COGNITIVE NEUROSCIENCE

Does inappropriate behavior hurt or stink? The interplay between neural representations of somatic experiences and moral decisions

G. Sharvit[1,2,3], E. Lin[1,4], P. Vuilleumier[1,2,5]*, C. Corradi-Dell'Acqua[1,5,6]*†

Embodied models suggest that moral judgments are strongly intertwined with first-hand somatic experiences, with some pointing to disgust, and others arguing for a role of pain/harm. Both disgust and pain are unpleasant, arousing experiences, with strong relevance for survival, but with distinctive sensory qualities and neural channels. Hence, it is unclear whether moral cognition interacts with sensory-specific properties of one somatic experience or with supramodal dimensions common to both. Across two experiments, participants evaluated ethical dilemmas and subsequently were exposed to disgusting (olfactory) or painful (thermal) stimulations of matched unpleasantness. We found that moral scenarios enhanced physiological and neural activity to subsequent disgust (but not pain), as further supported by an independently validated whole-brain signature of olfaction. This effect was mediated by activity in the posterior cingulate cortex triggered by dilemma judgments. Our results thus speak in favor of an association between moral cognition and sensory-specific properties of disgust.

INTRODUCTION

Recent years have seen a proliferation of research investigating the cognitive and neural mechanisms underlying moral cognition. Seminal theoretical perspectives posit that moral processing and judgments are not achieved through only formal reasoning but might result from intuitions generated by "gut" feelings (1), possibly linked through one's current emotional state. This idea was supported by evidence associating moral decisions with the experience of physical disgust. For instance, participants evaluate moral transgressions as less acceptable when sitting at a dirty desk, smelling a disgusting odor (2), or tasting a bitter beverage (3). Furthermore, individuals exposed to transgressions become more sensitive to physical disgust (4) and express more pronounced thoughts of cleansing (5, 6). These results have often been interpreted within the framework of embodied accounts and conceptual metaphor theories, according to which moral behavior is represented in terms of "purity" (5, 6), with conducts processed as "clean" or "dirty, rotten, stinky" according to whether they comply with ethical norms.

However, such a link between moral cognition and physical disgust has been challenged under both methodological and theoretical grounds. First, recent meta-analyses and high-powered replications failed to confirm early empirical findings, pointing to effects in individual behavioral responses that are (at best) of small magnitude (7, 8). Second, past studies reporting a positive effect often did not use adequate control conditions to assess whether moral cognition is specifically associated with disgust, or if it generalizes also to other affective experiences. In this perspective, one study showed that

moral decisions can be influenced also by nondisgusting chemosensory stimulations, such as neutral or positive odorants, consistent with an overall role played by olfactory intensity (9). Furthermore, few studies comparing the effects of different emotions on moral cognition found stronger effects of disgust than sadness (2) but not stronger than other arousing states such as anger, fear, or even excitement (10). Yet, the most theoretically relevant control condition is represented by harm/pain (11): Unlike many aversive states that often occur in the absence of norm violations (including chemosensory disgust), pain can also result from transgressions such as murder, assault, or rape and is therefore imbued with a "normative" significance. In this view, scholars recently suggested that people's moral judgments are far better explained in terms of considerations about harm (this conduct hurts someone), rather than by an association with the experience of disgust (this conduct stinks) (12).

Nevertheless, several studies pointed out that moral cognition shares with physical disgust similar facial responses (13) and neural activity in brain structures associated with negative affect, such as the anterior insula and amygdala (14, 15) [but see (16)]. These measures are advantageous in terms of sensitivity, as they capture subtle processes that might not be sufficient to affect overt behavior. On the other hand, they suffer an interpretability limitation, given that the same facial muscles and brain areas are frequently engaged under a wide range of states/conditions, including pain (17, 18). Hence, in the absence of a reliable model of what constitutes a disgust (or pain) signal, it is impossible to verify a full similarity with moral cognition. However, the neural structures implicated in ethical dilemmas extend beyond the insula and amygdala, as meta-analyses systematically implicate a broad network comprising also the medio prefrontal cortex (MPFC), temporoparietal junction (TPJ), precuneus (PC), and posterior cingulate cortex (PCC) (19, 20), possibly reflecting the integration of different processes related to theory of mind, decision-making, personal memory, affect, etc. Therefore, it is plausible to assume that the interplay between somatic states and moral cognition suggested in the literature might not necessarily reflect shared neural representations but rather enhanced connectivity between segregated networks.

[1]Laboratory of Behavioural Neurology and Imaging of Cognition, Department of Neuroscience, University Medical Center, University of Geneva, Geneva, Switzerland. [2]Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland. [3]Haas School of Business, University of California, Berkeley, Berkeley, CA, USA. [4]Department of Management, Technology, and Economics, ETH, 8006, Zurich, Switzerland. [5]Geneva Neuroscience Center, University of Geneva, Geneva, Switzerland. [6]Theory of Pain Laboratory, Department of Psychology, Faculty of Psychology and Educational Sciences (FPSE), University of Geneva, Geneva, Switzerland.
*These authors contributed equally to this work.
†Corresponding author. Email: corrado.corradi@unige.ch

Here, we sought to test whether moral cognition might influence (and be influenced by) the representation of comparably unpleasant feelings of disgust and pain. Across two experiments, participants were exposed to olfactory [of either high or low disgust (HD or LD)] and thermal stimuli [high or low pain (HP or LP)] and asked to evaluate their perceived unpleasantness. In half of the trials, stimuli were delivered following an anticipatory cue (reference trials), whereas in the remaining trials, a short text–based dilemma was presented between the cue and the stimulation, describing either a morally challenging situation (moral trials) or a control scenario with no ethical quandary (nonmoral trials). This experimental design (Fig. 1) allowed us to investigate not only the effect of moral transgressions (relative to nonmoral controls) on the subsequent experience of pain/disgust (dilemma ➜ stimuli) but also the influence of disgust/pain expectancy on the dilemma assessment (cue ➜ dilemma). In either case, the critical question is whether moral processing interacts specifically with the experience of disgust (as predicted with embodied accounts) or whether a comparable (or possibly stronger) interplay is observed with pain.

In line with the literature reviewed above, we measured behavioral responses to stimuli and dilemmas. Furthermore, in experiment 1,

we complemented this analysis by measuring galvanic skin response (GSR), whereas in experiment 2, we recorded brain activity using functional magnetic resonance imaging (fMRI). This allowed us to compare the networks evoked by comparably unpleasant pain and disgust, derive neural responses specific for each, and see how they relate with moral cognition and its associated brain activity.

RESULTS

For the purpose of this study, we provide data from a population of $N = 25$ for experiment 1 (18 females; age average = 25.60, SD = 5.14), and $N = 27$ for experiment 2 (14 females; age average = 24.22, SD = 4.23 years). To duly address the research question, we included only a selection within the original sample recruited, in whom comparable pleasantness between pain and disgust was established (see Materials and Methods). Such selection was carried out based only on reference trials and not trials preceded by dilemmas, which were the real focus of the study.

Unpleasantness ratings

Having established that pain and disgust were comparably unpleasant in the reference trials, we tested how the same stimuli were rated
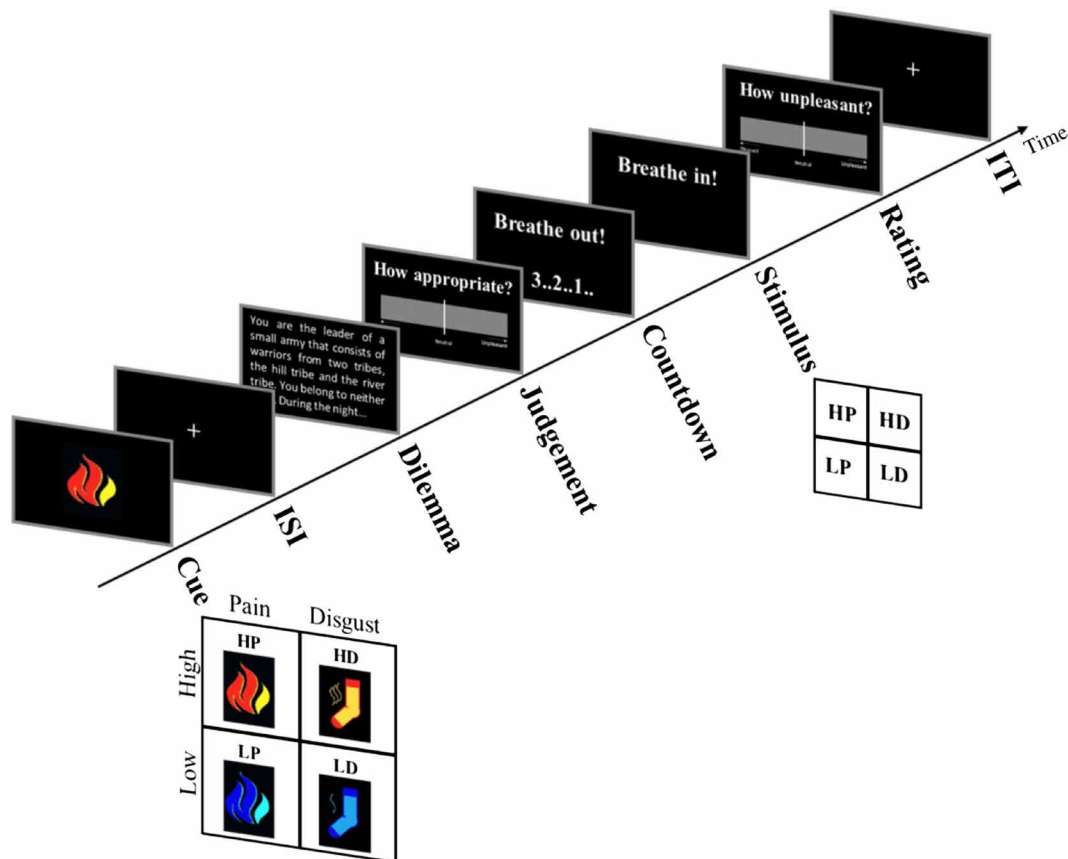
Fig. 1. Experimental design. Each trial began with a pictorial cue predicting either a pain or disgust stimulation. After a variable interstimulus interval (ISI), participants had to read a dilemma and judge a presented course of action on a continuous visual analog scale (VAS) of appropriateness (from extremely inappropriate to extremely appropriate). Next, participants were instructed to slowly breathe out during a countdown period. Then, a breathe in instruction appeared together with the stimulus delivery, which could be either olfactory or thermal. Consequently, participants rated the unpleasantness associated with the stimulus on a VAS (from extremely unpleasant to extremely pleasant), which was followed by an empty intertrial interval (ITI). Four different kinds of predictive cues were presented, indicating the unpleasantness (high/low) and modality (pain/disgust) of the upcoming stimulation (thermal pain/olfactory disgust). Please see Materials and Methods for details about duration and number of trials in each condition of experiments 1 and 2.

after reading a text-based dilemma using a linear mixed model with modality (thermal, olfactory), unpleasantness (neutral, unpleasant), and dilemma (moral, nonmoral) and their interaction as fixed factors, and subjects' identity as a random factor (with random slope and intercept for the fixed factors; see Materials and Methods). This revealed a main effect of unpleasantness (Exp. 1: $t_{24.58}$ = 12.97, P < 0.001; Exp. 2: $t_{20.80}$ = 9.64, P < 0.001). As shown in Fig. 2, in both experiments, HP and HD stimuli were perceived as more unpleasant than their corresponding controls LP and LD. Furthermore, in experiment 1 only, a modality effect was also found (Exp. 1: $t_{25.81}$ = −2.75, P = 0.011; Exp. 2: $t_{21.96}$ = −0.12, P = 0.091). Hence, although in the reference trials the two modalities were carefully matched, at least in experiment 1, the postdilemma thermal events were generally rated as more unpleasant than olfactory ones. Critically, we found no main effect of the content of the preceding dilemma nor any interaction with the other factors (Exp. 1: $|t|$ ≤ 1.18, P ≥ 0.248; Exp. 2: $|t|$ ≤ 1.24, P ≥ 0.223).

Galvanic skin response

We analyzed of GSR in experiment 1 by estimating single trial phasic activity through model-based deconvolution (21) and by feeding them to a generalized linear mixed model, with Tweedie compound Poisson distribution suited for datasets with inflated zero responses (see Materials and Methods) (22). This approach was sufficiently sensitive to capture differences between HP versus LP (t_{398} = 2.12, P = 0.043) and HD versus LD (t_{398} = 3.64, P < 0.001) in the reference trials. Furthermore, when combining both kinds of stimulations in a factorial design with unpleasantness and modality, we found a main effect of unpleasantness (t_{796} = 2.44, P = 0.015) and a main effect of modality (t_{796} = −4.04, P < 0.001), suggesting that, despite matched unpleasantness, GSR was overall larger for thermal (rather than olfactory) stimuli, consistent with previous observations (23, 24). No unpleasantness × modality interaction was observed (t_{796} = −0.18, P = 0.858).

Having established that both pain and olfactory disgust display enhanced GSR in the reference trials (albeit with different magnitude), we tested the extent to which such response was modulated by the content of the preceding dilemma. A generalized linear mixed model with unpleasantness, modality, and dilemma as fixed factors revealed a main effect of unpleasantness (t_{791} = 3.78, P < 0.001) and an unpleasantness × modality × dilemma three-way interaction (t_{791} = −2.62, P = 0.009; all other effects $|t|$ ≤ 1.31, P ≥ 0.189). We explored this interaction by running two separate analyses for thermal and olfactory stimulations, respectively. For thermal stimuli, only a main effect of unpleasantness was observed (t_{396} = 3.81, P < 0.001; all other effects $|t|$ ≤ 1.38, P ≥ 0.167). As shown in Fig. 3A (left subplot), enhanced GSR for HP (relative to LP) was observed following both moral and nonmoral scenarios (t ≥ 2.14, P ≤ 0.032). Furthermore, HP responses did not change between moral and nonmoral dilemmas (t_{198} = 1.26, P ≥ 0.126). In contrast, for olfactory events, we found a significant main effect of unpleasantness (t_{395} = 3.11, P = 0.002), a main effect of dilemma (t_{395} = 2.24, P = 0.025) and an unpleasantness × dilemma interaction (t_{395} = −3.39, P < 0.001). This suggests that, differently from the case of pain, the galvanic response to disgust was influenced by the prior dilemmas, as clearly visible in Fig. 3B (right subplot) which
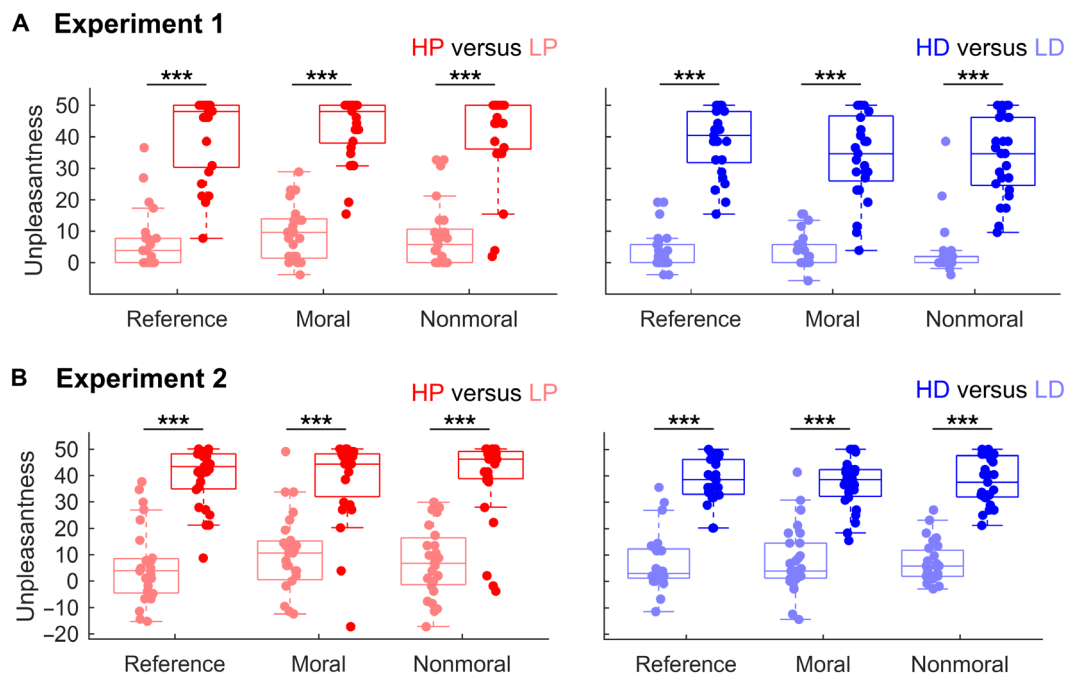
Fig. 2. Thermal and olfactory events: Subjective unpleasantness. Ratings for thermal (left) and olfactory (right) stimuli in reference trials, as well as trials following moral and nonmoral dilemmas, in experiment 1 (**A**) and in experiment 2 (**B**). Data from 25 subjects for experiment 1 and 27 subjects for experiment 2. Red boxplots refer to thermal stimulations, whereas blue boxplots refer to olfactory events. Darker colors refer to unpleasant conditions (HP and HD), whereas lighter tones refer to neutral controls (LP and LD). Each boxplot describes the median value (central mark), the interquartile range (boxes' edges), and the extreme points of the distribution (whiskers) without considering outliers. Single-subject data points are also plotted over each boxplot as colored circles. "***" refers to P < 0.001 in a linear mixed model probing differential responses across conditions.
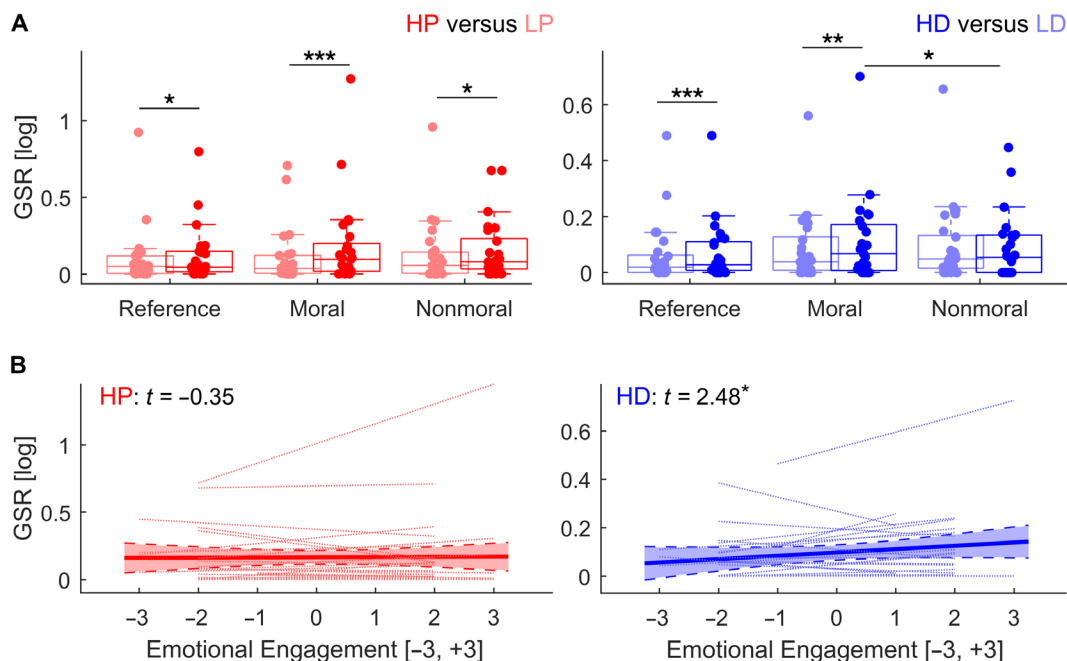
Fig. 3. Thermal and olfactory events: GSR. (**A**) Average GSR for thermal (left) and olfactory (right) stimuli in reference trials, as well as following moral and nonmoral dilemmas. Each boxplot describes the median value (central mark), the interquartile range (boxes' edges), and the extreme points of the distribution (whiskers) without considering outliers. Single-subject data points are also plotted over each boxplot as colored circles. "***," "**," and "*," refer to $P < 0.001$, $P < 0.01$, and $P < 0.05$, respectively, for a generalized linear mixed model test probing differential responses across conditions. (**B**) Multilevel regression describing the relationship between GSR to HP (left) and HD (right) and ratings of Emotional Engagement associated with each dilemma from an independent validation study (see Materials and Methods). Each plot shows a group-wise linear dependency (with a shaded area describing the 95% confidence interval), overlaid by individual regression dotted lines. The t scores of the multilevel regression are also displayed, with "*" referring to significance at $P < 0.05$. Please note that GSR is here displayed in log scale only for readability purposes, as they were analyzed in their raw values.

reports enhanced GSR for HD (relative to LD) following moral ($t_{197} = 3.81$, $P < 0.001$), but not nonmoral stories ($t_{198} = -1.63$). Furthermore, HD responses were larger following moral, as opposed to nonmoral, dilemmas ($t_{197} = -2.39$, $P = 0.017$).
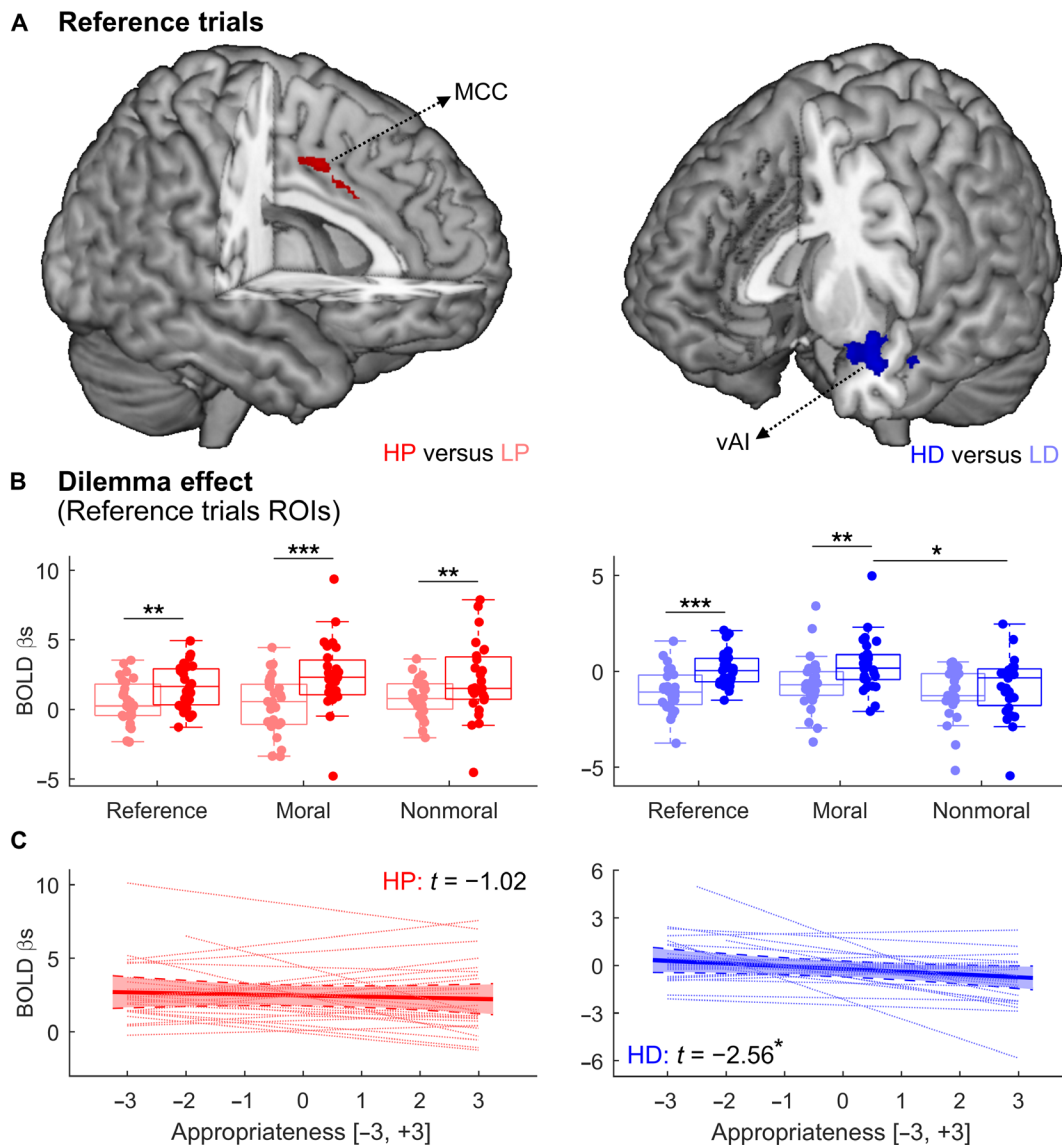
Neural response
Table S1 reports results from the whole-brain fMRI analysis (from experiment 2). In particular, when testing the main effect of pain (HP-LP) across all conditions (pooling both reference, moral, and nonmoral trials), we found increased activity in a well-known network involving the right AI and middle cingulate cortex (MCC). For olfactory events, when testing the main effect of disgust (HD-LD) across all conditions, we found increased activity in the ventral portion of the left anterior insula (vAI). These effects are consistent with our previous neuroimaging study using the same kinds of stimulations (*18*). We then assessed whether the neural response of pain or disgust was modulated by the moral content of the preceding dilemma with direct interaction contrasts ([HP$_{Moral}$ − LP$_{Moral}$] ≠ [HP$_{Nonmoral}$ − LP$_{Nonmoral}$] and [HD$_{Moral}$ > LD$_{Moral}$] ≠ [HD$_{Nonmoral}$ > LD$_{Nonmoral}$]) but found no effect.

Given our predictions that moral-related modulations should affect brain areas selectively engaged by aversive stimuli, we focused our attention on predefined regions of interest (ROIs) sensitive to pain and disgust. Hence, we first identified pain-sensitive regions from the reference trials and confirmed the involvement of the MCC ($x = 2$, $y = 2$, $z = 43$, $t_{130} = 3.83$; 82 consecutive voxels) as in the whole-brain analysis above (albeit at an uncorrected threshold).

Figure 4B (left plot) shows how pain-related activity in this MCC cluster was also confirmed in postdilemma trials, regardless of the moral content of the previous scenario ($t \geq 3.14$, $P \leq 0.004$). Furthermore, HP responses did not change between moral and nonmoral dilemmas ($t_{103.80} = -0.59$, $P = 0.559$). Last, a linear mixed model with unpleasantness and dilemma as fixed factors revealed only a main effect of unpleasantness ($t_{34.25} = 5.62$, $P < 0.001$), whereas no modulation of the previous dilemma was observed, neither as main effect not in interaction with unpleasantness ($|t| \leq 1.18$, $P \geq 0.249$).

For disgust, the analysis of reference trials confirmed the key role played by the left vAI ($x = -38$, $y = 14$, $z = -14$, $t_{130} = 5.29$; 571 consecutive voxels, $P = 0.001$ familywise corrected at the cluster level). Disgust-related activity in this region was also found following moral ($t_{162} = 2.63$, $P = 0.009$) but not nonmoral dilemmas ($t_{52.32} = 0.76$, $P = 0.453$). Furthermore, when focusing on HD responses, these were larger following moral, as opposed to nonmoral, dilemmas ($t_{55.43} = -2.33$, $P = 0.023$). Unfortunately, when analyzing vAI activity in a factorial scheme through a linear mixed model, only a main effect of unpleasantness was found ($t_{101.15} = 2.65$, $P = 0.009$) without any main effect/interaction associated with the factor dilemma ($|t| \leq 1.33$, $P \geq 0.187$). Hence, although exposure to transgressions enhanced insular responses to disgusting odorants, we found no evidence of dissociation between HD and LD.

Overall, the neural results from experiment 2 are in line with those found for the GSR in experiment 1, as they show a differential brain response to pain (in the MCC) regardless of the moral content of the

Fig. 4. Thermal and olfactory events: Neural responses. (A) Whole-brain maps displaying differential response to HP-LP (left) and HD-LD (right) in reference trials. **(B)** Signal associated with thermal (left) and olfactory (right) stimuli in reference trials, as well as trials following moral and nonmoral dilemmas. Individual responses were calculated by averaging the parameter estimates (βs) from the MCC (left) and vAI (right) in the analysis of reference trials. Each boxplot describes the median value (central mark), the interquartile range (boxes' edges), and the extreme points of the distribution (whiskers) without considering outliers. Single-subject data points are also plotted over each boxplot as colored circles. "***," "**," and "*" refer to $P < 0.001$, $P < 0.01$, and $P < 0.05$, respectively, for a generalized linear mixed model test probing differential responses across conditions. **(C)** Multilevel regression describing the relationship between neural activity to HP (left) and HD (right) and ratings of appropriateness from an independent validation study (see Materials and Methods). Each plot shows a group-wise linear dependency (with a shaded area describing the 95% confidence interval), overlaid by individual regression dotted lines. The t scores of the multilevel regression are also displayed, with "*" referring to significance at $P < 0.05$. BOLD, blood-oxygen-level dependent signal.

preceding dilemma. Instead, the response to disgust (in left vAI) was consistently stronger following moral (as opposed to nonmoral) dilemmas.
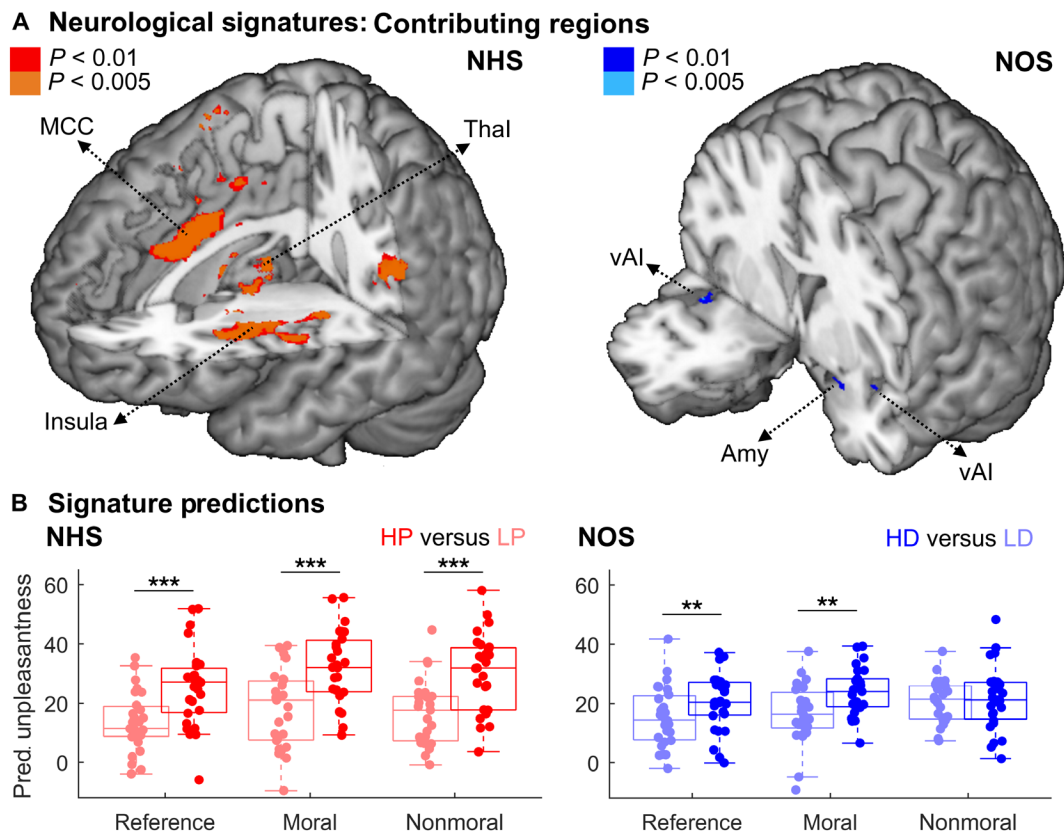
Neurological whole-brain signatures

Rather than focusing on isolated brain regions, we further tested whether moral cognition influenced the overall brain response to thermoception and olfaction across distributed networks. For this purpose, we followed recent advances in neuroscience modeling, which has successfully used multivariate pattern regression to obtain whole-brain neurological signatures to pain (25) and negative affect (26). Using a similar logic, we reanalyzed data from an independent study (18), characterized by identical settings for thermal and olfactory stimulations, except for the absence of text-based dilemmas. A support vector machine algorithm for multivariate regression (under a radial basis function kernel) was then used to model from these data two independent whole-brain signatures predictive of the stimulus unpleasantness in each modality (see Materials and Methods and the Supplementary Materials).

Figure 5A shows the brain regions most relevant for the prediction, which are reminiscent of the contribution maps for previous

A Neurological signatures: Contributing regions

Fig. 5. Neurological whole-brain signatures. (A) Whole-brain maps displaying regions contributing the most to the prediction of the unpleasantness associated with heat (left) and olfaction (right) using an independent dataset from Sharvit *et al*. (*18*). Red and orange blobs refer to regions contributing to the prediction of heat, whereas blue and cyan blobs refer to regions contributing to the prediction of olfaction. Color tones indicate the significance of the contribution as assessed through bootstrap resampling approaches (see Materials and Methods). **(B)** Unpleasantness for thermal (left) and olfactory stimuli (right) from the current study, as predicted by the corresponding signatures. Each boxplot describes the median value (central mark), the interquartile range (boxes' edges), and the extreme points of the distribution (whiskers) without considering outliers. Single-subject data points are also plotted over each boxplot as colored circles. Thal, thalamus; Amy: amygdala. "***" and "**" refer to $P < 0.001$ and $P < 0.01$, respectively, for a generalized linear mixed model test probing differential responses across conditions.

models of thermal pain (*27*), as well as known regions implicated in olfaction (*28*). As a sanity check, we tested whether the estimated models could correctly classify the reference trials from the present dataset. Accordingly, on the basis of our neurological heat signature (NHS), HP was estimated as far more unpleasant than its tailored control LP ($t_{26} = 5.70$, $P < 0.001$), with approximately 12 unpleasantness points of difference along a scale from 0 to 50 (see also Fig. 5B). Likewise, according to our neurological olfactory signature (NOS), HD was more unpleasant than its tailored control LD ($t_{26} = 2.80$, $P = 0.009$), with approximately five unpleasantness points of difference.

Having established that each model could efficiently predict reference trial data from the present experiment, we tested how this sensitivity was influenced by the moral content of the preceding dilemma. Figure 5B shows how the NHS successfully discriminated HP from LP in postdilemma trials, but this was observed regardless of the moral content of the previous scenario ($t \geq 4.69$, $P \leq 0.001$). A linear mixed model revealed only a main effect of unpleasantness ($t_{33.62} = 4.38$, $P < 0.001$), and no significance was associated with the factor dilemma ($|t| \leq 0.99$, $P \geq 0.325$; Fig. 5B, left plot). In contrast, for disgust, the NOS discriminated between HD and LD following the moral ($t_{42.28} = 3.76$, $P = 0.008$) but not the nonmoral dilemmas ($t_{41.47} = -0.16$, $P = 0.866$). Furthermore, a linear mixed model revealed

a main effect of unpleasantness ($t_{29.43} = 2.51$, $P = 0.018$), a main effect of dilemma ($t_{41.15} = 2.26$, $P = 0.029$), and an unpleasantness × dilemma interaction ($t_{66.49} = -2.27$, $P = 0.026$). Hence, the ability of the model to predict disgust unpleasantness was modulated by previous moral processing.

We repeated this analysis by using the seminal pain model developed by Wager *et al.* (*25*). In line with our NHS, this model could also discriminate HP from LP in the reference trials, as well as following moral and nonmoral dilemmas ($|t| \geq 2.26$, $P \leq 0.031$). Furthermore, a linear mixed model run on postdilemma trials revealed only a main effect of unpleasantness ($t_{112} = 2.58$, $P = 0.011$), but there was no effect about the preceding dilemma ($|t| \leq 0.86$, $P \geq 0.394$). Hence, at least for thermal pain, our results displayed in Fig. 5B do not seem idiosyncratic to the model implemented or our own stimulation procedure and dataset. To our knowledge, no neurological model of olfaction, disgust, or chemosensation exists other than the one estimated in the current study.

Evaluative and affective components of the dilemmas
Up to now, we found that galvanic and neural responses of disgust were modulated by prior moral processing, whereas this was not the case of pain. However, it is still possible that pain might be influenced

by specific moral scenarios, such as those with strong emotional load due to vivid descriptions of physical harm. To partially address this issue, we exploited the data from a validation pilot study in which the same dilemmas were evaluated in terms of emotional engagement or the appropriateness of the action described (see Materials and Methods and fig. S1). Hence, we repeated all the analyses described above by replacing the categorical factor dilemma with either of these scores as a continuous predictor. In none of the measures of pain [behavioral rating, galvanic response, MCC activity, and NHS output, including the model from Wager *et al.* (*25*)] did we find an effect of appropriateness or emotional engagement ($|t| \leq 1.52$, $P \geq 0.132$). The only exception was provided by the GSR to thermal stimuli, which were positively associated with appropriateness as a main effect ($t_{396} = 2.31$, $P = 0.023$). However, the direction of such modulation was opposite to the one expected, as both LP and HP showed larger GSR the more the action described was considered appropriate (i.e., enhanced signals were observed for control nonmoral dilemmas).

A different picture is provided by disgust. Consistent with our previous analysis, behavioral responses were not influenced by any predictor from the validation pilot ($|t| \leq 0.41$, $P \geq 0.680$). Instead, in the analysis of galvanic and neural activity, all effects that were associated with the dichotomic factor dilemma were confirmed with the predictors of appropriateness/emotional engagement. More specifically, GSR analysis revealed an unpleasantness × appropriateness ($t_{395} = -3.65$, $P < 0.001$) and an unpleasantness × emotional engagement interaction ($t_{396} = 2.31$, $P = 0.023$): GSR to HD was negatively modulated by appropriateness ($t_{197} = -2.90$, $P = 0.004$; i.e., stronger response to disgust following inappropriate events), whereas it was positively influenced by emotional engagement ($t_{197} = 2.48$, $P = 0.013$; see Fig. 3B, right subplot). Similar interaction effects were observed also for the analysis of NOS output ($|t| \geq 2.18$, $P \leq 0.033$). Last, and consistent with previous analyses, vAI response was not associated with an interaction effect ($|t| \leq 1.33$, $P \geq 0.187$). However, the activity to HD was negatively modulated by appropriateness ratings ($t_{77.90} = -2.56$, $P = 0.012$; see Fig. 4B, left plot) and only marginally positively by emotional engagement ($t_{77.90} = 1.97$, $P = 0.054$).

Bayes factor

Up to now, we showed that participants' sensitivity to pain was not significantly modulated with the moral content of the preceding dilemma but appeared to be influenced only by main effect of unpleasantness. Instead, galvanic (Fig. 3) and neural (Figs. 4 and 5) responses to disgust were influenced by previous moral processing, with positive results regardless of the fact that we modeled dilemma as a dichotomic factor, or in terms of the continuous predictors appropriateness/emotional engagement from the validation pilot. We complemented our previous analyses by using formal model comparisons using the Bayes factor (BF). In particular, for pain responses, we considered behavioral ratings (from both experiments), GSR, and neural activity. For all these measures, we found evidence in favor of a model characterized only by a main effect of unpleasantness, as opposed to other models in which the factor dilemma, or the predictors of appropriateness/emotional engagement, were specified ($BF \geq 3.10$).

A different picture was observed for disgust. In this case, behavioral ratings (from both experiments) led to approximately the same output of pain responses, with a model characterized only by a main effect of unpleasantness exceeding any other model ($BF \geq 3.49$). In contrast, however, for GSR, the best model was characterized by an interaction between unpleasantness × emotional engagement, which moderately exceeded any other model ($BF \geq 2.73$), including one characterized by only the main effect of unpleasantness ($BF = 79.58$). As for the neural response in vAI, the best model was characterized by a main effect of unpleasantness and main effect of appropriateness, which marginally exceeded a model with unpleasantness alone ($BF = 2.60$; all other models $BF \geq 1.90$). Last, for the NOS output, the best model was characterized by both a main effect of unpleasantness and the interaction unpleasantness × appropriateness. As for the case of vAI, also for the NOS output, the best model marginally exceeded the one characterized by only unpleasantness ($BF = 2.61$; all other models $BF \geq 1.04$).

Overall, these data indicate that while pain responses are describable only in terms of the bottom-up properties of stimuli, both the galvanic and neural responses to disgust seem to involve models in which the unpleasant nature of the olfactory stimulations interacts with the moral content of the preceding scenario. Critically, whereas galvanic signals are best explained in terms of the emotional engagement of the dilemma (see Fig. 3B), neural responses to olfactory disgust appear more modulated by the appropriateness of the conduct described.

Dilemma events

For completeness, we also analyzed data from the moral judgments themselves. Appropriateness ratings in both experiments were analyzed using a linear mixed model with dilemma, as well as the modality (thermal, olfactory) and unpleasantness (neutral, unpleasant) of the previous cue. This revealed a dilemma main effect (Exp. 1: $t_{44.85} = 16.44$, $P < 0.001$; Exp. 2: $t_{52.26} = 12.67$, $P < 0.001$), confirming that moral scenarios were indeed rated as less appropriate than nonmoral controls. No other effects were found ($|t| \leq 1.86$, $P \geq 0.063$). See Supplementary Results for follow-up analyses in which we modeled appropriateness/emotional engagement predictors.

Brain activity patterns in periods during which participants read, and subsequently rated, the dilemmas in experiment 2 are shown in table S2 and Fig. 6A. In line with previous meta-analyses on moral processing (*19*, *20*), reading moral (versus nonmoral) dilemmas recruited a network containing the MPFC, TPJ, PC extending to PCC, and superior temporal sulcus, among others. MPFC and PC/PCC were also observed when testing the same contrast for the rating epochs. We then tested whether neural responses to the moral content of dilemmas were modulated by the preceding cue, which predicted the subsequent olfactory or thermal stimulation. However, an interaction contrast testing how the moral > nonmoral effects changed across different cues revealed no suprathreshold activity, neither when correcting for multiple comparisons for the whole brain nor when restricting the search to within the network sensitive to moral violations (as described for the moral > nonmoral main effect in Fig. 6A).

As the disgust-evoked neural activity (e.g., in left vAI, in Fig. 4) was modulated by the content of the preceding dilemmas, we tested whether this relationship was directly mediated by neural responses related to moral processing (Fig. 6A and table S1). We therefore ran voxelwise mediation analysis, using the type of dilemma as the independent variable (0, nonmoral; 1, moral), disgust-evoked activity in vAI as the dependent variable (HD-LD βs), and dilemma-evoked activity preceding the olfactory stimulations as the mediator (see Materials and Methods). When focusing on brain regions engaged
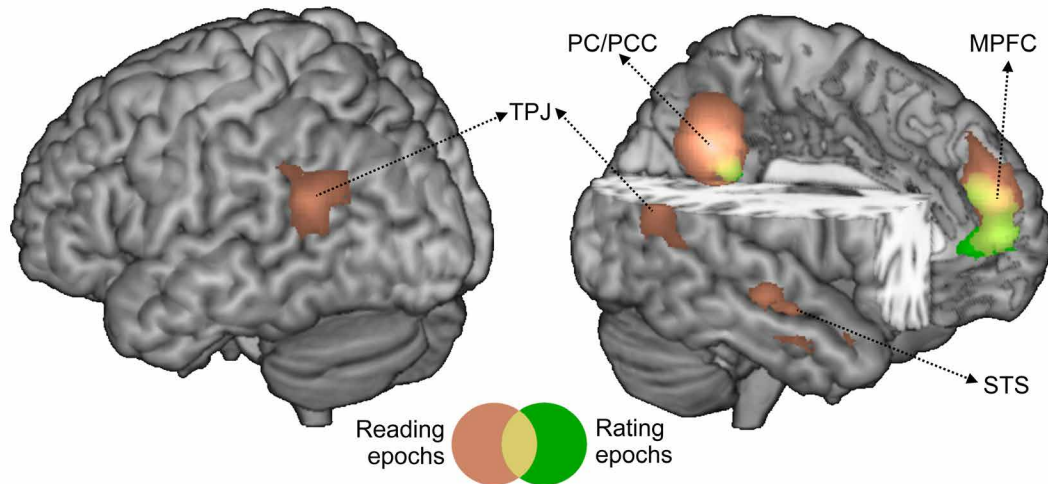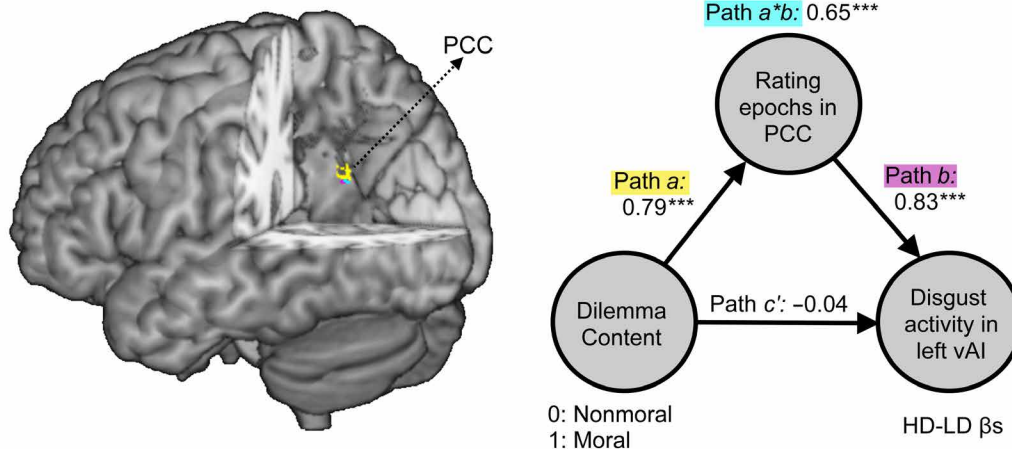
Fig. 6. Dilemma events: Neural responses. (**A**) Whole-brain maps displaying differential response to moral-nonmoral dilemmas in both reading (brown blobs) and rating epochs (green blobs). (**B**) Whole-brain maps displaying regions implicated in the mediation analysis. The parameters extracted from the portion of the PCC showing conjoint activity for paths *a* (yellow blobs), *b* (purple blobs), and *a* × *b* (cyan blobs). "***" refers to parameters being significantly higher than 0 under bootstrap testing. STS, superior temporal sulcus.

by reading the dilemmas (Fig. 6A, brown blobs), no significant mediation effect was found. However, when focusing on the epochs during which individuals rated the appropriateness of the described conducts, we found a significant effect in the PCC (Fig. 6B and table S3). PCC activity was not only modulated by the moral content of the dilemmas (path *a*, as shown in Fig. 6A) it was also positively coupled with the subsequent disgust-related activity (path *b*) and formally mediated the relationship between the former and the latter (path *a* × *b*).

DISCUSSION
In two separate experiments, participants evaluated moral dilemmas and subsequently experienced painful and disgusting stimuli of comparable unpleasantness. We found that galvanic (experiment 1) and neural (experiment 2) responses to disgust were enhanced by prior exposure to moral transgressions, as opposed to acceptable

behavior. In particular, converging evidence was obtained both by looking at specific brain ROIs (left vAI) and by modeling a whole-brain signature of olfactory unpleasantness from independent data (*18*). Furthermore, neural activity in PCC during the evaluation of scenarios (rating epochs) mediated the relationship between the moral content of the dilemma and the subsequent disgust-evoked activity in the vAI. Last, unlike disgust, moral-related information did not appear to influence participants' responses to pain, which were only determined by their bottom-up properties without any modulation by moral transgressions. Overall, our data show that moral cognition interacts in a privileged fashion with a representation of (olfactory) disgust not pain.

The role of pain and disgust in moral cognition
Even in its core component, disgust is a heterogeneous experience characterized by sensory (e.g., olfactory) aspects, as well as by contextual appraisals and memories, allowing for the assessment of

actual or potential intoxication/contamination, and the preparation of coping withdrawal reactions (29). In this perspective, previous studies shed little light about which component of disgust is related to moral cognition, and whether it could extend to other somatic or affective experiences. For this issue, pain represents the most adequate control, as it shares with disgust an intrinsic unpleasantness (here carefully matched), high arousal, and strong relevance for one's well-being. To our knowledge, no research directly compared pain and disgust in relation to different contextual appraisals. However, Meuleman et al. (30) investigated cross-cultural semantic representations of different affective states and found common evaluations of obstructiveness (i.e., a threat for one's goals) between "being hurt" and "disgust," something that may also extend to physical injuries. Our study exploits the similarity between these two aversive experiences and shows that moral cognition interacts with a component of physical disgust that does not generalize with pain, thus ruling out influences related to unpleasantness, arousal, or common contextual evaluations.

Yet, previous research suggested that representations of pain/harm might be strongly tied with moral judgments (11, 12). Our data do not necessarily contradict these earlier studies, as long as one assumes that such influence occurs at a different level than the one observed for chemosensory disgust. In line with many experiments using negative odorants/tastes (2–4, 7, 8), the present study tested the sensitivity to somatic events that were completely unrelated to the dilemmas described. Hence, disgust might interact with moral cognition at a prenormative (implicit) level (31) by associating the to-be-evaluated conduct with a representation of "stain," but independent of an explicit evaluation of context or deliberate associations. In sharp contrast, previous research often investigated the role of harm in direct relationship with sanctioned behavior, for instance, by testing whether individuals based their condemnations on the physical/psychological damage caused to others (11, 12, 31). This, however, was not the case in our study, which used transgressions with extremely harmful (and emotionally engaging) consequences (e.g., the well-known "trollex problem") (32) but tested potential effects on the sensitivity to pain through the delivery of an unrelated noxious temperature. Hence, differently from disgust, pain could influence moral cognition only in a normative fashion, for instance, through an explicit appraisal of the harmful outcomes caused by inappropriate conduct.

Last, although the present study provides convergent evidence that physical disgust plays a role in moral cognition, such a role might be more circumscribed than assumed in some theoretical accounts. First, we found significant effects only when measuring physiological and brain responses, possibly underlining a modulation that is too subtle to affect overt behavior. Thus, our data add to recent failures at replicating early positive evidence at the behavioral level (7, 8). Second, the present and previous experiments might be confounded by the sensory channel used to deliver disgust, as studies using olfaction/gustation show a stronger link with moral cognition than those using vision (7). Furthermore, recent research shows that deontological judgments can be influenced also by neutral odorants (9, 33). In this perspective, it is unclear whether the effect described in the present study might generalize also to nonchemosensory disgust. Third, our research documents only modulations of moral cognition on disgust experiences, although it was designed to also capture effects in the opposite direction (through the use of predictive cues, see Fig. 1). Such discrepancy might reflect the different nature of the disgust experiences implemented here: one involving direct inputs to the olfactory system and the other involving only indirect expectancy effects (although expectancy can modulate both disgust and pain processing) (18, 23). Overall, our data seem in keeping with the idea that moral cognition is heterogeneous, partly influenced by "intuitions" and "gut feelings" related to the experience of disgust (1), but not uniquely explainable only in these terms. Other processes, such as those implicated in assessing the intentionality of the conduct or the sufferance of the people involved, are likely to play an important role.

The moral network and PCC

The neural structures underlying moral cognition have been studied in a wide range of paradigms and meta-analyses, systematically implicating the lateral and medial portions of the parietal, temporal, and prefrontal cortex (see also Fig. 6) (19, 20). Yet, the functional role of these brain structures is still under debate, with different regions possibly underlying different cognitive and affective processes. For instance, it has been argued that part of this network could underlie mechanisms for the assessment of intentions and beliefs (in perpetrators), as evidenced by meta-analytic conjunction in the TPJ and MPFC between tasks probing moral cognition and those testing theory of mind and mentalizing abilities (19). Furthermore, transient disruption of the right TPJ was found to influence people's moral decisions by diminishing the severity with which a conduct was judged inappropriate if ill-intentioned (versus unintentional) (34).

However, there has been little insight on how the network underlying moral cognition relates to disgust processing. A few studies reported that the amygdala, insula, and orbitofrontal cortex were jointly active during exposure to transgressions and physical disgust (14, 15), in line with the idea that ethical violations are grounded in those same neural processes underlying experience of nausea, aversion, etc. (1, 5, 6). However, such overlap was presumably biased by the presence of explicit elicitors in the description of violations (blood, urine, etc.), as no evidence of shared activations was found when using experimental materials carefully controlled for this aspect (16).

In this perspective, our experiment stands off with respect to the extant literature, by investigating the neural interplay between moral cognition and physical disgust in a previously unexplored way. Rather than identifying activation overlaps, our data reveal enhanced coupling between different regions (PCC versus vAI) in distinct epochs (appropriateness judgment versus olfactory stimulation). Hence, rather than contributing to moral cognition through the evaluation of others' states/intentions, the PCC might serve to connect a representation of the to-be-evaluated conduct with that of personally experienced disgust. In the literature, PCC has been frequently reported in self-processing and episodic memory (35, 36), as well as moral cognition (19, 20). Our data bridge the gap between these two independent lines of research, by suggesting how this region could underlie a pathway to moral decisions grounded on self-related somatic experiences, rather than overt evaluation of the events' context.

Neurological olfactory signature

Part of our neuroimaging results were based on the development and validation of a multivariate model that could predict the ratings evoked by olfactory disgust from a predefined network of interest. To our knowledge, this is the first attempt to derive a neurological signature of olfaction (or chemosensation) using a similar approach

as models of pain (*25*) and negative affect (*26*). Defined by an independent sample (*18*), the model proved sufficiently reliable in predicting participants' unpleasantness in the present study. In addition, when using the same processing pipeline to predict the unpleasantness of painful temperatures (rather than odorants), the outcome was fairly similar to a previous model in the literature (*25*), both in terms of regions implicated (Fig. 5) and predictive effectiveness. This provides further support to the reliability of our approach.

This reliability of modeling was possibly due to the fact that training and testing cohorts were collected under almost identical settings, in terms of stimulation used, synchronization of respiratory activity, stimuli duration, rating scale, etc. (see Materials and Methods) (*18*). Such similarity in the methodology served our purpose well, as it allowed us to estimate a multivariate pattern that was independent of our research question, and yet precisely tailored to experimental parameters of the present study. Hence, when testing the degree to which the olfactory network was influenced by prior moral considerations, we were confident that we used the most sensitive model for our dataset. However, a drawback of this approach might concern its generalizability, as we do not know whether the estimated neural signature pertains to specific experimental parameters or whether it also effectively predicts any other kinds of experiences. Future research will need to investigate the degree to which the present signature detects specifically the engagement of the olfactory system or more general features common to other sensory modalities such as gustation or vision. However, these considerations do not undermine the main result of the study that the current olfactory signature was reliably affected by moral judgments, unlike the model trained on heat.

Conclusions

Notwithstanding its limitations, our study extends previous investigations about the relationship between physical disgust and moral cognition in important ways. First, we show that exposure to transgressions enhances the representation of olfactory disgust, as evidenced by measures of galvanic response and neural activity in a predefined network including the ventral insula. Second, we show that such effect is mediated by the activity of the PCC when assessing the acceptability of conducts, thus highlighting a direct functional interaction between regions responsive to moral cognition and those responsive to unpleasant odors. Critically, all the effects observed for disgust were not found for a comparably unpleasant pain stimulation, for which the associated physiological and neural responses were exclusively determined by bottom-up inputs. Overall, our data favor theories suggesting a privileged association between moral cognition and physical disgust and rule against general confounds related to aversiveness.

MATERIALS AND METHODS
Experimental design
Both experiments from the present study were carried out under the same broad experimental design (with minor changes in experiment 2 to comply with requirements for neuroimaging investigations, see below). Participants were exposed to predictive cues anticipating an upcoming olfactory (of either HD or LD) or thermal stimulation (HP or LP). In half of the trials (reference trials), these cues were followed by the anticipated stimulation, and participants were then asked to evaluate its perceived unpleasantness. In the remaining tri-

als, a short scenario was presented between the cue and the stimulation, describing either an ethical dilemma or a control story with no morally challenging elements. This experimental design (Fig. 1) allowed us to (i) analyze the reference trials to ensure that thermal and olfactory events were indeed comparably unpleasant in the absence of any moral/nonmoral scenario, (ii) investigate the effect of moral transgressions (versus nonmoral controls) on the subsequent experience of pain/disgust (dilemma ➔ stimuli), and (iii) assess the disgust/pain expectancy on the dilemma assessment (cue ➔ dilemma).

Participants
We recruited a total of 60 participants. Thirty-three (23 females; aged 18 to 39, mean = 26.00, SD = 4.78 years) took part in experiment 1, whereas 27 (14 females; aged 18 to 33, mean = 24.22, SD = 4.23 years) took part in experiment 2. All participants were native speakers of either French or English and were naïve to the purpose of the experiment. They were right-handed, reported no history of neuropsychological or psychiatric disorders, and were sensitive to the odorants used in the present study. None had any history of neurological/psychiatric illness or reported any olfactory deficit. In addition, all participants of experiment 2 passed a screening for MRI safety. Written informed consent was obtained from all subjects. The study was approved by the local Institutional Review Board and conducted according to the Declaration of Helsinki.

Thermal and olfactory stimulations
We identified for each participant two odorants (expected to elicit HD and LD) and two thermal stimulations (expected to elicit HP and LP). Odorants arose from vials containing isovaleric acid or sclarymol at different concentrations and were delivered to the subjects' nostrils by means of rubber cannulas connected to a computer-controlled, multichannel, custom-built olfactometer. Thermal stimuli were delivered through a computer controlled thermal stimulator with an MRI-compatible 25 × 50 mm fluid-cooled Peltier probe (MSA, Thermotest), attached to participants' legs. Odorants and temperatures were selected on a participant-by-participant basis from preexperimental sessions. In particular, participants first underwent two random presentations of 13 olfactory stimulations, comprehending isovaleric acid and sclarymol (embedded in a solution of odorless dipropylene glycol at five different concentrations: 0.1, 0.5, 1, 5, and 10%), two positive odorants (shampoo and lavender, at 10%) and an odorless control (dipropylene glycol alone). Participants provided unpleasantness ratings on a visual analog scale (VAS) ranging from 50 (extremely unpleasant) to −50 (extremely pleasant). This allowed the selection of two main odorants of interest, one unpleasant (~40 of unpleasantness) and one neutral (~5), plus a positive stimulation of no interest (<0), used only to give participants relief from the aversive stimuli and to reduce the occurrence of putative habituation or sensitization effects. Subsequently, thermal stimulations were selected through a modified double random staircase session aimed at identifying a temperature whose unpleasantness was the closest possible to that of the previously identified negative odorant. This approach led to a highly unpleasant temperature, which averaged across participants at 48.87°C (SD = 1.50) for experiment 1 and 47.15°C (SD = 1.51) for experiment 2. On the basis of this temperature, an additional neutral stimulus was selected, corresponding to an average value of 46.81°C (SD = 1.61) for experiment 1 and 44.68°C (SD = 2.01) for

experiment 2. See previous studies for detailed description of the selection procedure (*18*, *23*).

Text-based dilemmas

For the purpose of this study, we created a database of 32 text-based dilemmas (with moral and nonmoral content). Each text was available both in French and in English. The English version of the dilemmas was obtained from the same database of 44 stimuli used by Greene *et al.* (*32*). These dilemmas were translated ad hoc by a native speaker proficient in English and modified to incorporate cultural differences (e.g., changing "$" to "CHF"). In a pilot study, we asked proficient speakers in either French (20 volunteers: 12 females; aged 19 to 50, mean = 28.8, SD = 6.82 years) or English (37 volunteers: 19 females; aged 20 to 43, mean = 28.81, SD = 4.89) to evaluate each of the 44 dilemmas (in their corresponding language) according to the following dimensions: (i) "How much is the course of action described in the story appropriate for you?" (marked on a VAS ranging from "extremely inappropriate" to "extremely appropriate"); (ii) "How emotionally engaged were you when reading the vignette?" (marked on a VAS ranging from "not engaged at all" to "extremely engaged"); (iii) "How comprehensible was the vignette?" (marked using a scale ranging from "extremely incomprehensible" to "extremely comprehensible"). The data from these rating tasks were used to select 32 dilemmas that displayed, in both their English and French formulation, the following properties: (i) half of the dilemmas (*16*) were expected to describe strong violations of moral norms and thus had to be associated with the lowest appropriateness ratings and with the highest emotional engagement ratings; (ii) the remaining dilemmas were expected to describe ordinary (non-morally challenging) behavior and were associated with the highest appropriateness ratings and the lowest emotional engagement ratings; and (iii) all dilemmas were associated with high comprehensibility ratings. Full information about the selected dilemmas and the pilot data are available in the Supplementary Materials (fig. S1) and under the Open Science Framework at the following link: https://osf.io/jkrvp/.

Task setup

The main experimental task consisted of 76 trials (see Fig. 1). On each trial, a 1.5-s predictive cue was presented and followed by an interstimulus interval (ISI) with a fixation cross at the screen center. In experiment 1, this ISI had a fixed duration of 2 s, whereas in experiment 2, the ISI duration was jittered, ranging from 1.75 to 6.25 s (average 4 s) with an incremental step of 0.25 s, to accommodate the standard requirements of MRI research. Next, the instruction "breathe out" was presented together with a numerical 3-s countdown. When the countdown reached 0, participants had to breathe in evenly while the "breathe in" instruction was presented and the olfactory/thermal stimulus delivered. These breathing instructions ensure that the stimuli were synchronized with the inspiration cycle and stabilized intra- and interparticipant breathing pattern variability (*18*, *23*). Olfactory stimuli lasted 2 s (experiment 1) or 3 s (experiment 2; the duration of the olfactory stimulus was longer in experiment 2 following pilot testing in the MRI setting for the olfactometer). Thermal stimuli always lasted 2 s, although additional 3 s was necessary for the thermal stimulator to reach the target temperature. After stimulation, participants rated its unpleasantness on a VAS with their right hands using the appropriate response keys. The VAS remained onscreen until a response was delivered for a maximum of 6 s. The scale was followed by an intertrial interval (ITI) with a fixation cross at the center of the screen [experiment 1: duration, 4 s; experiment 2: duration, ranging from 1.75 to 6.25 s (average 4 s) with an incremental step of 0.25 s].

In this paradigm, the four stimulus conditions (HP, LP, HD, and LD) were presented following their corresponding cues, which were schematic representations of either a smelly sock or a flame (see Fig. 1). Thus, each cue was always correctly predictive of the upcoming stimulation. Furthermore, in 32 of the 76 trials (50%), a dilemma was presented between the cue and the stimulus. For each participant, the 32 dilemmas (16 moral versus 16 nonmoral, selected from the pilot results) were randomly associated with each of the four cue/stimulus conditions, to minimize putative idiosyncratic confounds of the vignettes. This yielded eight balanced conditions of interest in which moral and nonmoral dilemmas were preceded by each of the four possible cues. For experiment 2 only, a graphical representation of the previously presented cue was also displayed on the top-left corner of the screen during the presentation of the dilemmas (this was performed to enhance any expectancy effect and minimize distraction due to a noisy/stressful environment such as the MRI). The dilemma remained on screen until participants pressed a key, for a maximum duration of 60 s. Subsequently, participants rated how much a course of action associated with the story was appropriate on a VAS ranging from −50 (extremely inappropriate) to +50 (extremely appropriate). The VAS remained on screen until a response was delivered for a maximum of 10 s.

The experimental structure is fully described in Fig. 1. Participants received 32 reference trials in which cues were followed directly by the predicted stimuli (eight trials HP, eight trials LP, eight trials HD, and eight trials LD) and 32 trials in which dilemmas were presented between the cues and the predicted stimuli according to the eight conditions described above. These 32 postdilemma trials were the main objective of the experiment. Last, these 64 trials (32 reference trials + 32 postdilemma trials) were intermingled with 12 trials in which the positive odor was administered following a corresponding cue (schematic flower).

Experiment 1 was organized in one unique block of about 40 min, in which all 76 trials were presented in random order. Experiment 2 was instead split into four independent blocks, each lasting about 12 min and comprising one-fourth (*19*) of the overall trials, to minimize potential movement artifacts and signal drop in MRI data. The experiments were all run using Cogent 2000 (Wellcome Department, London, UK), as implemented in MATLAB R2015b (MathWorks, Natick, MA).

Procedure and apparatus

Participants sat in a chair in front of a computer screen (experiment 1) or laid supine on the MRI patient table with their head fixed by firm foam pads (experiment 2). They were then connected to both the olfactometer and the thermode and underwent the olfactory and thermal stimuli selection sessions (as described above; see the Supplementary Materials for full details), followed by the main experiment. In experiment 1, visual stimuli were projected from a PC (Dell) on a screen (1024 × 768 resolution). Keypresses were recorded on a keyboard (Dell). In experiment 2, the visual stimuli were projected inside the scanner bore with a LCD projector (CP-SX1350, Hitachi, Japan) on a screen (1024 × 768 resolution). Keypresses were recorded on an MRI-compatible bimanual response button box (HH-2 x 4-C, Current Designs Inc., USA).

Statistical analysis

The experiments were carried out under the assumption that HP and HD stimuli were both perceived as more unpleasant than their corresponding neutral counterparts. To ensure this, we focused on the reference trials and excluded all subjects/sessions in which the HP/HD stimuli were rated as neutral (median unpleasantness: $HP \leq -5$ or $HD \leq -5$), or considered as equally (or less) unpleasant than the corresponding controls ($HP \leq LP$ or $HD \leq LD$). In addition, as in the remaining sample, thermal pain was rated as slightly more unpleasant than disgusting odors; we removed subjects/blocks where such divergence was too extreme ($HP-HD \geq 18$). As a result, 8 participants of 33 were excluded from the analysis of experiment 1 (final sample $N = 25$). Likewise, 26 of 108 blocks (27 subjects × 4 blocks per subject) were excluded from experiment 2 (final sample $N = 27$). These exclusion criteria ensured the best trade-off between sample size and matched unpleasantness between modalities. A test comparing directly HP versus HD differences provided strong support in favor of the null hypothesis (Exp. 1: BF = 3.16; Exp. 2: BF = 5.31). Hence, our selection procedure ensured matched unpleasantness between thermal and olfactory stimuli in the reference trials and was applied to all subsequent analytical steps.

Subjective ratings

The analysis of the behavioral responses in the postdilemma trials was carried out as follows. For each subject, for each condition, single-trial ratings of interest (unpleasantness ratings from the stimuli epochs; appropriateness ratings from dilemma epochs) were fed into a linear mixed model with modality (thermal, olfactory), unpleasantness (neutral, unpleasant), and dilemma (moral, nonmoral) as fixed factors and subject identity as a random factor (with random intercept and slope for the fixed factors). Furthermore, we exploited the data from the validation pilot (fig. S1) by replacing the factor dilemma with continuous predictors describing the appropriateness/ emotional engagement associated with each dilemma. This was achieved by using as predictors the median ratings from the validation experiment (subjects who read the French version of the scenarios were modeled as a function of the French validation data, whereas subjects who read the English version of the modeled as function of the English validation data). In case of model misconvergence, the random structure of the model was simplified until convergence was reached. The analysis was carried out with R 4.0.2 software (https://cran.r-project.org/), with the aid of the *lmerTest* package. *P* values associated with the estimated parameters and *t* test were calculated through approximation of the degrees of freedom, as implemented in *lmerTest*. This analysis was complemented with formal model comparison through the estimation of the BF for linear mixed models (with subjects' identity specified as a random factor), as implemented in the *BayesFactor* package for R (https://richarddmorey. github.io/BayesFactor/).

Galvanic skin response

In experiment 1, GSR was recorded through Beckman Ag-AgCl electrodes (8-mm-diameter active area) filled with an isotonic, 0.05 M NaCl, electrode paste, attached to the participant's left hand on the palmar side of the middle phalanges of the second and third fingers. The electrodes were connected to the MP150 Biopac System (Santa Barbara, CA) for GSR recording at a 1000-Hz sampling rate. For each subject, single-trial estimates of GSR were calculated using the MATLAB package Ledalab (www.ledalab.de) (21). More specifically, the raw time course was down sampled to 50 Hz, preprocessed through adaptive Gaussian smoothing, and visually inspected for

potential movement artifacts, which were corrected through spline interpolation. The resulting signal was then deconvolved using continuous decomposition analysis, which separates traces into tonic and physic signal of galvanic activity. For the purpose of the present study, we considered a galvanic phasic response as reliable if exceeding 0.02 μS. Hence, single-trial event-related responses were calculated as the sum amplitude of all suprathreshold phasic responses occurring between 1 and 7 s from the stimulus onset (in olfactory stimulations) or from the time in which temperature reached plateau (in thermal stimulations). These values were analyzed with a mixed model framework similar to that of behavioral ratings. Notably, however, given the high amount of zero responses in GSRs (see fig. S2B), we implemented a generalized linear mixed model with Tweedie compound Poisson distribution (link-log), which allows us to account for an inflated amount of zero values in the dataset (22). The analysis was carried out with the *cplm* package of R. *P* values associated with the estimated parameters and *t* test were calculated through approximation of the degrees of freedom, as implemented in the *parameter* package. This analysis was complemented with formal model comparison through the estimation of the BF. However, as the *BayesFactor* package does not allow modeling generalized linear mixed model with Tweedie distribution, the BF of GSR was estimated through BIC approximation (37).

Imaging data: Acquisition

In experiment 2, functional images were acquired using a 3T whole-body MRI scanner (Trio TIM, Siemens) with a 32-channel head coil. We used an echo planar imaging (EPI) sequence with repetition time (TR) = 2100 ms, echo time (TE) = 30 ms, flip angle = 50°, 36 interleaved slices, 64 × 64 pixels, 3 mm × 3 mm × 3 mm voxel size, and 3.9-mm slice spacing. A field map was also estimated through the acquisition of two functional images with different echo times (short TE = 5.19 ms; long TE = 7.65 ms). Last, structural images were acquired with a T1 weighted three-dimensional sequence (MPRAGE, TR/TI/TE = 1900/900/2.27 ms, flip angle = 9°, parallel accelleration (PAT) factor = 2, 192 sagittal slices, 1 mm × 1 mm × 1 mm voxel sizes, 256 × 256 pixels).

Imaging data: Preprocessing

Preprocessing of functional images was carried out with the software SPM12 (www.fil.ion.ucl.ac.uk/spm/). For each subject, the volumes were realigned, unwrapped using a field map image, coregistered to the structural image, normalized to a template based on 152 brains from the Montreal Neurological Institute with a resolution of 2 mm × 2 mm × 2 mm, and smoothed by convolution with an 8-mm full-width at half-maximum isotropic Gaussian kernel.

Imaging data: First-level analysis

Data were then fed into a first-level analysis using the general linear model framework implemented in SPM12. For each experimental block, we modeled each kind of stimulus event as follows: Olfactory stimuli were modeled as events of 3 s whose onset corresponded to the estimated time in which odorants reached participants' noses; thermal stimuli were modeled as events of 2 s whose onset corresponded to the time of the plateau temperature. As for dilemma epochs, we fitted each dilemma reading period and each dilemma rating period, with a boxcar function with a duration corresponding to the dilemma reading/rating time. This led to 29 regressors on each block [eight dilemma reading epochs, eight dilemma rating epochs, eight stimuli events following the dilemmas, five "reference trials" (LP, HP, LP, HD, and positive odor)], which were convolved with a canonical hemodynamic response function and associated

with regressors describing their first-order time derivative. We also included nine covariates of no interest: These were the six differential realignment parameters, an estimate of inspiration-based changes in the signal (based on a response function from the PhysIO toolbox: www.tnu.ethz.ch/en/software/tapas/documentations/physio-toolbox), and the average time courses extracted from anatomical masks of white matter and cerebrospinal fluid. Low-frequency signal drifts were filtered using a cutoff period of 128 s.

Imaging data: Second-level analysis
The average parameter estimates from the first-level model were fed into separate second-level group analyses testing the effects associated with thermal stimuli, olfactory stimuli, dilemma reading epochs, and dilemma rating epochs. For the analysis of thermal/olfactory epochs, the parameters associated with both reference and postdilemma trials were fed into a second-level flexible factorial analyses with a within-subject factor with six levels (2 unpleasantness × 3 dilemma) and subjects as a random factor. For the analysis of dilemma epochs, the parameters were fed into a second-level flexible factorial analyses with within-subject factor with eight levels (2 modality × 2 unpleasantness × 2 dilemma) and subjects as a random factor. In modeling the variance components, we allowed the factor condition to have unequal variance between its levels, whereas the factor subjects was modeled with equal variance. Activations were considered significant if exceeding an extent threshold allowing $P < 0.05$ correction for multiple comparison for the whole brain, with an underlying height threshold corresponding to $P < 0.001$ uncorrected.

Imaging data: Mediation analysis
We investigated the interplay between neural processes underlying moral and disgust processing, through mediation analysis combined with robust iteratively reweighted least squares (*38*). This was achieved by assessing whether the moral content of the dilemma (coded as a dichotomic variable 0: nonmoral, 1: moral) influenced the subsequent disgust-evoked activity HD > LD in predefined regions (left vAI, see Results), and whether such relationship was mediated by the neural response evoked by the dilemmas preceding the olfactory stimulations. We ran a voxelwise mediation, modeling moral processing activity for each brain coordinate of interest, to estimate three parameters: brain regions that show increased activity for moral dilemmas (path *a*); brain regions that predict changes in subsequent disgust-related response, when controlling for path *a* (path *b*); and brain regions that formally mediate the relationship between dilemmas and disgust signals (path *a* × *b*). Given that we modeled dilemma reading and rating epochs separately, we repeated the mediation analysis twice, once for each epoch. Each analysis thus led to three whole-brain parameter maps (one for each path), whose deviance from zero was assessed with bootstrap techniques (10,000 resamples). As this analysis was conceived to map regions sensitive to moral processing, we constrained our hypothesis by focusing only on the structures that were implicated in the main effect of moral > nonmoral dilemmas in the flexible factorial analysis (see the "Dilemma events" section). Within this mask (5748 voxels for reading, 1262 for rating epochs), we report as significant those effects surviving $q < 0.05$ under false discovery rate correction. The analysis was run with the Mediation Toolbox (https://github.com/canlab/MediationToolbox).

Neurological whole-brain signatures
Consistent with previous studies using the same methods (*25*, *26*), this analysis involved the following steps. (i) Identifying a suitable network of interest for pain and disgust: For this purpose, the auto- mated meta-analysis toolbox Neurosynth (https://neurosynth.org/) (*39*) was used to a create a mask of regions preferentially implicated in studies about pain (18,759 coordinates at a 2 mm × 2 mm × 2 mm resolution) or by chemosensory disgust (combining the terms "disgust," "olfactory," and "taste"; 4443 coordinates). (ii) Identifying an independent dataset in which neural responses to pain and disgust are estimated in the absence of any previous dilemma (training sample): For this purpose, we took advantage of our previous study, in which an independent group of participants was subjected to thermal (painful) or olfactory (disgusting) responses, with three district unpleasantness levels (low, medium, and high) matched between the two modalities (*18*). The trial structure (expectancy cue, sniffing event, stimuli duration, ramp-up of thermal responses) was almost identical to that of the reference trials from the present study. (iii) Extracting pain and disgust data (*18*) from the meta-analytically defined masks. The extracted values were fed into a principal components analysis to identify a limited number of components that retained ~99.9% of the variance of the original data (95 components for pain, 96 components for disgust). (iv) Feeding the components to a machine learning algorithm for the prediction of individual unpleasantness ratings. The algorithm's proficiency was assessed through leave-one-out cross-validation to ensure prediction for an independent group of subjects than the ones used for the modeling. Among the different algorithms implemented (see the Supplementary Materials), the best prediction of either pain or disgust was provided by a support vector machine regression under a radial basis function kernel as implemented in the LIBSVM 3.18 software (www.csie.ntu. edu.tw/~cjlin/libsvm/). (v) Identifying the regions contributing the most to each model, by assessing the impact of the removal of each feature on the model's predictive ability (*40*). This led to a contribution map, whose values were assessed statistically through bootstrap techniques (10,000 resamples).

The estimated pain and disgust signatures (available under Open Science Framework at: https://osf.io/jkrvp/) were applied to fMRI data from the present experiment. As a first step, we assessed whether each model could distinguish between corresponding conditions in the reference trials, with the pain signature successfully distinguishing between HP and LP and the disgust signature distinguishing between HD and LD. Subsequently, each model was used to assess the neural activity in postdilemma trials for each of the corresponding modality, and the resulting prediction values were fed into a linear mixed model probing for differences between unpleasantness and dilemma.

REFERENCES AND NOTES
1. J. Haidt, The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychol. Rev.* **108**, 814–834 (2001).
2. S. Schnall, J. Haidt, G. L. Clore, A. H. Jordan, Disgust as embodied moral judgment. *Pers. Soc. Psychol. Bull.* **34**, 1096–1109 (2008).
3. K. J. Eskine, N. A. Kacinik, J. J. Prinz, A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychol. Sci.* **22**, 295–299 (2011).
4. K. J. Eskine, N. A. Kacinik, G. D. Webster, The bitter truth about morality: Virtue, not vice, makes a bland beverage taste nice. *PLOS ONE* **7**, e41159 (2012).
5. S. Schnall, J. Benton, S. Harvey, With a clean conscience: Cleanliness reduces the severity of moral judgments. *Psychol. Sci.* **19**, 1219–1222 (2008).
6. C.-B. Zhong, K. Liljenquist, Washing away your sins: Threatened morality and physical cleansing. *Science* **313**, 1451–1452 (2006).

7. J. F. Landy, G. P. Goodwin, Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspect. Psychol. Sci.* **10**, 518–536 (2015).

8. E. Ghelfi, C. D. Christopherson, H. L. Urry, R. L. Lenne, N. Legate, M. A. Fischer, F. M. A. Wagemans, B. Wiggins, T. Barrett, M. Bornstein, B. de Haan, J. Guberman, N. Issa, J. Kim, E. Na, J. O'Brien, A. Paulk, T. Peck, M. Sashihara, K. Sheelar, J. Song, H. Steinberg, D. Sullivan, Reexamining the Effect of Gustatory Disgust on Moral Judgment: A Multilab Direct Replication of Eskine, Kacinik, and Prinz (2011). *Adv. Methods Pract. Psychol. Sci.* **3**, 3–23 (2020).

9. C. Cecchetto, R. I. Rumiati, V. Parma, Relative contribution of odour intensity and valence to moral decisions. *Perception* **46**, 447–474 (2017).

10. J. S. Cheng, V. C. Ottati, E. D. Price, The arousal model of moral condemnation. *J. Exp. Soc. Psychol.* **49**, 1012–1018 (2013).

11. C. Schein, K. Gray, The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Pers. Soc. Psychol. Bull.* **41**, 1147–1163 (2015).

12. C. Schein, R. S. Ritter, K. Gray, Harm mediates the disgust-immorality link. *Emotion* **16**, 862–876 (2016).

13. H. A. Chapman, D. A. Kim, J. M. Susskind, A. K. Anderson, In bad taste: Evidence for the oral origins of moral disgust. *Science* **323**, 1222–1226 (2009).

14. J. Moll, R. de Oliveira-Souza, F. T. Moll, F. A. Ignácio, I. E. Bramati, E. M. Caparelli-Dáquer, P. J. Eslinger, The moral affiliations of disgust: A functional MRI study. *Cogn. Behav. Neurol.* **18**, 68–78 (2005).

15. J. Schaich Borg, D. Lieberman, K. A. Kiehl, Infection, incest, and iniquity: Investigating the neural correlates of disgust and morality. *J. Cogn. Neurosci.* **20**, 1529–1546 (2008).

16. M. Oaten, R. J. Stevenson, M. A. Williams, A. N. Rich, M. Butko, T. I. Case, Moral violations and the experience of disgust and anger. *Front. Behav. Neurosci.* **12**, 179 (2018).

17. G. Dirupo, P. Garlasco, C. Chappuis, G. Sharvit, C. Corradi-Dell'Acqua, State-specific and supraordinal components of facial response to pain. *IEEE Trans. Affect. Comput.* **2020**, 1 (2020).

18. G. Sharvit, C. Corradi-Dell'Acqua, P. Vuilleumier, Modality-specific effects of aversive expectancy in the anterior insula and medial prefrontal cortex. *Pain* **159**, 1529–1542 (2018).

19. D. Bzdok, L. Schilbach, K. Vogeley, K. Schneider, A. Laird, R. Langner, S. Eickhoff, Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct.* **217**, 783–796 (2012).

20. R. Eres, W. R. Louis, P. Molenberghs, Common and distinct neural networks involved in fMRI studies investigating morality: An ALE meta-analysis. *Soc. Neurosci.* **13**, 384–398 (2018).

21. M. Benedek, C. Kaernbach, A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* **190**, 80–91 (2010).

22. Y. Zhang, Likelihood-based and Bayesian methods for Tweedie compound Poisson linear mixed models. *Stat. Comput.* **23**, 743–757 (2013).

23. G. Sharvit, P. Vuilleumier, S. Delplanque, C. Corradi-Dell'Acqua, Cross-modal and modality-specific expectancy effects between pain and disgust. *Sci. Rep.* **5**, 17487 (2015).

24. L. Antico, E. Cataldo, C. Corradi-Dell'Acqua, Does my pain affect your disgust? Cross-modal influence of first-hand aversive experiences in the appraisal of others' facial expressions. *Eur. J. Pain* **23**, 1283–1296 (2019).

25. T. D. Wager, L. Y. Atlas, M. A. Lindquist, M. Roy, C.-W. Woo, E. Kross, An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* **368**, 1388–1397 (2013).

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

27. A. Xu, B. Larsen, E. B. Baller, J. C. Scott, V. Sharma, A. Adebimpe, A. I. Basbaum, R. H. Dworkin, R. R. Edwards, C. J. Woolf, S. B. Eickhoff, C. R. Eickhoff, T. D. Satterthwaite, Convergent neural representations of experimentally-induced acute pain in healthy volunteers: A large-scale fMRI meta-analysis. *Neurosci. Biobehav. Rev.* **112**, 300–323 (2020).

28. J. Seubert, J. Freiherr, J. Djordjevic, J. N. Lundström, Statistical localization of human olfactory cortex. *Neuroimage* **66**, 333–342 (2013).

29. P. Rozin, J. Haidt, C. R. McCauley, Disgust, in *Handbook of emotions*, M. Lewis, J. M. Harviland, Eds. (Guilford Press, 1993), pp. 575–594.

30. B. Meuleman, A. Moors, J. Fontaine, O. Renaud, K. Scherer, Interaction and threshold effects of appraisal on componential patterns of emotion: A study using cross-cultural semantic data. *Emotion* **19**, 425–442 (2019).

31. R. Giner-Sorolla, T. Kupfer, J. Sabo, What makes moral disgust special: An integrative functional review, in *Advances in Experimental Social Psychology*, J. M. Olson, Ed. (Academic Press, 2018), vol. 57, pp. 223–289.

32. J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, J. D. Cohen, An fMRI investigation of emotional engagement in moral judgment. *Science* **293**, 2105–2108 (2001).

33. C. Cecchetto, E. Lancini, D. Bueti, R. I. Rumiati, V. Parma, Body odors (even when masked) make you more emotional: Behavioral and neural insights. *Sci. Rep.* **9**, 1–14 (2019).

34. L. Young, J. A. Camprodon, M. Hauser, A. Pascual-Leone, R. Saxe, Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6753–6758 (2010).

35. R. N. Spreng, R. A. Mar, A. S. Kim, The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *J. Cogn. Neurosci.* **21**, 489–510 (2009).

36. A. E. Cavanna, M. R. Trimble, The precuneus: A review of its functional anatomy and behavioural correlates. *Brain* **129**, 564–583 (2006).

37. E.-J. Wagenmakers, A practical solution to the pervasive problems ofp values. *Psychon. Bull. Rev.* **14**, 779–804 (2007).

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

39. T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, T. D. Wager, Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).

40. I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines. *Mach. Learn.* **46**, 389–422 (2002).

Citation: G. Sharvit, E. Lin, P. Vuilleumier, C. Corradi-Dell'Acqua, Does inappropriate behavior hurt or stink? The interplay between neural representations of somatic experiences and moral decisions. *Sci. Adv.* **6**, eaat4390 (2020).