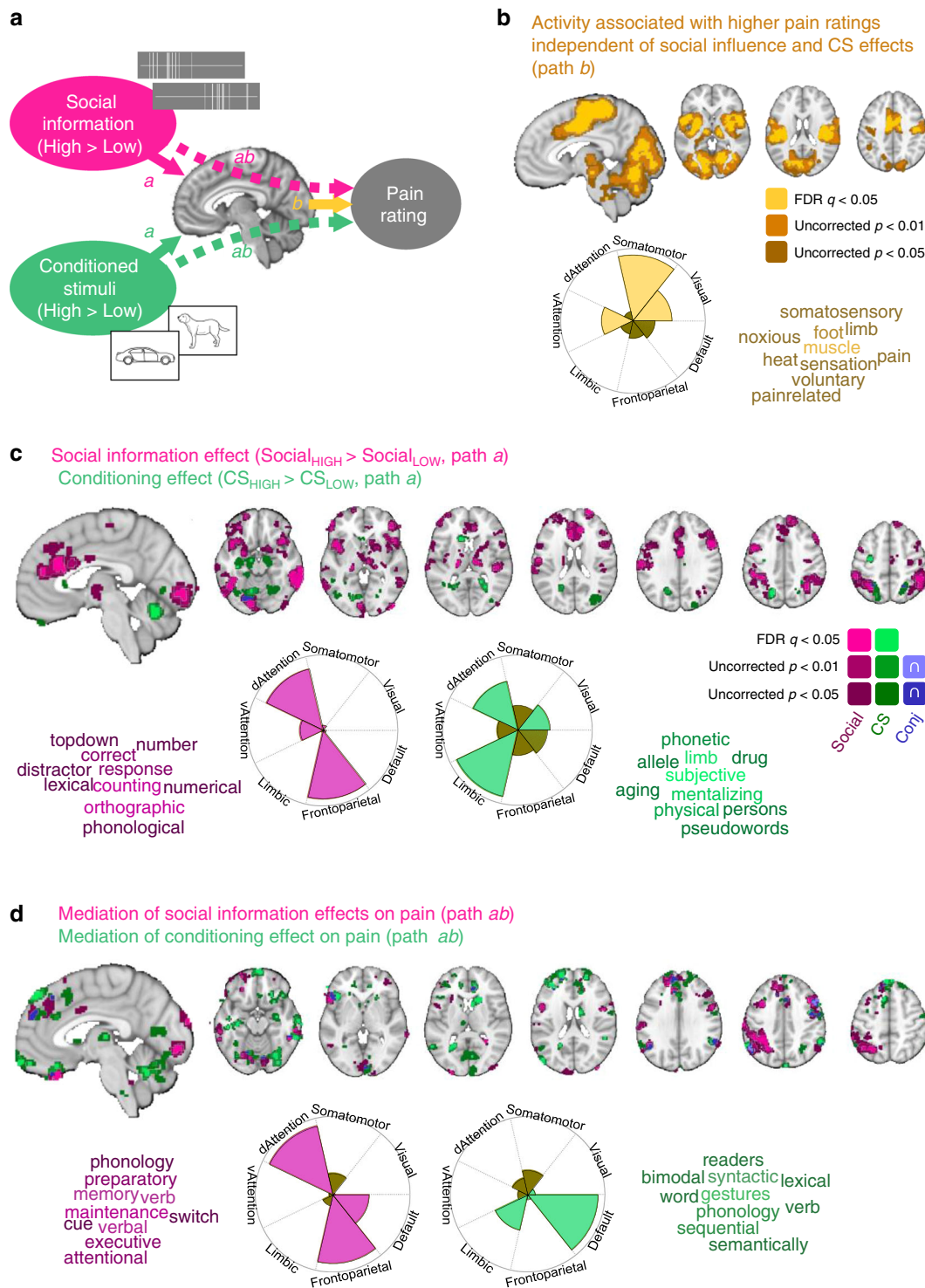
**Social information effects on pain.** Increased activity for the Social<sub>HIGH</sub> compared to the Social<sub>LOW</sub> condition was found in several cortical and subcortical areas (path *a* effects, purple in Fig. 2c, FDR  $q < 0.05$  corrected, Supplementary Table 2). First, they comprised areas associated with cognitive control and top-down attention, including dlPFC, inferior parietal lobule (IPL), and intraparietal sulcus (IPS). Second, they included brain areas associated with salience, affect, and pain processing, i.e., ACC, anterior insula and operculum, and ventrolateral thalamus. Third, increased activity was found in brain areas involved in somatosensory integration, such as mid insula and parietal operculum. In addition, inferior temporal and visual cortex showed increased activation for Social<sub>HIGH</sub> > Social<sub>LOW</sub>.

These social information effects on pain were spatially correlated mostly with the dorsal attention and frontoparietal networks (see Supplementary Table 6). Paralleling the network findings, the top ten decoding results for Social<sub>HIGH</sub> > Social<sub>LOW</sub> using Neurosynth meta-analytic maps included terms associated with attention and cognitive control (e.g., “topdown”, “distractor”) and with word and number processing (e.g., “counting”, “orthographic”, “number”, “lexical”, Supplementary Table 7). Thus, though some regions identified are reliably involved in social cognition, we do not have strong reasons to believe that the processes engaged are uniquely social; they may reflect effects of control and attention on pain more generally.

Mediators of social information effects on pain ratings (path *ab*, purple in Fig. 2d; Supplementary Table 3) included a similar set of regions, especially vlPFC, dorsomedial prefrontal cortex (dmPFC), dlPFC, IPL/IPS, visual cortex, inferior temporal sulcus, orbitofrontal cortex, and visual cortex. In close correspondence to path *a* effects, the networks most prominently involved in mediation were the dorsal attention and frontoparietal networks, with a smaller contribution of the default mode network (see Supplementary Table 6). The top ten decoding terms based on Neurosynth included words associated with cognitive control and attention (e.g., “memory”, “maintenance”, “switch”, “executive”, “attentional”, see Supplementary Table 7).

We next used a conjunction analysis to identify spatial overlap in brain areas responding to the Social<sub>HIGH</sub> > Social<sub>LOW</sub> manipulation (path *a*) and those mediating the effects of this manipulation on pain outcomes (path *ab*). The mediation effect is driven by a combination of responses to the experimental manipulation, correlations with pain, and correlations in individual differences between these two effects (i.e., stronger responses to the Social<sub>HIGH</sub> > Social<sub>LOW</sub> manipulation correlated with stronger effects of the brain region on pain). Thus, the *ab* effect encompasses all the elements required to link the manipulation, brain, and behavior, but does not guarantee that all significant regions show significant cue effects. The conjunction analysis therefore shows regions that show both Social<sub>HIGH</sub> > Social<sub>LOW</sub> effects and mediation effects (purple in Fig. 3a). It confirmed the important contribution of the dlPFC, IPS, dmPFC, vlPFC, and visual cortex to social information effects on pain. Higher single-trial activity averaged over this set of regions was correlated with higher single-trial expectation ratings ( $\beta = 0.86$  [0.33, 1.40],  $t(35) = 3.15$ ,  $p = 0.003$ , Cohen's  $d = 0.53$ ), suggesting that these areas were involved in the generation of explicit expectations based on the social





information. Individual differences in how much social information influenced expectation ratings did not moderate this relationship between brain activity and expectations ( $t$  test,  $p > 0.10$ ).

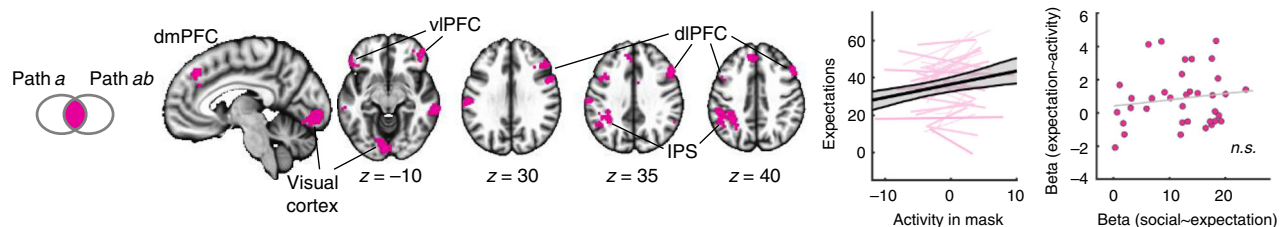
**Learning effects on pain.** Increased activity for CS<sub>HIGH</sub> compared with CS<sub>LOW</sub> was found in a largely different set of brain regions (path *a*, green in Fig. 2c), including “limbic” areas such as hippocampus, caudate, and retrosplenial cortex, as well as cerebellum, precentral gyrus, and left IPS. These activations mapped

weakly onto a combination of limbic, dorsal attention and visual networks (Fig. 2c; Supplementary Tables 4, 6). Accordingly, the top ten Neurosynth associations for the CS<sub>HIGH</sub> > CS<sub>LOW</sub> map spanned over various topics including language (“pseudowords”, “phonetic”) and mentalizing (“persons”, “mentalizing”) (see Supplementary Table 7).

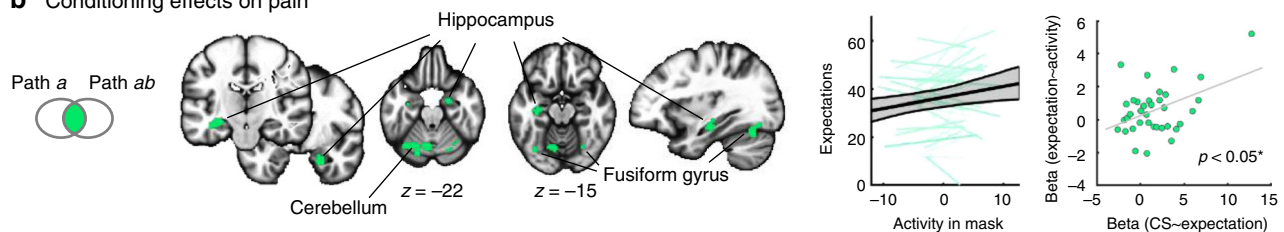
Significant mediation of learning effects on pain (path *ab*) was seen in many areas, including dmPFC, medial and lateral orbitofrontal cortex, left anterior insula/operculum, caudate, hippocampus, retrosplenial cortex, fusiform gyrus, dPFC, IPL, and cerebellum (green in Fig. 2d; Supplementary Table 5). Some

**Fig. 2** Whole brain mediation analysis. **a** Overview. A mass-univariate mediation analysis was performed to identify: (1) activity increases for high > low social information and CS (Path *a*), (2) activity associated with increased pain ratings when controlling for path *a* effects (path *b*), and (3) activity formally mediating the effects of social information and CS on pain ratings (dashed arrows, path *ab*). **b** Path *b* effects. Significant pain-related activity independent of the experimental manipulations was found in mid cingulate, posterior and mid insula, thalamus, cerebellum, and other regions. The wedge plot (with the radius of each wedge proportional to the correlation strength) indicates a high spatial correlation of this effect with the Somatomotor network (Pearson correlation coefficient  $r = 0.28$ , see Supplementary Table 6). The ten most strongly associated Neurosynth terms are shown on the right (decreasing brightness indicates order of associations, see Supplementary Table 7). **c** Path *a* effects for Social (purple), CS effects (green), and their conjunction (blue). Social information effects (increased activity for  $\text{Social}_{\text{HIGH}} > \text{Social}_{\text{LOW}}$ ) were found in ACC, anterior insula, dlPFC, and parietal areas. Those effects most strongly mapped on the frontoparietal ( $r = 0.13$ ) and dorsal attention networks ( $r = 0.12$ , see wedge plot and Supplementary Table 6) and were associated with terms reflecting cognitive tasks (see Supplementary Table 7). CS effects ( $\text{CS}_{\text{HIGH}} > \text{CS}_{\text{LOW}}$ ) were seen in limbic areas and cerebellum, and showed a more diffuse mapping on large-scale networks and meta-analytic terms. **d** Path *ab* effects for Social (purple), CS effects (green), and their conjunction (blue). Social influence effects on pain mapped on the frontoparietal and dorsal attention networks (both  $r$ 's = 0.06) and on terms associated with cognitive control. CS effects mapped on the default mode network ( $r = 0.06$ ) and were associated with terms related to semantic processing. All maps were thresholded at  $\text{FDR } q < 0.05$  corrected for multiple comparisons across the whole brain (gray matter masked) with adjacent areas thresholded at  $p < 0.01$  and  $p < 0.05$  (uncorrected) for display

### a Social information effects on pain



### b Conditioning effects on pain



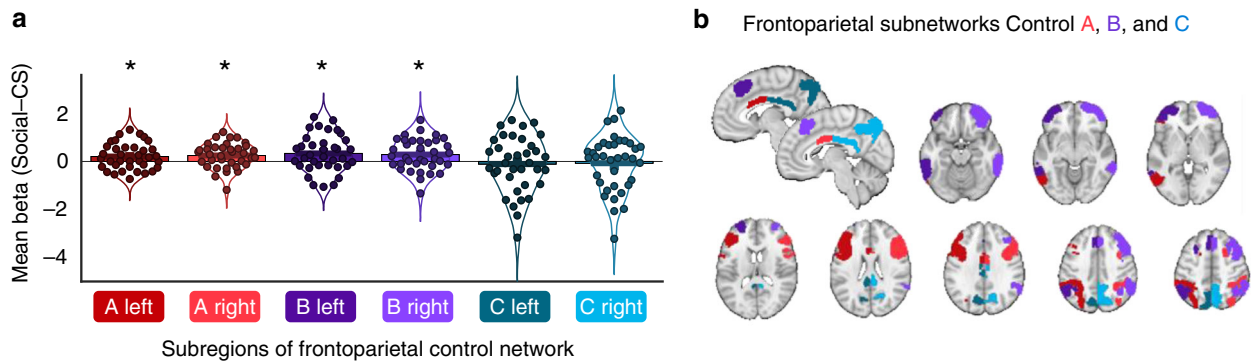
**Fig. 3** Conjunction of path *a* and *ab* effects. **a** Social information effects were consistent across paths *a* and *ab* in dmPFC, vIPFC, dlPFC, IPS, and visual cortex. Trial-wise average activity (betas) in these regions correlated significantly with trial-wise expectation ratings (individual slopes in purple). Average activity did not significantly predict individual differences in the social influence effects on expectations (scatter plot). **b** CS effects were consistent across paths *a* and *ab* in the hippocampus, cerebellum, and fusiform gyrus. Trial-wise average activity in these areas correlated significantly with trial-wise expectation ratings (individual slopes in green). Individual differences in the strength of the activation were correlated with individual differences in how much expectations ratings were driven by the CS. Conjunction effects are displayed as the intersection of activation for paths *a* and *ab*, each of them thresholded at  $P < 0.05$  FDR-corrected and adjacent voxels at  $p < 0.05$  uncorrected. Asterisk reflects significant Pearson correlation coefficient ( $p < 0.05$ ). Shaded error bands reflect bootstrapped 95% confidence intervals. Source data are provided as a Source Data file

of those effects partially overlapped or were neighboring with mediation effects of social influence. Yet, learning effects were mostly correlated with default and limbic networks and with meta-analytic maps associated with semantic processing (e.g., “lexical”, “word”, “semantically”, “gestures”, see Supplementary Tables 6, 7). As noted above, this does not strongly imply that these processes were involved, but suggests similarity to regions engaged in associative learning, including semantic associations.

We again used a conjunction analysis to illustrate the overlap in brain regions that showed increased activation to pain for  $\text{CS}_{\text{HIGH}} > \text{CS}_{\text{LOW}}$  (path *a* effects), and regions that statistically mediated the effects on pain ratings (path *ab* effects). This conjunction revealed three main regions of bilateral effects: hippocampus, fusiform gyrus, and cerebellum (green in Fig. 3b). In parallel to the social mediator regions, higher activity in these areas again correlated with higher trial-by-trial expectation ratings ( $\beta = 0.58$  [0.09, 1.06],  $t(35) = 2.34$ ,  $p = 0.025$ , Cohen's  $d = 0.39$ ; Fig. 3b, middle). This association was modulated by individual differences in how much expectations were driven by the learning cues ( $\beta = 0.21$  [0.07, 0.35],  $t(34) = 2.84$ ,  $p = 0.008$ ,

Cohen's  $d = 0.49$ ; Fig. 3b, right)—however, this relationship with individual differences was largely driven by a single participant (~3 STD above the mean on both brain and behavior). Larger sample sizes are needed to investigate individual differences in more detail. In sum, these results show the contribution of brain regions involved in memory, learning, and object recognition in learned pain modulation and (as with unconditioned social information) suggests a role for conscious (reportable) expectation in this process.

**Cue-related and interaction effects.** For completeness, we tested how social information and CS influenced brain activity at the time of cue presentation, again demonstrating largely different systems responding to each type of cue and mediating their effects on expectations (see Supplementary Fig. 3). We further explored interaction effects between social and CS conditions, shown in Supplementary Fig. 4. Broadly, these analyses confirmed the importance of frontoparietal, orbitofrontal, and temporal systems in pain expectancy, with ventral and dorsal systems



**Fig. 4** Difference between Social and CS effects in frontoparietal control subnetworks. **a** Difference in mean path *a* beta weight (Social–CS) in Control A (dark and bright red), Control B (dark and bright purple), and Control C (dark and bright blue) subnetworks in the left and right hemisphere (dark and bright colors, respectively). Each dot reflects the difference in mean beta estimates of one participant. Significantly greater activation for social information was found in the Control A (left:  $t(35) = 2.3$ ,  $p = 0.027$ , 95% CI = [0.03, 0.39], Cohen’s  $d = 0.38$ ; right:  $t(35) = 2.9$ ,  $p = 0.006$ , 95% CI = [0.08, 0.43], Cohen’s  $d = 0.48$ ) and the Control B (left:  $t(35) = 2.6$ ,  $p = 0.012$ , 95% CI = [0.08, 0.59], Cohen’s  $d = 0.44$ ; right:  $t(35) = 2.6$ ,  $p = 0.013$ , 95% CI = [0.07, 0.51], Cohen’s  $d = 0.44$ ), but not in the Control C network. Asterisks denote networks with significant differences between social information and CS effects (using *t*-tests,  $p < 0.05$ ). Source data are provided as a Source Data file. **b** Display of frontoparietal control subnetworks A (dark and bright red), B (dark and bright purple), and C (dark and bright blue) on sagittal and transversal brain slices

preferentially involved in associative learning and social information effects, respectively.

**Similarity versus separability of social influence and learning effects.** To test for commonalities between social information and learning effects, we performed a conjunction analysis (using a conjunction null<sup>50</sup>). At a lenient threshold ( $p < 0.05$  uncorrected voxels adjacent to FDR-significant voxels), CS and social influence showed a few small clusters of common path *a* effects in visual cortex and bilateral superior parietal lobule (see Fig. 2c). To quantify the degree of overlap vs. dissociation, we computed the Dice coefficient for voxels activated at 0.05 uncorrected. The Dice coefficient—which can theoretically range between 0 and 1, where 0 reflects complete separation and 1 reflects perfect overlap—for path *a* was 0.024, thus suggesting relatively little overlap between path *a* activations for Social and CS effects. For path *ab*, shared mediation effects of social influence and CS effects on pain were observed at  $p < 0.01$  and  $p < 0.05$  uncorrected, notably in dmPFC, dlPFC, vlPFC, and IPL (see Fig. 2d), with a Dice coefficient of 0.081. Thus, while these clusters of conjunction effects were relatively small in size, this suggests that parts of the frontoparietal network may be involved in top-down modulation of pain based on expectations irrespective of their sources. In sum, mediation maps suggest a dissociation in the regions mediating both learned and instructed cue effects on pain, along with some potential similarities in frontal and parietal regions.

While the peak areas involved in each type of pain modulation are distinct, it is conceivable that they reflect similar underlying activity patterns with distinct peaks surviving significance thresholds. In order to test this possibility, we performed several additional analyses.

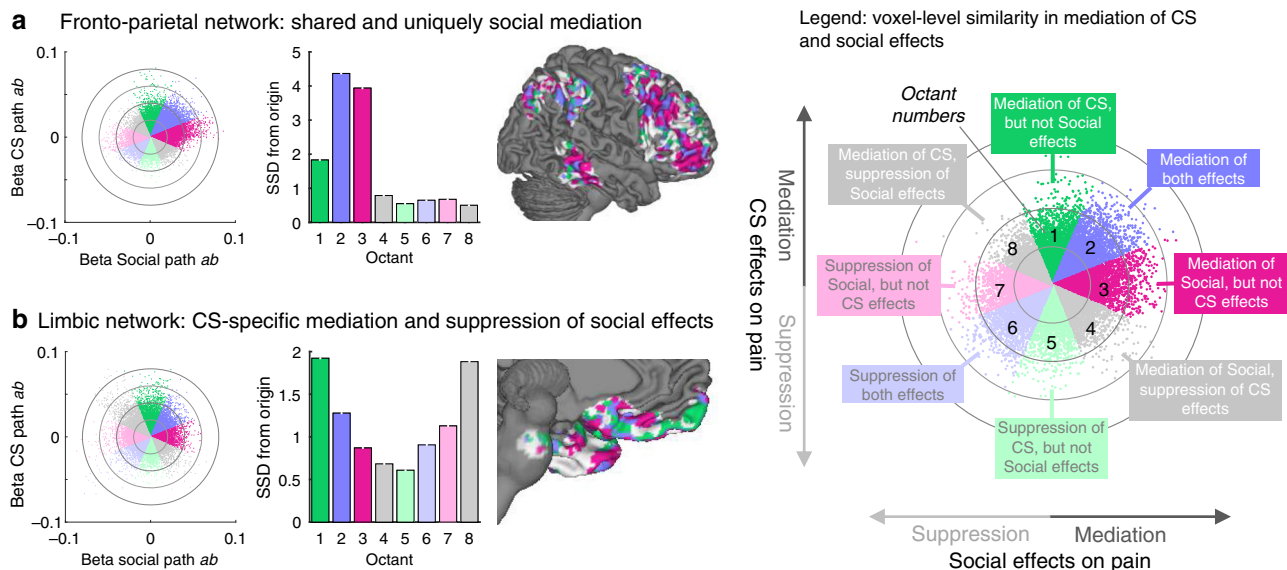
First, we assessed whether social influence and CS-related activity could be separated based on distributed multivariate activity. We trained a support vector machine (SVM) classifier using leave-one-subject-out cross-validation to separate individual beta images for  $\text{Social}_{\text{HIGH}} > \text{Social}_{\text{LOW}}$  from beta images for  $\text{CS}_{\text{HIGH}} > \text{CS}_{\text{LOW}}$  (path *a* effects) in all brain areas associated with mediation (in the *a* and *ab* conjunction images) for either social or learning effects (i.e., union of regions displayed in Fig. 3a, b). This SVM yielded significant and moderately accurate predictions (forced-choice, mean  $\pm$  SE = 72%  $\pm$  7.5%, binomial test,  $P = 0.011$ ), indicating separable, reliably distinct patterns of activity

for social and learning effects during pain in these brain regions (though it does not rule out potential similarities as well).

Second, since conjunction effects were found in frontal and parietal regions, we tested how individual (unthresholded and normalized) beta images for Social and CS path *a* effects engaged more fine-grained parcellations of the frontoparietal network (Fig. 4), based on an established 16-network cortical parcellation<sup>48,51</sup> (Supplementary Fig. 5). Overall, the frontoparietal network (but no other network) was significantly more activated for Social compared to CS path *a* effects ( $t(35) = 3.1$ ,  $p = 0.0042$ , Cohen’s  $d = 0.51$ ). Further, the frontoparietal subnetworks “Control A” and “Control B” in both left and right hemisphere were more activated for the Social compared with the CS path *a* effect (see Fig. 4a). In contrast, both left and right “Control C” subnetworks did not differ between Social and CS effects. This is consistent with the observation that strong Social effects were observed in the prefrontal, lateral parietal, and temporal parts of the frontoparietal network, but not in the medial parietal cortex or posterior cingulate cortex (which constitute the “Control C” subnetwork, see Fig. 4b).

Third, we analyzed the spatial covariation between the unthresholded weight maps for Social and CS mediation effects (path *ab*), and summarized the voxel-level covariation within each network. Here, voxels are the unit of analysis, and voxels within a network may have diverse functional relationships with CS and social information effects. In Fig. 5 and Supplementary Fig. 6, we plotted the joint distribution of both type of effects—i.e., the weight (effect magnitude) of each voxel for the Social mediation effect on the *x*- and for the CS mediation effects on the *y*-axis, separately for each network<sup>52,53</sup>. Effects in any given voxel could be positive, negative, or near-zero for each of the Social and CS mediation effects. This lends itself to classifying voxels within each network into eight equally-sized octants depending on the relative Social and CS effects. Voxels in Octants 1 and 3 were selectively related to positive mediation of CS and social influence respectively. Octants 5 and 7 showed selective negative effects of CS or Social cues, respectively. Voxels in Octants 2 and 6 show positive and negative mediation effects that are common and in the same direction for both cue types. Voxels in these octants drive positive spatial correlations across voxels within the network as a whole, indicating overlap. Finally, voxels in Octants 8 and 4 are those with positive CS effects, but negative Social cue effects, or vice versa. Voxels in these octants drive negative spatial





**Fig. 5** Voxel-level spatial covariation in Social and CS mediation (path  $ab$ ) effects. Scatter plots display unthresholded single-voxel beta weights for CS (y-axis) and Social (x-axis) mediation effects (path  $ab$ ). Each dot represents a voxel. Different colors are assigned to eight octants that reflect positive mediation for CS but not Social effects (Octant 1), positive mediation effects for both (Octant 2), positive mediation of Social, but not CS effects (Octant 3), and so on. Radial grids display distance from the origin (0,0) in 0.02 unit steps. These maps are descriptive, unthresholded illustrations of mediation beta weights. Units are arbitrary. **a** Voxel-level similarity in the frontoparietal network, showing the highest sum of squared distances (SSD, displayed in bar plots) from the origin in Octants 2 and 3, reflecting shared (CS and Social) positive and uniquely positive mediation of Social effects, respectively. Voxels in Octants 1–3 (green, blue, and purple, respectively) within the frontoparietal network (light gray area) are displayed on a lateral brain surface plot (far right). **b** Voxel-level similarity in the limbic network, showing highest SSD in Octants 1 and 8 (see bar plot), reflecting selective mediation of CS effects and suppression of Social effects. Suppression effects are in the opposite direction from, and thus “work against”, the overall effects of cues on pain ratings. Voxels in Octants 1–3 (green, blue, and purple, respectively) in the limbic network (light gray area) are plotted on a medial brain surface (far right). See Supplementary Fig. 6 for the results in all seven large-scale networks. Source data are provided as a Source Data file

correlations, indicating dissimilarity. Furthermore, to provide an overall measure for voxels in each octant, we computed the sum of squared distances (SSD) from the origin, thus combining a measure of both absolute numbers of voxels in each octant and their (squared) distance from the origin.

This analysis of the spatial covariation of Social and CS effects shows qualitatively different patterns across networks. Visual, ventral attention, and default mode networks (Supplementary Fig. 6) have peak SSDs in Octant 2, reflecting a disproportionately large number of voxels that show positive (if not necessarily significant) effects for both Social and CS mediation, indicating some overlap in these networks. However, these networks also contained many voxels with positive weights only for CS (Octant 1) or only for Social mediation effects (Octant 3). Fronto-parietal (Fig. 5) and dorsal attention networks show large effects in the shared positive (Octant 2) and in uniquely Social mediation effects (Octant 3). The limbic network (Fig. 5) shows a peak in voxels mediating CS, but not (or even suppressing) Social effects (Octant 1 and 8). Finally, the Somatomotor network was associated with many voxels showing suppressor effects for CS (negative slopes in mediation) and shared suppressor effects for CS and Social effects, in line with its role in primary nociceptive, but less in contextual pain modulation effects. Overall, this analysis provides more detailed evidence for shared and non-shared elements within each network. The networks with the strongest evidence for some shared processing include the dorsal and ventral attention, frontoparietal, and default mode networks, but these similarities are offset by the differential responses in the vast majority of voxels in these networks (the overall spatial covariance across voxels is relatively weak). Those with the strongest evidence for dissimilar effects of Social and CS

include the Limbic (CS preference) and Somatomotor networks.

## Discussion

The effects of both socially conveyed and learned expectations on human behavior and experience are pervasive. Dissecting the brain mechanisms of different sources of expectations—learned and socially instructed—is important for understanding a number of related and yet distinct phenomena such as placebo effects<sup>5–7</sup>, social influence and instruction effects<sup>22,54–56</sup>, the brain mechanisms underlying predictive coding<sup>2,3,7</sup>, and effects of expectations and mindsets more broadly<sup>4</sup>. Here, we found that both social information and learning independently influenced expectations and pain reports, with especially large and durable (nonextinguishing) effects of social information. Future studies should test whether the presence of salient social information can augment or hinder experience-based learning, by comparing CS effects in the presence versus absence of salient social cues. The large effects of social information have important implications for decision-making in real life, where information from social media or online ratings often competes with experience-based learning of factual contingencies.

On a behavioral level, expectations mediated both types of effects—yet the brain effects and mediators were found in largely different brain regions and mapped onto different large-scale networks, suggesting dissociable effects of social information and learning on pain. These findings have implications for predictive coding theories of pain and of information processing more broadly<sup>2</sup>. Whereas previous studies have established the effects of expectations in many domains, it is less clear where these expectations come from and which brain systems are at the “top”



of “top-down” modulation of information processing<sup>57</sup>. Our results suggest a distributed system of source-dependent (e.g., social or conditioned cues) and source-independent brain areas involved in representing expectations and modulating pain accordingly. Current predictive coding models focus on the functional form of the models (e.g., hierarchical Bayesian information), but not on potential differences across priors based on sources and context. We show that priors in predictive coding frameworks may arise from different neural sources, even if they have common psychological correlates.

Mediators of social information effects on pain centered on frontoparietal regions frequently activated during cognitive control and attention tasks, whereas mediators of conditioned cue effects centered on the hippocampal, inferior occipitotemporal, and cerebellar circuits frequently identified in studies of memory and conditioning. This suggests that dlPFC and parietal activations often associated with placebo effects<sup>5,6,41,58</sup> may reflect direct effects of instructions and suggestions, whereas limbic or orbitofrontal activations may reflect learning effects, i.e., integration of experiences over time. The two systems also interact, and we found indications of common mediation effects, including in vmPFC, dmPFC, frontoparietal, and visual areas, suggesting potential common processes related to expectations or attention allocation to pain. Both types of pain modulation and their brain mediators were correlated with reported pain expectancies, suggesting that the differences across them are more related to the type of information presented and cognitive inference performed, rather than the engagement of conscious vs. unconscious predictions.

In addition, both types of effects differed to a large degree from activity that predicted pain independent of our experimental manipulations (Fig. 2b), which engaged primary nociceptive and sensory pathways. This suggests a direct influence of context and cognitive inference on reported pain experience, largely independent of influences on ascending nociceptive input. Such findings are consistent with a number of prior studies demonstrating similar direct effects of perceived control<sup>46,59</sup>, reappraisal<sup>60</sup>, and placebo<sup>41,44</sup>.

Our results resonate with a large body of literature that has demonstrated the power of social context in shaping behavior, health, and affect<sup>22,61,62</sup>, yet they are among the first to show how abstract social information can dramatically alter pain reports and physiological responses to pain<sup>42,63</sup>—even in the presence of predictive learning cues. Interestingly, while social information effects were seen in some areas associated with mentalizing, such as TPJ, dmPFC and adjacent ACC, other areas often associated with mentalizing (e.g., precuneus) were not key mediators of social information effects on pain. Instead, social information exerted its effects mostly via the recruitment of areas of the frontoparietal and the dorsal attention network such as dlPFC and IPS. These brain regions (especially dmPFC, vlPFC, and dlPFC) have been shown to play a role in social conformity effects on decision-making and value<sup>55,64,65</sup> and in instruction effects on aversive<sup>16,31</sup> and appetitive learning<sup>29,30</sup>. For instance, dmPFC responds to social conflict<sup>55</sup> and social dissonance<sup>64</sup>, which lead to adjustments in decision-making and preferences. DmPFC has been associated with individual differences in confirmation biases that prevent extinction of instructed learning<sup>66</sup> and placebo-induced suppression of prediction errors in the ventral striatum<sup>31</sup>. Slightly more ventral areas in rACC have been proposed to be important for social learning and tracking other people’s motivational states<sup>67</sup>. Further, learning by instructions is thought to be mediated by frontal systems, especially dlPFC<sup>68</sup>. Interactions of dlPFC with value-related and learning-related brain regions such as vmPFC, amygdala, and striatum may underlie instruction effects on associative learning<sup>16,29,31</sup>. dlPFC and IPS

have also been shown to play a role in the modality-independent representation of meaning and expectations<sup>69</sup>. Thus, though social influence did not have significant effects on primary targets of nociceptive afferents or validated brain predictors of nociceptive pain, they do alter circuits important for expectations, meaning, and choice. This study also shows that these circuits are directly linked to trial-by-trial pain experience.

What are the functional mechanisms underlying these effects? Comparisons with existing meta-analytic maps<sup>49</sup> and large-scale functional brain networks<sup>48</sup> suggest a primary mapping of social influence effects on pain to frontoparietal and dorsal attention networks associated with cognitive control, attention, and multiple forms of top-down influence<sup>33,34,70</sup>. These networks may influence representations in sensory integration areas, such as mid- and anterior insula. In other words, these systems may provide top-down attentional filters to align experience with socially instructed expectations. This interpretation is in line with a recent model of social information effects that suggests that dlPFC represents an “instructed state” that biases affective processing across other brain regions<sup>22</sup>. While frontoparietal and dorsal attention networks are not primarily thought of as “social cognition” or “mentalizing” regions, it is noteworthy that their functions (e.g., sustained attention and cognitive control) are thought to develop in a social context<sup>71,72</sup>, and are associated with social effects on behavior such as social facilitation<sup>73</sup>.

Contrary to intuition, associative learning effects were—while significant and meaningfully large ( $d = \sim 0.5$ )—much smaller in size than social information effects on pain. Despite strong effects of the different temperatures associated with high and low cues on pain ratings, not all participants learned the contingency between cues and pain (cf. Figure 1). In line with our previous findings<sup>42</sup>, self-reported expectations mediated CS effects on pain. This reflects an important role of contingency awareness in this type of learned pain modulation, which is based on trace conditioning<sup>74,75</sup> (i.e., involving a delay between CS and reinforcer). In contrast, paradigms that use delay conditioning can elicit subliminal learning effects on pain<sup>76,77</sup>.

Brain mediators of CS effects on pain included bilateral hippocampus, inferior occipital cortex, and cerebellum—regions distinct from those mediating social influence effects and those involved in basic nociceptive processing. Cerebellum and hippocampus are important structures for learning and conditioning. While the cerebellum is critically involved in both delay and trace conditioning<sup>78,79</sup>, the hippocampus is necessary for trace conditioning, which requires contingency awareness<sup>80,81</sup>. Given its role in explicit learning and trace conditioning, the hippocampus is well positioned to associate painful stimuli with learned contingencies. While not a classical part of the pain-modulating network<sup>82</sup>, the hippocampus is involved in pain processing across different studies<sup>83–85</sup>, and our findings confirm previous evidence of hippocampal involvement in conditioning effects on pain<sup>77,86</sup> and in pain modulation more broadly<sup>46</sup>. Structural and functional alterations of the hippocampus are also a hallmark of chronic pain conditions<sup>87–90</sup>, suggesting potential bidirectional effects between learning-circuits and pain.

In some respects, the distinction between social influence and conditioning effects shown here parallels other work on top-down versus bottom-up information effects<sup>33,91</sup>. Yet, there were also commonalities and overlapping brain effects. First, self-reported expectations mediated both effects. Second, activation in both sets of brain mediators was correlated with expectation ratings, demonstrating a presumed link with conscious awareness. Third, while there was no strong overlap in regions activated by each manipulation (path *a* effects), several networks (including default, frontoparietal, ventral attention, and visual) showed a large number of subthreshold voxels that showed positive mediation

effects for both experimental manipulations. Finally, a shared theme in the associated Neurosynth decoding results were terms associated with language and meaning (e.g., “verb”, “phonology”, “gestures”), again in line with the idea that conceptual processing may contribute to both learning and social influence effects on pain. Together, these results suggest that direct learning shapes expectations and pain via an interplay of limbic learning areas with frontal and parietal associations areas, whereas social instruction may engage some these circuits, particularly fronto-parietal control networks, more directly.

Social influence, instructions, and different types of learning constitute a large space of manipulations and effects. Here, we contrasted two specific manipulations that differed in several ways; matching them on all sensory and cognitive aspects is likely impossible and would reduce ecological validity. Whereas learning cues were reinforced using classical trace conditioning, the social information was not reinforced, i.e., not predictive of actual stimulus temperature. The learning cues were two drawings, whereas the social information was presented as ratings on a visual analog scale. While it is unlikely that those aspects caused the observed effects, future studies may use different cue types (e.g., numerical instead of visual representations of others’ ratings) and other modalities of cue presentation<sup>69</sup> to replicate the present findings.

Further, in order to avoid a blurring of cue-evoked activity into the pain-related brain activity, pain stimuli were presented several seconds (jittered) after the cues, similar to previous studies<sup>17,42,75,92</sup>. This delay induces trace conditioning, which involves partially different mechanisms than “delay conditioning”, which has been used in some other pain learning studies<sup>21,77,86,93</sup>. Different delays between cue and pain stimuli may involve different neurophysiological mechanisms. Future studies could also test whether trace and delay conditioning interact in different ways with social information and instructions. For instance, social information effects may be stronger or weaker when participants see the CS cues at the same time. This possibility is unlikely, however, since several studies have shown robust effects of social information on affect ratings and brain responses in the absence of conditioning cues<sup>63,94</sup>.

We chose abstract social information (ratings), that are similar to online ratings or verbal information provided by doctors about other patients’ experiences. The advantage of such abstract information is that its effects should be driven by purely conceptual processes without any intrinsic affective value (i.e., they are not “biologically prepared”<sup>95</sup>). Other types of social learning effects on pain<sup>96–99</sup> often involve more direct social interactions, including facial expressions or body language, which could be investigated in future studies. Such observational learning may be more similar to direct learning<sup>100–102</sup>.

Further, we tested the similarity of our brain results with a 7-network parcellation of resting-state activity<sup>48</sup>, and complemented this approach by investigating the contribution of subnetworks in each hemisphere and with analyses of the distribution of CS and social effects across individual voxels. Whether there is an optimal parcellation and/or degree of granularity for these pain-modulatory effects is an open question.

Finally, we tested effects on ratings of pain intensity, but not pain affect (unpleasantness). These two measures are often highly correlated<sup>58,103</sup>. Yet, they can be differentially modulated by contextual manipulations<sup>104,105</sup> and may correlate with different brain areas<sup>106,107</sup>. Future studies could test how social information and learning influence pain affect—a measure that is often more sensitive to social and contextual manipulations of pain<sup>104</sup>.

## Methods

**Participants.** Thirty-eight healthy volunteers with no history of psychiatric, neurological, or pain disorders participated in the experiment. The data from two participants were excluded from all analyses due to large movements artifacts and delays in the experiment, leaving a final sample of 36 participants (20 female and 16 male, mean age = 27.1 years, range 18–50 years). All participants provided informed consent, and were paid for their time. The study was conducted according to the Declaration of Helsinki and approved by the Institutional Review Board of the Department of Psychology and Neuroscience at the University of Colorado Boulder.

**Materials and procedures.** At the beginning of each of 96 trials, participants saw two types of cues (see Fig. 1): Simple visual displays of putatively social information indicated the pain ratings of ten fictitious previous participants. These vicarious pain ratings were either low (Social<sub>LOW</sub>) or high (Social<sub>HIGH</sub>) on average, but were actually completely nonpredictive of subsequent heat stimulation, i.e., both Social<sub>LOW</sub> and Social<sub>HIGH</sub> cues were followed equally often by low (48 °C), medium (49 °C), or high (50 °C) heat pain stimulation. At the same time, participants were presented with one of two conditioning stimuli (CS, drawings of an animal or vehicle, see Fig. 1), which were partially reinforced with moderately painful heat (CS<sub>LOW</sub>, followed by 48 °C or 49 °C stimulation) and more intensely painful heat (CS<sub>HIGH</sub>, followed by 49 °C or 50 °C heat stimulation). Thus, learning cues (CS) were predictive of subsequent heat pain intensity and participants had to learn which CS predicted lower or higher pain on average.

For the social information, we generated 96 different stimuli (48 Social<sub>LOW</sub>, 48 Social<sub>HIGH</sub>), each depicting ten vertical white lines on a horizontal line with gray background that closely resembled the pain rating visual analog scale<sup>42,63</sup>. To generate a variable and plausible set of vicarious pain ratings, each stimulus was generated by sampling ten random values, restricted to be between 0 and 1, from one of two Gaussian distributions:  $N(0.3, 0.15)$  for Social<sub>LOW</sub> and  $N(0.7, 0.15)$  for Social<sub>HIGH</sub>. For the CS, each participant was assigned one drawing of an animal (either a dog, cow, or horse) and one drawing of a vehicle (either a car, truck, or train) as the CS<sub>LOW</sub> and CS<sub>HIGH</sub>, respectively (fully counterbalanced across participants).

Participants were instructed that we were interested in their subjective experience of pain, and how they would predict pain intensity by seeing pain ratings from previous participants and by learning about additional visual cues (animal or vehicle cartoons). They were told that one of the CS (animal or vehicle) was associated with higher pain on average, and the other cue with lower pain on average, and that their task was to learn these associations. Further, they were instructed that they would see the pain ratings of several other participants, presented as vertical lines on the same scale they would use to make their pain ratings. They then performed six runs (16 trials each) of the pain-learning task, corresponding to six nonoverlapping skin sites on the right calf chosen in randomized order. In order to habituate the participants to the stimulation on a new skin site<sup>108</sup>, each run was preceded by a 49 °C stimulus (1 s plateau) before the scanner started again<sup>42</sup>.

Each trial of the learning task started with the presentation (3 s) of one social information cue and one CS on a gray background (top or bottom position of social and CS cues counterbalanced across trials). After a jittered interval (2–4 s), participants used a visual analog scale (VAS, ranging from 0 to 100, anchored at “no pain at all” to “worst pain”) to rate how much pain they expected (5 s). Then, after a jittered gray screen (2–4 s), they received a short (~2 s, 1 s plateau) heat pain stimulation of low (48 °C, 25% of trials), medium (49 °C, 50% of trials), or high (50 °C, 25% of trials) intensity to their right leg (see Fig. 1), using a 27-mm diameter fMRI compatible CHEPS thermode controlled by a Medoc Pathway system (Medoc, Israel). Following another 4–8 s jittered interval, they were asked to rate how much pain they actually felt (pain rating, 5 s), using the same VAS as before. Between trials, participants saw a small white fixation cross on the same gray background (3–8 s). After the learning task, participants performed four runs of a generalization task, which will be reported elsewhere.

**Psychophysiological measures.** Electrodermal (skin conductance) activity was measured at the middle and ring finger of the (non-dominant) left hand. Pulse rate (not analyzed here) was measured at the left index finger. Both signals were recorded using a BIOPAC MP150 system and Acknowledge software at a sampling rate of 2000 Hz. The data were low-pass filtered offline with a cutoff of 1 Hz.

**fMRI data acquisition and preprocessing.** Functional brain activity was acquired using a Siemens TrioTim 3T scanner, covering the brain in 26 interleaved transversal slices (3.4-mm isotropic voxels), with a T2\* weighted EPI GRAPPA sequence (TR = 1.3, TE = 25 ms, flip angle = 50°, FOV = 220 mm). Prior to preprocessing of functional data, time points that are potential global outliers (spikes) were identified based on meeting any of the following criteria: (a) absolute value of global signal > 10 median absolute deviations (m.a.d.), or (b) mahalanobis distance across slice-specific global means and spatial standard deviation > 10 median absolute deviations. These time points are identified on a per-run basis using recursive exclusion of outliers in a step-down test, so that outliers are removed before recursively identifying additional outliers (three iterations). SPM8 was used

for preprocessing for functional images, using a standard pipeline of motion correction, slice-time correction, spatial normalization to MNI space, and spatial smoothing of images using an 8-mm FWHM Gaussian kernel. For spatial normalization, T1-structural MPRAGE images (1 mm isomorphic voxels) were first coregistered to the mean functional image and then normalized to the SPM template using unified segmentation. Preprocessed functional images were resampled at a voxel size of  $3 \times 3 \times 3$  mm.

**Analysis.** We used multilevel general linear models (GLM) to assess the effects of experimental conditions on single-trial expectation and pain ratings. Temperature effects (48, 49, 50 °C, coded as  $-1$ ,  $0$ , and  $1$ ) on pain ratings were assessed across all trials. Effects of CS and social information (both coded as  $-1$  and  $1$ , for low and high, respectively) on pain were assessed only in 49 °C trials (48 out of 96 trials), given that cues were not independent of stimulus temperature in 48 and 50 °C trials. Individual differences in Social and CS effects on expectations (first-level beta weights) were forwarded to correlations with average brain activity in mediator regions (see conjunction analysis). Multilevel-mediation analysis tested whether expectation ratings mediated Social and CS effects on pain. All statistical tests used a significance threshold of  $p < 0.05$  (two-sided), and tests were planned contrasts with paired  $t$  tests unless otherwise specified. Code for all analyses is available at <https://github.com/canlab>.

Single-trial SCR estimates were obtained using the SCRalyze toolbox<sup>109</sup>, which uses a GLM approach similar to standard event-related fMRI analysis. Each participant's skin conductance time series was low-pass filtered (1 Hz cutoff) and normalized. Regressors for single-trial pain onsets were convolved with a canonical SCR function<sup>110</sup> and fitted to the time series data, yielding single-trial beta estimates for SCRs. Parallel to the behavioral data analysis, multilevel GLMs were employed to test the effects of experimental conditions on SCRs in 49 °C trials.

Each participant's functional imaging data was submitted to a GLM that contained single-trials regressors for each heat pain stimulus (modeled as a stick function). The GLM further included event-related regressors for the four different cue conditions (2 Social  $\times$  2 CS), expectations ratings, and pain ratings (boxcar regressors with the duration of the response time). A total of 24 movement regressors (movement estimates for displacement and rotation in three dimensions, their derivatives, squared movement estimates, and derivatives of squared movement estimates) and spike regressors (coded as  $1$  for the outlier time point and zero for all other time points) were added as regressor of no interest to control for motion artifacts. Single-trial regressors that had variance inflation factors  $> 2.5$  (indicating potential multicollinearity) were excluded from subsequent analyses. SPM8 and Matlab2014b were used for parameter estimation.

Single-trial beta estimates of pain-evoked functional BOLD activity in 49 °C trials (to control for temperature) were submitted to multilevel brain mediation analysis (code available at <https://github.com/canlab>). Brain mediation analysis identifies three statistical paths to characterize the causal effects of experimental variables (here: CS and social information) on outcomes measures (pain): (1) path  $a$  characterizes the effect of the experimental factor ( $CS_{HIGH} > CS_{LOW}$  and  $Social_{HIGH} > Social_{LOW}$ ) on brain activity in each voxel, (2) path  $b$  reflects the relationship between brain activity and the outcome, i.e., pain ratings, and (3) path  $ab$  reflects brain activity formally mediating the causal link between experimental factor and brain outcome, thus significantly reducing the strength of the direct path  $c'$ . Resulting statistical maps were thresholded at  $P < 0.05$  FDR-corrected across the whole brain and mediation paths<sup>17,41</sup> (corresponding to a voxel-level of  $p = 0.0014$  for Social and  $p = 0.0010$  for CS effects). Adjacent activity was thresholded at  $p < 0.01$  and  $p < 0.05$  uncorrected for displaying purposes. For completeness, additional mediation analyses were performed to examine the brain activity during the presentation of the cues (Social and CS) and their effects on expectation ratings. For this purpose, single-trial beta estimates of cue-evoked functional BOLD activity across all trials were used for two further multilevel mediation models (otherwise paralleling the models described above). The results of this analysis are shown in the Supplementary Fig. 5) at uncorrected threshold of  $p < 0.001$  and adjacent activity at  $p < 0.01$  and  $p < 0.05$ .

To illustrate the overlap of the brain effects with large-scale brain networks, we computed the spatial correlation (Pearson's  $r$ ) of the thresholded maps with seven previously published resting-state networks as identified in a large ( $N = 1000$  participants) sample<sup>48</sup>. Further, thresholded maps were compared with the Neurosynth database (525 term-based meta-analytic images, 2013 data release from neurosynth.org, Yarkoni, et al.<sup>49</sup>) to identify the ten unique most positively correlated terms associated with the activation maps. The Matlab text mining toolbox was used to create wordcloud plots of the decoded terms, with color brightness reflecting increasing correlation strength of terms with maps (arbitrarily scaled). To investigate the contribution of different parts of the frontoparietal network, we first normalized ( $z$ -scored) individual (unthresholded) path  $a$  beta images and then computed the average activation in the frontoparietal network and in its three subnetworks on each side (right and left Control A, B, and C) for each subject, using a 17-network parcellation<sup>51</sup>.

To test for overlap between social information and CS effects, we computed the intersection of brain areas that showed effects for both type of cues at  $P < 0.05$  FDR-corrected and adjacent voxels at  $p < 0.05$  uncorrected, for both path  $a$  and path  $ab$  effects (Fig. 2). Further, to test for brain areas that showed both mediation (path  $ab$ ) and path  $a$  effects, we computed the intersection of brain areas that

showed both path  $a$  and path  $ab$  effects at  $P < 0.05$  FDR-corrected and adjacent voxels at  $p < 0.05$  uncorrected, separately for social information (Fig. 3a) and CS effects (Fig. 3b). This test uses the conjunction null<sup>49</sup>, which requires that each effect be significant. We visualized results from the conjunction analysis that were significant at an uncorrected level, but only if those were adjacent to voxels exhibiting effects at the more stringent FDR-corrected level. To test for the role of expectations within and between participants, we used spatially averaged beta estimates within these conjunction regions for each trial and regressed them to single-trial expectation ratings (using multilevel GLM, see slope plots in Fig. 3).

In order to illustrate the spatial similarity of Social and CS mediation effects in each large-scale network, we plotted each voxel's beta weight for the CS mediation effect (on the  $y$ -axis) as a function of the Social mediation effect (on the  $x$ -axis). We then divided the resulting scatter plot in eight octants that coarsely define directions of shared positive (Octant 2), shared negative (Octant 6), and uniquely Social or CS-positive and -negative unthresholded activations (unique Social positive in Octant 3, negative in Octant 7; unique CS positive in Octant 1, negative in Octant 5). Negatively correlated effects of Social and CS mediation would be seen in Octants 4 and 1. For illustration purposes, patterns of unique and shared effects, voxels falling in Octants 1 (uniquely CS mediation), 2 (shared), and 3 (uniquely Social mediation), were then plotted on lateral and medial brain surfaces (unthresholded).

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

## Data availability

The source data underlying Figs. 1c-e, 3a-b, 4a, 5a-b, and Supplementary Figs. 2a-c, 5, and 6b are provided as a Source Data file. Contrast images are available on NeuroVault (<https://identifiers.org/neurovault.collection:5677>). Other data can be obtained from the authors upon reasonable request.

## Code availability

Code is available at <https://github.com/canlab>.

Received: 10 December 2018 Accepted: 6 August 2019

Published online: 10 September 2019

## References

- Knill, D. C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
- Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
- Summerfield, C. & de Lange, F. P. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* **15**, 745–756 (2014).
- Crum, A. & Zuckerman, B. Changing mindsets to enhance treatment effectiveness. *JAMA* **317**, 2063–2064 (2017).
- Wager, T. D. & Atlas, L. Y. The neuroscience of placebo effects: connecting context, learning and health. *Nat. Rev. Neurosci.* **16**, 403–418 (2015).
- Benedetti, F. Placebo effects: from the neurobiological paradigm to translational implications. *Neuron* **84**, 623–637 (2014).
- Büchel, C., Geuter, S., Sprenger, C. & Eippert, F. Placebo analgesia: a predictive coding perspective. *Neuron* **81**, 1223–1239 (2014).
- Kok, P., Jehee, J. F. M. & de Lange, F. P. Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* **75**, 265–270 (2012).
- Li, V., Hecce Castañón, S., Solomon, J. A., Vandormael, H. & Summerfield, C. Robust averaging protects decisions from noise in neural computations. *PLoS Comput Biol.* **13**, e1005723 (2017).
- Carlino, E. et al. Role of explicit verbal information in conditioned analgesia. *Eur. J. Pain.* **19**, 546–553 (2015).
- Frisaldi, E. et al. The placebo effect on bradykinesia in Parkinson's disease with and without prior drug conditioning. *Mov. Disord.* **32**, 1474–1478 (2017).
- Doll, B. B., Jacobs, W. J., Sanfey, A. G. & Frank, M. J. Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res.* **1299**, 74–94 (2009).
- Biele, G., Rieskamp, J., Krugel, L. K. & Heekeren, H. R. The neural basis of following advice. *PLoS Biol.* **9**, <https://doi.org/10.1371/journal.pbio.1001089> (2011).
- Staudinger, M. R. & Büchel, C. How initial confirmatory experience potentiates the detrimental influence of bad advice. *NeuroImage* **76**, 125–133 (2013).
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).



16. Atlas, L. Y., Doll, B. B., Li, J., Daw, N. D. & Phelps, E. A. Instructed knowledge shapes feedback-driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *eLife Sci.* **5**, e15192 (2016).
17. 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).
18. Atlas, L. Y. & Wager, T. D. How expectations shape pain. *Neurosci. Lett.* **520**, 140–148 (2012).
19. Sharvit, G., Vuilleumier, P., Delplanque, S. & Corradi-Dell'Acqua, C. Cross-modal and modality-specific expectancy effects between pain and disgust. *Sci. Rep.* **5**, 17487 (2015).
20. Enck, P., Bingel, U., Schedlowski, M. & Rief, W. The placebo response in medicine: minimize, maximize or personalize? *Nat. Rev. Drug Disco.* **12**, 191–204 (2013).
21. Colloca, L. & Miller, F. G. How placebo responses are formed: a learning perspective. *Philos. Trans. R. Soc. Lond. B* **366**, 1859–1869 (2011).
22. Koban, L., Jepma, M., Geuter, S. & Wager, T. D. What's in a word? How instructions, suggestions, and social information change pain and emotion. *Neurosci. Biobehav. Rev.* **81**, 29–42 (2017).
23. Barrett, L. F. & Simmons, W. K. Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* **16**, 419–429 (2015).
24. Schaefer, S. M., Geuter, S. & Wager, T. D. Mechanisms of placebo analgesia: a dual-process model informed by insights from cross-species comparisons. *Prog. Neurobiol.* **160**, 101–122 (2017).
25. Summerfield, C. et al. Predictive codes for forthcoming perception in the frontal cortex. *Science* **314**, 1311–1314 (2006).
26. Rahnev, D., Lau, H. & de Lange, F. P. Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *J. Neurosci.* **31**, 10741–10748 (2011).
27. Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W. & Forstmann, B. U. Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *J. Neurosci.* **32**, 2335–2343 (2012).
28. Heekeren, H. R., Marrett, S., Bandettini, P. A. & Ungerleider, L. G. A general mechanism for perceptual decision-making in the human brain. *Nature* **431**, 859–862 (2004).
29. Li, J., Delgado, M. R. & Phelps, E. A. How instructed knowledge modulates the neural systems of reward learning. *PNAS* **108**, 55–60 (2011).
30. Doll, B. B., Hutchison, K. E. & Frank, M. J. Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J. Neurosci.* **31**, 6188–6198 (2011).
31. Schenk, L. A., Sprenger, C., Onat, S., Colloca, L. & Büchel, C. Suppression of striatal prediction errors by the prefrontal cortex in placebo hypoalgesia. *J. Neurosci.* **37**, 9715–9723 (2017).
32. Plassmann, H., O'Doherty, J., Shiv, B. & Rangel, A. Marketing actions can modulate neural representations of experienced pleasantness. *PNAS* **105**, 1050–1054 (2008).
33. Gazzaley, A. & Nobre, A. C. Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* **16**, 129–135 (2012).
34. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
35. Forstmann, B. U., Brown, S., Dutilh, G., Neumann, J. & Wagenmakers, E.-J. The neural substrate of prior information in perceptual decision making: a model-based analysis. *Front. Hum. Neurosci.* **4**, 40 (2010).
36. Schacter, D. L., Addis, D. R. & Buckner, R. L. Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* **8**, 657–661 (2007).
37. Andrews-Hanna, J. R., Smallwood, J. & Spreng, R. N. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N.Y. Acad. Sci.* **1316**, 29–52 (2014).
38. Geuter, S., Koban, L. & Wager, T. D. The Cognitive neuroscience of placebo effects: concepts, predictions, and physiology. *Annu. Rev. Neurosci.* **40**, 167–188 (2017).
39. Bingel, U., Lorenz, J., Schoell, E., Weiller, C. & Büchel, C. Mechanisms of placebo analgesia: rACC recruitment of a subcortical antinociceptive network. *Pain* **120**, 8–15 (2006).
40. Tinermann, A., Geuter, S., Sprenger, C., Finsterbusch, J. & Büchel, C. Interactions between brain and spinal cord mediate value effects in nocebo hyperalgesia. *Science* **358**, 105–108 (2017).
41. Koban, L., Kross, E., Woo, C.-W., Ruzic, L. & Wager, T. D. Frontal-brainstem pathways mediating placebo effects on social rejection. *J. Neurosci.* **37**, 3621–3631 (2017).
42. Koban, L. & Wager, T. D. Beyond conformity: social influences on pain reports and physiology. *Emotion* **16**, 24–32 (2016).
43. Wager, T. D. et al. An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* **368**, 1388–1397 (2013).
44. Zunhammer, M., Bingel, U., Wager, T. D. & Consortium, P. Placebo effects on the neurologic pain signature: a meta-analysis of individual participant functional magnetic resonance imaging data. *JAMA Neurol.* **75**, 1321–1330 (2018).
45. Kragel, P. A., Koban, L., Barrett, L. F. & Wager, T. D. Representation, pattern information, and brain signatures: from neurons to neuroimaging. *Neuron* **99**, 257–273 (2018).
46. Woo, C.-W. et al. Quantifying cerebral contributions to pain beyond nociception. *Nat. Commun.* **8**, 14211 (2017).
47. Kenny, D. A., Korchmaros, J. D. & Bolger, N. Lower level mediation in multilevel models. *Psychol. Methods* **8**, 115–128 (2003).
48. Thomas Yeo, B. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
49. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
50. Nichols, T., Brett, M., Andersson, J., Wager, T. & Poline, J.-B. Valid conjunction inference with the minimum statistic. *NeuroImage* **25**, 653–660 (2005).
51. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
52. 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).
53. Carter, R. M., Bowling, D. L., Reeck, C. & Huettel, S. A. A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science* **337**, 109–111 (2012).
54. Zaki, J., Schirmer, J. & Mitchell, J. P. Social influence modulates the neural computation of value. *Psychol. Sci.* **22**, 894–900 (2011).
55. Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A. & Fernández, G. Reinforcement learning signal predicts social conformity. *Neuron* **61**, 140–151 (2009).
56. Izuma, K. & Adolphs, R. Social manipulation of preference in the human brain. *Neuron* **78**, 563–573 (2013).
57. Roepstorff, A. & Frith, C. What's at the top in the top-down control of action? Script-sharing and 'top-top' control of action in cognitive experiments. *Psychol. Res.* **68**, 189–198 (2004).
58. Wager, T. D. et al. Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* **303**, 1162–1167 (2004).
59. Bräscher, A.-K., Becker, S., Hoepli, M.-E. & Schweinhardt, P. Different brain circuitries mediating controllable and uncontrollable pain. *J. Neurosci.* **36**, 5013–5025 (2016).
60. 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).
61. Cialdini, R. B. & Goldstein, N. J. Social influence: compliance and conformity. *Annu. Rev. Psychol.* **55**, 591–621 (2004).
62. Necka, E. A. & Atlas, L. Y. The role of social and interpersonal factors in placebo analgesia. *Int. Rev. Neurobiol.* **138**, 161–179 (2018).
63. Yoshida, W., Seymour, B., Koltzenburg, M. & Dolan, R. J. Uncertainty increases pain: evidence for a novel mechanism of pain modulation involving the periaqueductal gray. *J. Neurosci.* **33**, 5638–5646 (2013).
64. Izuma, K. The neural basis of social influence and attitude change. *Curr. Opin. Neurobiol.* **23**, 456–462 (2013).
65. Campbell-Meiklejohn, D. K., Bach, D. R., Roepstorff, A., Dolan, R. J. & Frith, C. D. How the opinion of others affects our valuation of objects. *Curr. Biol.* **20**, 1165–1170 (2010).
66. Jepma, M., Koban, L., van Doorn, J., Jones, M. & Wager, T. D. Behavioural and neural evidence for self-reinforcing expectancy effects on pain. *Nat. Hum. Behav.* **2**, 838–855 (2018).
67. Apps, M. A. J., Rushworth, M. F. S. & Chang, S. W. C. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron* **90**, 692–707 (2016).
68. Cole, M. W., Laurent, P. & Stocco, A. Rapid instructed task learning: a new window into the human brain's unique capacity for flexible cognitive control. *Cogn. Affect. Behav. Neurosci.* **13**, 1–22 (2013).
69. Lobanov, O. V., Zeidan, F., McHaffie, J. G., Kraft, R. A. & Coghill, R. C. From cue to meaning: brain mechanisms supporting the construction of expectations of pain. *Pain* **155**, 129–136 (2014).
70. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
71. Yu, C. & Smith, L. B. The social origins of sustained attention in one-year-old human infants. *Curr. Biol.* **26**, 1235–1240 (2016).
72. Doebel, S. & Munakata, Y. Group influences on engaging self-control: children delay gratification and value it more when their in-group delays and their out-group doesn't. *Psychol. Sci.* **29**, 738–748 (2018).
73. Monfardini, E. et al. Others' sheer presence boosts brain activity in the attention (but not the motivation) network. *Cereb. Cortex* **26**, 2427–2439 (2015).
74. Montgomery, G. H. & Kirsch, I. Classical conditioning and the placebo effect. *Pain* **72**, 107–113 (1997).



75. Koban, L., Kusko, D. & Wager, T. D. Generalization of learned pain modulation depends on explicit learning. *Acta Psychol.* **184**, 75–84 (2018).
76. Jensen, K. B. et al. Nonconscious activation of placebo and nocebo pain responses. *PNAS* **109**, 15959–15964 (2012).
77. Jensen, K., Kirsch, I., Odmalm, S., Kaptchuk, T. J. & Ingvar, M. Classical conditioning of analgesic and hyperalgesic pain responses without conscious awareness. *PNAS* **112**, 7863–7867 (2015).
78. Thompson, R. F. & Steinmetz, J. E. The role of the cerebellum in classical conditioning of discrete behavioral responses. *Neuroscience* **162**, 732–755 (2009).
79. Kalmbach, B. E., Ohya, T., Kreider, J. C., Riusech, F. & Mauk, M. D. Interactions between prefrontal cortex and cerebellum revealed by trace eyelid conditioning. *Learn Mem.* **16**, 86–95 (2009).
80. Clark, R. E. & Squire, L. R. Classical conditioning and brain systems: the role of awareness. *Science* **280**, 77–81 (1998).
81. Moyer, J. R., Deyo, R. A. & Disterhoft, J. F. Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behav. Neurosci.* **104**, 523–532 (1990).
82. Bushnell, M. C., Čeko, M. & Low, L. A. Cognitive and emotional control of pain and its disruption in chronic pain. *Nat. Rev. Neurosci.* **14**, 502–511 (2013).
83. Bingel, U. et al. Subcortical structures involved in pain processing: evidence from single-trial fMRI. *Pain* **99**, 313–321 (2002).
84. Derbyshire, S. W. G. et al. Pain processing during three levels of noxious stimulation produces differential patterns of central activity. *Pain* **73**, 431–445 (1997).
85. Liu, M.-G. & Chen, J. Roles of the hippocampal formation in pain information processing. *Neurosci. Bull.* **25**, 237–266 (2009).
86. Ploghaus, A. et al. Exacerbation of pain by anxiety is associated with activity in a hippocampal network. *J. Neurosci.* **21**, 9896–9903 (2001).
87. Vachon-Preseu, E. et al. The stress model of chronic pain: evidence from basal cortisol and hippocampal structure and function in humans. *Brain* **136**, 815–827 (2013).
88. Apkarian, A. V. et al. Role of adult hippocampal neurogenesis in persistent pain. *Pain* **157**, 418–428 (2016).
89. Mutso, A. A. et al. Abnormalities in hippocampal functioning with persistent pain. *J. Neurosci.* **32**, 5747–5756 (2012).
90. Berger, S. E. et al. Hippocampal morphology mediates biased memories of chronic pain. *NeuroImage* **166**, 86–98 (2018).
91. Ochsner, K. N. et al. Bottom-up and top-down processes in emotion generation: common and distinct neural mechanisms. *Psychol. Sci.* **20**, 1322–1331 (2009).
92. Schafer, S. M., Colloca, L. & Wager, T. D. Conditioned placebo analgesia persists when subjects know they are receiving a placebo. *J. Pain.* **16**, 412–420 (2015).
93. Colloca, L., Sigauo, M. & Benedetti, F. The role of learning in nocebo and placebo effects. *Pain* **136**, 211–218 (2008).
94. Willroth, E. C., Koban, L. & Hilimire, M. R. Social information influences emotional experience and late positive potential response to affective pictures. *Emotion* **17**, 572–576 (2017).
95. Seligman, M. E. P. Phobias and preparedness. *Behav. Ther.* **2**, 307–320 (1971).
96. Colloca, L. & Benedetti, F. Placebo analgesia induced by social observational learning. *Pain* **144**, 28–34 (2009).
97. Hunter, T., Siess, F. & Colloca, L. Socially induced placebo analgesia: a comparison of a pre-recorded versus live face-to-face observation. *Eur. J. Pain.* **18**, 914–922 (2014).
98. Egorova, N. et al. Not seeing or feeling is still believing: conscious and non-conscious pain modulation after direct and observational learning. *Sci. Rep.* **5**, 16809 (2015).
99. Schenk, L. A., Krimmel, S. R. & Colloca, L. Observe to get pain relief: current evidence and potential mechanisms of socially learned pain modulation. *Pain* **158**, 2077–2081 (2017).
100. Olsson, A. & Phelps, E. A. Social learning of fear. *Nat. Neurosci.* **10**, 1095–1102 (2007).
101. Haaker, J., Yi, J., Petrovic, P. & Olsson, A. Endogenous opioids regulate social threat learning in humans. *Nat. Commun.* **8**, 15495 (2017).
102. Lindström, B., Haaker, J. & Olsson, A. A common neural network differentially mediates direct and social fear learning. *NeuroImage* **167**, 121–129 (2018).
103. Coghill, R. C., Sang, C. N., Maisog, J. M. & Iadarola, M. J. Pain intensity processing within the human brain: a bilateral, distributed mechanism. *J. Neurophysiol.* **82**, 1934–1943 (1999).
104. López-Solà, M., Koban, L. & Wager, T. D. Transforming pain with prosocial meaning: a functional magnetic resonance imaging study. *Psychosom. Med.* **80**, 814–825 (2018).
105. Villemure, C. & Bushnell, M. C. Mood influences supraspinal pain processing separately from attention. *J. Neurosci.* **29**, 705–715 (2009).
106. Rainville, P., Duncan, G. H., Price, D. D., Carrier, B. & Bushnell, M. C. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* **277**, 968–971 (1997).
107. Singer, T. et al. Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157–1162 (2004).
108. Jepma, M., Jones, M. & Wager, T. D. The dynamics of pain: evidence for simultaneous site-specific habituation and site-nonspecific sensitization in thermal pain. *J. Pain.* **15**, 734–746 (2014).
109. Bach, D. R., Flandin, G., Friston, K. J. & Dolan, R. J. Time-series analysis for rapid event-related skin conductance responses. *J. Neurosci. Meth* **184**, 224–234 (2009).
110. Bach, D. R., Flandin, G., Friston, K. J. & Dolan, R. J. Modelling event-related skin conductance responses. *Int J. Psychophysiol.* **75**, 349–356 (2010).

### Acknowledgements

We thank Megan Powell, Marianne Reddan, and Choong-Wan Woo for help with data acquisition. This work was funded by grants from the NIMH (R01 MH076136) and a VENI grant of the Netherlands Organization for Scientific Research (to M.J.). Matlab code for analyses is available at: <https://github.com/canlab>.

### Author contributions

L.K., M.J. and T.D.W. conceptualized the experiment. L.K. acquired and analyzed the data. T.D.W. obtained funding and supervised the project. L.K. wrote the original draft and created the figures. M.L.S., M.J. and T.D.W. contributed to the interpretation of the data and to the editing and revising of the paper.

### Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-11934-y>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer review information:** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019