JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles | Behavioral/Cognitive

Parieto-Frontal networks mediate contextual influences in the appraisal of pain and disgust facial expressions

https://doi.org/10.1523/JNEUROSCI.2233-23.2024

Received: 30 November 2023
Revised: 8 October 2024
Accepted: 22 October 2024

Title: “Parieto-Frontal networks mediate contextual influences in the appraisal of pain and disgust facial expressions”

Short title: “Neural networks of contextual face processing”

Giada Dirupo[1,2], Vincent Di Paolo[1], Emilie Lettry[1], Kevin Schwab[1], Corrado Corradi-Dell’Acqua[1,2,3]

[1] Theory of Pain Laboratory, Department of Psychology, Faculty of Psychology and Educational Sciences (FPSE), University of Geneva, Geneva, Switzerland

[2] Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland

[3] Center for Mind/Brain Sciences (CIMeC), University of Trento, Trento, Italy

Corresponding author: Giada Dirupo, giada.dirupo@gmail.com

Number of pages: 34

Number of figures: 5

Number of tables: 6

Number of words in the abstract: 199

Number of words in the introduction: 643

Number of words in the discussion: 1491

Conflict of interest statement: The authors declare no competing financial interests.

Acknowledgments:

The present study was funded by the Swiss National Science Foundation (grant number PP00P1_157424 and PP00P1_183717) awarded to Prof. Corradi- Dell’Acqua. The funding sources had no involvement in any step of the project including the study design, data collection, analyses and interpretation of the data or writing.

Abstract

We appraise other people's emotions by combining multiple sources of information, including somatic facial/body reactions and the surrounding context. A wealthy literature revealed how people take into account contextual information in the interpretation of facial expressions, but the mechanisms mediating such influence still need to be duly investigated. Across two experiments, we mapped the neural representations of distinct (but comparably unpleasant) negative states, pain and disgust, as conveyed by naturalistic facial expressions or contextual sentences. Negative expressions led to shared activity in fusiform gyrus and superior temporal sulcus. Instead, pain contexts recruited supramarginal, postcentral and insular cortex, whereas disgust contexts triggered the temporo-parietal cortex and hippocampus/amygdala. When pairing the two sources of information together, we found higher likelihood of classifying an expression according to the sentence preceding it. Furthermore, networks specifically involved in processing contexts were re-enacted whenever a face followed said context. Finally, the perigenual medial prefrontal cortex showed increased activity for consistent (*vs*. inconsistent) face-contexts pairings, suggesting that it integrates state-specific information from the two sources. Overall, our study reveals the heterogeneous nature of face-context information integration, which operates both according to a state-general and state-specific principle, with the latter mediated by the perigenual medial prefrontal cortex.

Significance Statement

With the aid of controlled database and a comprehensive paradigm, our study provides new insights of the brain and behavioral processes mediating contextual influences on face emotion-specific processing. Our results reveal that context operates both in face-

independent and face-conditional fashion, by biasing the interpretation of any face towards the state implied by associated context, and also triggering processes that monitor the consistency between the different sources of information. Overall, our study unveils key neural processes underlying the coding of state-specific information from both face and context and sheds new light on how they are integrated within the medial prefrontal cortex.

Keywords

"emotional expression", "contextual sentences", "MPFC", "fusiform gyrus", "insular cortex"

Introduction

People appraise others' affect by integrating multiple pieces of information. In particular, facial expressions are not processed exclusively from the inspection of perceivable muscular displacements, but also according to their consistency with the surrounding context (Aviezer et al., 2012; Righart & Gelder, 2008; Stewart et al., 2019; Wieser et al., 2012). For instance, expressions like disgust, fear and joy are classified more rapidly/accurately when preceded by a short text providing congruent information (Carroll & Russell, 1996; Stewart et al., 2019). Likewise, individuals underestimate the intensity of painful expressions if told that the displayed person has been successfully treated (Lamm et al., 2007), that the he/she is simulating the facial reaction (Zhao et al., 2021), or that the pain could not be explained by any medical condition (De Ruddere et al., 2016). Accordingly, faces of pain are likely to be judged as more intense if embedded in a consistent posture or background (Aviezer et al., 2012). Overall, these effects suggest that facial expressions have a degree of ambiguity, especially if evoked by states of comparable unpleasantness, like pain and disgust (Dirupo et al., 2022; Kunz et al., 2013). As such, context represents a critical source of disambiguation (Carroll & Russell, 1996; Stewart et al., 2019). This opens the question on which subprocess is influenced by contextual information, and whether it involves neural mechanisms of facial expressions decoding, or high-order representations of affective states arising from multiple sources of information.

It is known that static faces are processed by ventral portions of the occipital and fusiform cortex (Duchaine & Yovel, 2015; Haxby et al., 2000), while the middle temporal-occipital cortex and the superior temporal sulcus seem to process dynamic information (Deen et al., 2015; Duchaine & Yovel, 2015; Schobert et al., 2018). Critically, the sight of facial expressions of pain and disgust implicates the anterior insular cortex and inferior frontal gyrus (Gan, Zhou, Li, Jiao, Jiang, Biswal, et al., 2022; Jauniaux et al., 2019). These regions might encode both domain-general and domain-specific information, with some components being specific for pain and disgust and others processing supra-ordinal dimensions such as unpleasantness (Corradi-Dell'Acqua et al., 2016).

Previous studies have investigated the degree with which the neural response to faces in these regions is influenced by contextual information. For instance, Vrtička et al (2009, 2011) found that portions of the fusiform gyrus, amygdala, temporo-occipital cortex and inferior frontal gyrus responded more strongly to facial expressions when associated with contextual cues of opposite valence, possibly underlying error-like signals about the inconsistency. Furthermore, the anterior insula exhibits altered connectivity with the supramarginal gyrus and olfactory midbrain, respectively whenever painful and disgusted faces were associated with cues suggestive that expressions were simulated (Zhao et al., 2021, 2022). Of particular insterest however, is the perigenual portion of the medial prefrontal cortex (mPFC), as this region discloses activity patterns responding coherently to specific emotional states across different sources of information (face, postures, comic-like vignettes, etc.; Peelen et al., 2010; Skerry & Saxe, 2014a). Importantly, however, the perigenual MPFC response might be influenced by supraordinal coding of valence, as previous studies employed only positive vs. negative comparisons (Skerry & Saxe, 2014), or found the strongest pattern differentiation between positive and negative states (Peelen et al., 2010). It is therefore unclear whether perigenual MPFC integrates contextual and facial information according to a state-specific or valence coding.

In the present study we used fMRI to investigate the behavioral and neural mechanisms underlying contextual influences on facial expression processing. To this aim, we run two experiments where video-clips of naturalistic facial expressions of pain and disgust (matched for unpleasantness) were associated to contextual sentences either consistent or inconsistent with face information. Hence, we tested contextual effects independently from supraordinal coding of unpleasantness. Based on the literature reviewed above, we hypothesized that the integration of the contextual and facial state-specific information involves the MPFC in its perigenual section.

Methods

Population

Thirty-eight participants (20 males, mean age = 24.13 ± 7.53 SD) were recruited for Experiment 1, whereas twenty-six (10 males, mean age = 23.88 ± 3.91) were recruited for Experiment 2. All were native French speakers, declared no history of psychological/psychiatric illness and were naïve to the purpose of the study. Furthermore, they signed an informed consent prior to the experiment. This research was conducted in accordance with the Declaration of Helsinki and was approved by the local ethical committee.

Stimuli

We used a video-database of naturalistic facial reactions of individuals exposed to comparably unpleasant painful or disgusting stimulations. This is composed of 81 video-clips, organized into 27 triplets in each of which the same person reacts to a thermal painful temperature (*fP*), a disgusting olfactory stimulation (*fD*), or a thermal/olfactory stimulation eliciting a neutral reaction (*fN*). *fP* and *fD* were matched for unpleasantness from the point of view of both the video-recorded person, and an independent sample of observers. Furthermore, they were sufficiently similar to be confused at times with one another, but sufficiently different to be discriminated with ~65% accuracy, thus minimizing potential ceiling/floor effects in the main tasks. We also used a database of 81 phrases describing contextual scenarios of individuals in situations eliciting pain (*cP;* e.g., “*walking on a sharp nail*”), disgust (*cD*; “*walking on cat vomit*”) or a neutral situation (*cN*; “*walking on a soft carpet*”). The sentences from these three categories were comparable in length and lexical frequency. Furthermore, pilot validation ensured that *cP* sentences elicited larger association with pain than the other two categories, whereas *cD* sentences elicited larger association with disgust. Finally, *cP* & *cD* sentences elicited similar unpleasantness ratings, both reliably larger than those associated with neutral context.

Facial Expressions Database

We used a video-database of naturalistic facial expressions of pain, disgust and neutral state. Full details about how these videos were created and validated are available in previous research (Dirupo et al., 2020, 2022). In summary, twenty-nine participants (10 males, average

age = 25.00, SD = 3.46) were video recorded whilst undergoing olfactory and thermal stimuli, respectively triggering disgust and pain at different levels of unpleasantness. These video-recordings were used to create a pool of 123 short clips with no sound, organized in 41 triplets with the same person video-recorded whilst experiencing pain and disgust events as matched as possible for unpleasantness, and a third thermal/olfactory stimulation with unpleasantness rated as close as possible to the ideal point of 0 (corresponding to a neutral state). These videos were validated by an independent sample of 24 participants (7 males, average age = 23.54, SD = 4.12), who underwent a classification task in which they had to guess whether the portrayed person experienced pain, disgust or a neutral state. For pain/disgust choices, participants were also asked to subsequently rate the associated unpleasantness (Dirupo et al., 2020). Based on the performance of this independent sample we selected a portion of 81 videos (organized in 27 triplets) which were matched for unpleasantness from the point of view of both the video-recorded person, and an independent sample of observers. Furthermore, spontaneous expressions of pain and disgust are sufficiently similar to be confused at-times with one another, and yet sufficiently different to be discriminated with ~65% accuracy. This minimizes the emergence of ceiling/floor effects in the main experiment. See Table 1 and Figure 1 for full details on the video-database.

[Figure 1 here]

[Table 1 here]

Contextual Sentences

We created short French-written sentences describing painful, neutral and disgusting contexts in the infinitive form. As for the videos, also texts were organized in triplets, describing individuals embedded in situations eliciting pain (e.g., "*walking on a sharp nail*"), disgust ("*walking on cat vomit*"), or a neutral situation ("*walking on a soft carpet*"). As our aim was to describe plausible contexts of affect, without describing directly the affective states themselves, we ensured that none of the sentences reported explicitly words like "pain", "disgust" (or synonyms). Hence, in our experiment we explored the association between a facial reaction and the context perceived with it, rather than relationship between the lexical and facial representation of the same affective state, as is the case of priming-like experiments (see Weingarten et al., 2016). Furthermore, for disgust sentences, we avoided descriptions related to moral transgressions or "violation of the body envelope" (Haidt et al., 1994), which might recall also violence or physical harm. This dataset was validated on an independent sample of native-French speaker volunteers (*Pilot 1*: N = 24, 11 men, age= 27 ± 9.25), who were asked to evaluate the event described in each sentence in terms of pain, disgust and unpleasantness on a Visual Analogue Scale (VAS). Following the results from this pilot, we selected a subportion of 81 sentences (27 triplets containing one sentence for each state), with the following characteristics. First, painful contexts elicited larger pain ratings than disgust and neutral contexts, whereas disgust contexts elicited larger disgust ratings than pain/neutral contexts. Second, pain and disgust contexts elicited similar unpleasantness ratings, both reliably larger than those associated with neutral contexts. Finally, all three categories were matched in terms of text length (number of characters) and lexical frequency, referring to the database Lexique 3.83 (New et al., 2004) which exploits a corpus of 218 books (135'000 words) published between 1950 and 2000. See Table 2 and Figure 1 for full details.

[Table 2 here]

Experimental Set-up

We employed two experiments using the video-clips and sentences in such way that the state described in the context and the one displayed in the face could be congruent or incongruent. Both experiments were programmed and run with Matlab R2012a (Mathworks, Natick, MA) with the aid of the Cogent 2000 toolbox (Wellcome Dept., London, UK).

Experiment 1

Experiment 1 was a behavioral study organized in two independent experimental sessions (Figure 1C), the order of which was counterbalanced across participants. The first task (Face Classification) was chosen in keeping with prior research investigating contextual influences on facial expressions (Carroll & Russell, 1996; Stewart et al., 2019). Participants first read one contextual sentence, and subsequently were shown a facial expression, which they had to classify by pressing one of three keys corresponding to “Pain”, “Disgust” or “Neutral” (the association key-state was balanced across participants), within a time-window of 5 seconds. Importantly, as the experiment required participants to evaluate only facial expressions, contexts were manipulated as task-irrelevant competing information. In other words, while we ensured attention towards both facial expressions and the contexts (see catch trials described below), participants were explicitly instructed to ignore the sentence during the execution of the main task. Within this paradigm, the 81 contexts (27 per state) and the 81 expressions (27 per state) were matched pseudo-randomly to get 9 independent conditions (each with 9 repetitions), where each facial expression type was associated with each type of context, leading to a 3 *Expressions (fP, fD, fN*) x 3 *Contexts (cP, cD, cN*) factorial design.

The second task (Unpleasantness Rating) was organized in almost identical fashion to the classification paradigm, with the only difference that participants were asked to quantify the degree of unpleasantness experienced by the person depicted in the video. The evaluation happened through a VAS where the two extremities were labelled as “neutral” and “extremely unpleasant” (the side of the anchors was balanced across participants) and a cursor could be moved along the scale by pressing two different keys. This rating task was chosen to account for a potential limitation of the classification task. Indeed, based on the literature participants are expected to classify faces as function of the preceding context (Carroll & Russell, 1996; Stewart et al., 2019). Such effect, however, could be explained either in terms of contextual influences on facial processing (e.g., *I see more pain in the face*), or in terms of pre-activation of a given response selection (e.g., *I am more ready to select "pain"*) regardless of the observed face. The unpleasantness rating task was designed in such way that response selection occurred along a dimension that was orthogonal to (and matched between) pain/disgust categories. This would allow to test whether individuals respond to faces as function of the preceding context without any response pre-selection confound.

As, in both tasks, participants were required to evaluate only facial expressions, we included a control condition to ensure that they were paying attention also to the sentences presented before each video. We randomly embedded in each session nine *catch trials* in which contextual sentences were followed by a question aiming at testing participants' comprehension of the situation described. These nine sentences were chosen from those excluded from the contextual validation pilot, and therefore shared similar properties with the 81 used in the main conditions. The question was: *"How many living beings are there in the situation described by this sentence?"*; the possible answers were *"one living being"* or *"more than one living beings"*. Participants had 5 seconds to press one of two keys corresponding to the two possible answers (the association key-response was balanced across participants). Overall, each session comprehended 90 trials (81 experimental trials + 9 catch trials) and lasted approximately 25 minutes.

Experiment 2

In this experiment we recorded neural activity through functional Magnetic Resonance Imaging (fMRI) while participants underwent a modified version of the "Unpleasantness Rating" session from Experiment 1. In particular, we selected the Unpleasantness Rating task (as opposed to the Classification task), as this would allow for the most unbiased investigation of contextual effects on facial processing (e.g., by testing differences in the neural response to the videos when these were congruent *vs*. incongruent with the previous context) independent of any response pre-selection confound. Furthermore, the paradigm was modified to (1) overcome limitations from the previous experiment (see Behavioral Results), (2) optimize the design sensitivity for the analysis of neural activity, (3) include high-level control condition where faces and contexts were presented in isolation.

Hence, the core of the experiment was simplified to a 2 *Expressions (fP, fD*) x 2 *Contexts (cP, cD*) design, with four independent conditions, where pain and disgust expressions were displayed following pain or disgust contexts, in either a consistent or inconsistent fashion. Additionally, we included six high-level control conditions: in three of those participants saw *fP*, *fD* and *fN* in absence of any previous context, whereas in the remaining three participants read *cP*, *cD* and *cN* phrases followed immediately by a rating scale. This led to an overall of 10 independent conditions, with 9 repetitions each. Trials in each of these conditions were followed by a jittered interstimulus interval ranging between 2 and 5 seconds. The same jittered interval was presented in-between contexts and faces in the main trials where the two sources of information were integrated. Please note that, as this modified paradigm contained context-only trials, participants knew that they had to pay attention also to contexts throughout the experiment. It was therefore not necessary to include any control catch trial as in Experiment 1.

Procedure and apparatus

After having read and signed the consent form and MRI security checklist (for Experiment 2), participants underwent the experiment as described above. In Experiment 1, they sat comfortably on an office chair and watch the stimuli displayed on a Dell PC screen. Keypresses were recorded on Dell keyboard where the relevant responses keys were highlighted. In Experiment 2, participants lay supine in the scanner with their head fixed by firm foam pads and underwent a unique scanning session of about 25 minutes. The visual stimuli were presented on a 23" MRI compatible LCD screen (BOLDScreen23; Cambridge Research Systems, UK). Keypresses were recorded on an MRI-compatible bimanual response button box (HH-2X4-C; Current Designs Inc, Philadelphia, PA). Following the experimental session, participants filled demographic questionnaires and were formally debriefed. Experiment 1 took place at the Brain and Behavior Laboratory of the University of Geneva and required approximately 60 minutes. Experiment 2 took place at the Human Neuroscience Platform of the Campus Biotech in Geneva and required approximately 90 minutes.

Data Analysis

Behavioral Data

In the analysis of rating stimuli from both experiments, the cursor position on the scale was converted into a scalar ranging from 1 (the position associated with the label "neutral") to 10 (the opposite position, associated with "extremely unpleasant"). Behavioral data were analyzed through a (Generalized) Linear Mixed Model with *Expressions* and *Context* as fixed factor. As random factors we modeled the identity of the participants, the identity of the people displayed in the video-clips and the contextual sentences. In particular, we privileged those converging model with the most complex random structure (see Tables 1-3), in order to account for possible idiosyncratic effects of the experimental materials. For the analysis of correct Response Times (from Face Classification session, in Experiment 1) and Ratings (from Experiments 1 & 2) we used a Linear Mixed Model and significance of the fixed effects was calculated using the Satterthwaite approximation of the degrees of freedom. For the analysis of classification accuracy (from Experiment 1), we employed a Generalized Linear Mixed Model with a binomial distribution and Laplace approximation. The analysis was carried out as implemented in the *lmerTest* package (Kuznetsova et al., 2017) from R.3.4.4 software (https://cran.r-project.org/).

Neural Activity

In Experiment 2, brain structural and functional images were acquired by the means of a 3T Siemens Magnetom Prisma whole-body MRI scanner with a 64-channel head-and-neck coil. The sequence was multiband with time to recovery = 1100 ms, (TE) = 32 ms, flip angle = 50˚, 66 interleaved slices, 112 x 112 in-plane resolution, 2 x 2 x 2 mm voxel size, and no inter-slice gap. We used no parallel acquisition technique and multiband acceleration factor 6. We estimated a field map based on the acquisition of 2 functional images with different echo times (short TE = 4.92 ms; long TE = 7.38 ms). A structural image of each participant was also recorded with a T1-weighted MPRAGE sequence (192 slices, TR = 2300 ms, TE = 2.32 ms, flip angle = 8°, slice thickness of 0.9 mm, in-plane resolution = 256 × 256, 0.9 × 0.9 × 0.9 mm voxel size).

Statistical analysis was performed using the SPM12 software (http://www.fil.ion.ucl.ac.uk/spm/). For each subject, functional images were realigned, unwrapped and slice-time corrected. The Artifact Detection Tools (embedded in the CONN21 toolbox, Whitfield-Gabrieli & Nieto-Castanon, 2012) were then used for the identification of outlier scans in terms of excessive subject motion and signal intensity spikes. Finally, the images were normalized to a template based on 152 brains from the Montreal Neurological Institute with a voxel-size resolution of 2 × 2 × 2 mm and smoothed by convolution with an 8 mm full width at half-maximum Gaussian.

Preprocessed volumes were fed into a first-level analysis using the general linear model framework implemented in SPM. In particular, our design had 7 kinds of face conditions: 3 conditions in which painful, disgusted, and neutral facial expressions were presented in absence of a preceding contextual information, and 4 conditions in which painful and disgusted expressions were presented following either a consistent or inconsistent context. These seven conditions were modeled through a boxcar function corresponding to each video duration. Furthermore, 3 kinds of contexts were modeled separately as 3 seconds events. We accounted for habituation effects in neural responses by using the time-modulation option implemented in SPM, which creates, for each condition, an additional regressor in which the trial order is modulated parametrically. This led to a total of 20 regressors (10 main conditions + 10 time-modulators) that were convolved with a canonical hemodynamic response function and associated with regressors describing their first-order time derivative. To account for movement-related variance, physiological-related artifacts, and other sources of noise, we also included the 6 realignment parameters, dummy variables' signaling outlier scans (from Artifact Detection Tools), and an estimate of cardiac- and inspiration-induced changes in the signal based on PhysIO toolbox (Kasper et al., 2017). Low-frequency signal drifts were filtered using a cutoff period of 128 seconds. Serial correlations in the neural signal were accounted through exponential covariance structures, as implemented in the 'FAST' option of SPM. Global scaling was applied.

Functional contrasts, testing differential parameter estimates images associated with one experimental condition *vs.* the other were then fed in a second level, one-sample t-test using random-effect analysis. Effects were considered significant if exceeded $p < 0.05$, family-wise correction for multiple comparisons at the cluster level, with an underlying height threshold of $p < 0.001$, uncorrected (Flandin & Friston, 2019).

Results

Behavioral Data

Preliminary analysis

Of the 38 participants recruited for Experiment 1, 8 did not carry out the unpleasantness rating task due to technical issues. On the remaining population, we first analyzed participants' performance in the catch control condition, where they were asked to respond to properties of the contextual phrases. The overall accuracy was 68% across the two sessions (for those who carried out only the classification task, the accuracy was calculated only on one session). However, there was an important inter-individual variability in the performance of this control, with 8 individuals at chance level (50% or less), who were excluded from the final analysis. Hence, the final sample for Experiment 1 was 30 participants (17 males, mean age = 23.43 ± 4.57) for the Face Classification and 22 (12 males, mean age = 23.86 ± 5.06) for the Unpleasantness Rating. The high number of excluded people reveals the suboptimal nature of the catch control condition from Experiment 1. Consequently, Experiment 2 was a modified version of the Unpleasantness Rating session from Experiment 1, without such control, but with ~30% of trials involving rating the Unpleasantness of the contexts themselves, rather than the facial expressions (see methods section). This ensured that the participants paid attention also to contexts throughout the experimental session.

Classification Task

Table 3 reports full details on the statistical analysis and associated results. When analyzing accuracy as function of *Expression*, we found no difference between the classification of pain (Accuracy: 63.23% ± 17.87) and disgust (64.83% ± 19.00; *fP* – *fD*, $z = -1.30$, $p = 0.193$). Instead, neutral expressions were classified with significantly higher accuracy (91.90% ± 13.65; *fN* – *fD,* $z = 3.71$, $p < 0.001$). Furthermore, accuracy was influenced by the preceding *Context*. Specifically, when processing disgusted expressions, participants were less accurate when the video-clips were preceded by a pain (58.60% ± 27.35) as opposed to a disgust context (71.08% ± 19.96; *cDfD* – *cPfD*, $z = 2.12$, $p = 0.033$). We found an effect with opposite direction when participants processed painful faces, leading to a significant *Expression*Context* interaction ([*cDfD* – *cPfD*] – [*cDfP* – *cPfP*], $z = 2.22$, $p = 0.026$) revealing that pain facial expressions were processed with higher accuracy when preceded by a pain (66.84% ± 22.54) as opposed to a disgust context (62.98% ± 26.72). We found no interaction effect associated to neutral expressions.

This interaction effect confirms previous results showing that contextual sentences can influence the accuracy in subsequent face classification (Carroll & Russell, 1996; Stewart et al., 2019). In our task, participants were presented with two sources of information (facial expression and sentence on a contextual information). While explicitly instructed to answer only accordingly to the former, our results suggest that also the latter information source (contextual sentences) was processed and contributed to the participants' performance. However, this effect could be interpreted in two different ways: on the one side, context might influence the evaluation of the expression, reflecting true context-face integration; on the other side context could merely facilitate the pre-selection of a given response regardless of the facial information available. To shed more light on the mechanisms underlying contextual effects on facial processing, we repeated the analysis by modeling the occurrence of pain/disgust/neutral responses instead of accuracy. Results are described in Figure 2A and Table 3 and confirm that the response likelihood is influenced by *Expression* and *Context*, without any interaction. More specifically, participants were more likely to select pain responses when processing a face expression following any context (*fP* – *fD*: z = 4.81, *p* < 0.001) and, independently, when any face was preceded by painful contexts (*cP* – *cD: z* = 2.04, *p* = 0.042; Figure 2A, left subplot). Likewise, participants were more likely to select disgust responses when processing a disgust expression following any context (*fD* – *fP*: z = 4.40, *p* < 0.001) and, independently, when any face was preceded by disgusting contexts (*cD* – *cP: z* = 2.06, *p* = 0.039; Figure 2A, middle subplot). Instead, neutral responses were modulated exclusively by the facial expressions, with higher likelihood of correct answer when processing neutral faces (*fN* – *fD*: z = 6.61, *p* < 0.001, Figure 2A, right subplot), without any context effect. Overall, contexts influenced the appraisal of videos in an *additive* fashion, that is by increasing the likelihood of selecting the response suggested by the context, independently of the subsequent face. No significant effect was associated with the Response Times of correct responses.

Unpleasantness Rating Task

The Unpleasantness Rating task was devised as a most stringent (albeit indirect) way to assess whether context affected the processing of facial expressions. Indeed, as responses are labelled in terms of unpleasantness (matched and orthogonal between pain/disgust), any contextual influence in face processing could not have been interpreted in terms of response preselection. In this view, both Experiment 1 & 2 confirm that unpleasantness was influenced exclusively by *Expressions*, with no difference between pain and disgust faces (*fP* – *fD*, *t* ≤ 0.91, *p* ≥ 0.378), but less unpleasant ratings for neutral expressions (*fN* – *fD, t* ≤ -6.16, *p* < 0.001; see Figure 2B-C and Tables 2-3). In neither experiment, ratings were influenced by the *Context* main effect, or by the *Expression*Context* interaction.

[Figure 2 here]

[Table 3 here]

Neural Activity

Facial Expressions

In Experiment 2 we analyzed the neural activity evoked by the Unpleasantness Rating task. Table 4 lists the regions implicated in processing facial expressions in absence of previous contexts. When compared with neutral expressions, both pain and disgust expressions recruited bilaterally the fusiform gyrus and middle temporal gyrus, extending to the inferior frontal gyrus (Figure 3A). Pain expressions recruited also the right superior temporal sulcus and the right precentral gyrus extending to the inferior frontal gyrus. These same regions were also observed when contrasting directly pain vs. disgust expressions (Figure 3B), whereas no region displayed increased activity for the opposite contrast.

[Table 4 here]

[Figure 3 here]

Contextual Phrases

We also looked at the neural areas implicated in the processing of contextual sentences without associated facial expressions (Table 5). Pain contexts, as compared to neutral ones, implicated the supramarginal gyrus, postcentral gyrus, middle temporal gyrus, posterior insula and frontal operculum (Figure 4A, red blobs). Part of this network was observed also when contrasting pain contexts against disgust ones (Figure 4B, red blobs). Instead, disgust (vs. neutral) contexts recruited portions of the middle temporal gyrus and frontal operculum already observed for the case of pain (Figure 4A, purple blobs) plus the amygdala, extending posteriorly to hippocampus and parahippocampal gyrus (Figure 4A, blue blobs). Furthermore, when contrasting directly disgust contexts against pain ones, we found a network involving the bilateral angular gyrus, temporal pole, precuneus and dorsomedial prefrontal cortex (Figure 4B, blue blobs).

[Figure 4 here]

[Table 5 here]

Effects of Context in Face Processing

Table 3 displays the brain regions implicated in facial expressions followed by contextual phrases (see Table 6). As a first step, we tested for increased activity when an expression was preceded by a pain vs. disgust context (*cP – cD*) and we found increased activity in the precuneus and supramarginal/postcentral gyrus, extending to the central operculum and posterior insula (Figure 5, red blobs) in a subportion of the network implicated in the processing of contexts alone. Instead, the opposite contrast (*cD* – *cP*) showed an increased activity in the inferior frontal gyrus. Furthermore, under a slightly less conservative threshold (FDR cluster correction at $q < 0.05$), we found increased activity also at the level of the angular gyrus (Figure 5, blue blobs), over and around the area associated with the processing of disgust contexts alone.

As a last step, we tested the interaction, specifically the contrast comparing neural response to faces when associated with consistent vs. inconsistent contexts. When correcting for multiple comparisons for the whole brain, no suprathreshold effect was observed. However, following studies that repeatedly implicated the perigenual MPFC in the integration of facial and non-facial cues of affective states (Peelen et al., 2010; Skerry & Saxe, 2014), we computed a small volume correction analysis on ROI combining medial portions areas 10 and 14 (bilaterally) from the Brainnetome atlas (Fan et al., 2016). Within this search area we found a significant interaction effect (see Figure 5 green blobs). No region was implicated in the inverse contrast.

[Figure 5 here]

[Table 6 here]

Discussion

We investigated the role played by contextual information in the processing of spontaneous facial expressions of pain and disgust. We found that contextual cues have an *additive* influence on the classification of faces, by increasing the likelihood of selecting the response implied by the context, regardless of the expression displayed. In a separated experiment, we found that contextual information influenced the neural processing of expressions in multiple ways. The postcentral cortex and angular gyrus, heavily sensitive to painful and disgusting contexts respectively, were also strongly recruited when a face followed said contexts. Furthermore, the perigenual MPFC displayed increased activity when pain and disgust expressions followed consistent contexts, suggesting that the MPFC integrates state-specific information from both facial and non-facial cues.

Networks for facial expressions

The sight of pain and disgust expressions triggered a common set of regions involving the ventral occipital cortex and posterior superior temporal structures. This converges with previous literature describing these regions as part of a core network for face processing (Deen et al., 2015; Duchaine & Yovel, 2015; Said et al., 2010; Schobert et al., 2018). Furthermore, pain expressions preferentially activated the superior temporal sulcus in all its length. This possibly reflects the differential facial response patterns between the two states, as pain usually triggers more frequently mouth movements than disgust (Dirupo et al., 2022; Kunz et al., 2013), and the anterior-ventral superior temporal sulcus was found associated with movements of the lower portion of the face (Schobert et al., 2018). Alternatively, the pain-preferential activity might underlie a representation of the painful characteristics of the face, as suggested for the activity in the inferior frontal gyrus and neighboring insula (see Ding et al., 2019; Jauniaux et al., 2019; Timmers et al., 2018 as meta-analyses). Contrary to previous experiments (Jabbi et al., 2007; Wicker et al., 2003; Jauniaux et al., 2019; Timmers et al., 2018; Zhao et al., 2021; 2022), our study finds little response of insular and middle cingulate activity to affective facial responses (especially in the case of disgust). It should be stressed, however, that our dataset was characterized by entirely spontaneous expressions (without any extra-facial supporting information). Instead, previous studies relied often on actors which could have led to a more pronounced and stereotypical facial configuration and, in turn, different neural activations.

Networks for Contextual Information

The analysis of contextual sentences revealed a dissociation between supramarginal gyrus, postcentral gyrus and posterior insula, sensitive to pain-related phrases, and angular gyrus, temporo-parietal junction and hippocampus/amygdala, sensitive to disgust contexts. The bilateral frontal operculum appeared implicated in both states. These results converge with previous studies on verbal descriptions of physical pain (Bruneau et al., 2012, 2013; Corradi-Dell'Acqua et al., 2014, 2020; Gu & Han, 2007; Jacoby et al., 2016) which is thought to trigger similar neural responses to those observed for self-directed experiences (Corradi-Dell'Acqua et al., 2014, 2023). A similar interpretation could fit the hippocampus/amygdala, often implicated in first-hand experience of core disgust (Gan et al., 2023; Gan, Zhou, Li, Jiao, Jiang, Biswald, et al., 2022; Sharvit et al., 2020). As for the frontal operculum, previous studies suggest that the neural response of this region (and the neighboring dorsal anterior insula) might underlie a broad coding of unpleasantness shared between pain and disgust (Corradi-Dell'Acqua et al., 2016).

Previous studies consistently reported a dissociation between the supramarginal, postcentral and insular structures, responding to sentences of pain and unpleasant somatic sensations, and the angular gyrus and temporo-parietal cortex, responding to non-somatic affective (Bruneau et al., 2012, 2013; Corradi-Dell'Acqua et al., 2014) and mental states like thoughts and believes (Mar, 2011; Saxe & Powell, 2006). It has been suggested that temporo-parietal regions are involved in processing people's affective states *via* their beliefs/thoughts (Corradi-Dell'Acqua et al., 2014). This interpretation fits with our findings, as disgust is grounded on evaluations about potential intoxications/contaminations (Rozin et al., 1993), and therefore its inference in others might underlie our representation of people's beliefs about those risks.

Contextual influence in networks for facial expressions

We found that the precuneus, the supramarginal, poscentral and opercular gyrii, showed an *additive* effect for contextual cues, with enhanced activity when a facial expression was preceded by pain (vs. disgust) contexts. Importantly, this activation (Figure 5, red blob) is part of a larger cluster involved in pain-related sentences alone (Figure 4, red blobs), suggesting that representation of contexts is subsequently reinstated when processing an expression potentially in line with such information. Our results are in keeping with Zhao et al. (2021) who showed that contextual cues informing about whether painful expressions were genuine (vs. simulated) enhanced supramarginal/postcentral activity. Hence, the combined information between present and previous research suggests that this region plays a key role in interpreting facial information in light of pain-relevant prior knowledge, possibly reflecting a broader mechanism for matching pain representations from different sources of information (Lamm et al., 2016).

Also disgust context exerted an *additive* effect on face processing networks, enhancing the activity at the level of angular gyrus and IFG. These results converge partly with Zhao et al. (2022), who tested how reliability cues influenced the processing of facial responses to disgusting odors, and found as well that IFG activity was higher when contexts suggest the true nature of the expression. Importantly, this prior study implicated also other structures, like midbrain olfactory cortex. Please notice, however, that in our research disgust-related contexts described an ample range of eliciting events (visual, auditory, gustatory, etc.), whereas only facial expressions were manipulated through olfaction. Hence, our contextual modulations at the level of angular gyrus and temporo-parietal cortex should be interpreted as part of a general mechanisms for disgust and non-somatic emotion appraisal (Corradi-Dell'Acqua et al., 2014) which is not idiosyncratic to one sensory channel.

Most critically, perigenual MPFC showed enhanced activity whenever a facial expression was paired with a consistent (vs. inconsistent) context. Hence, MPFC operates in a state-conditional way, by distinguishing whether different sources of information are coherent with one another. Our results are in keeping with previous studies suggesting that this region represents people's affect across state-specific patterns, independently from the stimulus source (Peelen et al., 2010; Skerry & Saxe, 2014). However, previous effects could have been driven by a more general representation of valence, as Skerry and Saxe (2014) compared exclusively positive vs. negative affect, and Peleen et al. (2010) implemented a wide range of emotions but found stronger differentiation in MPFC between happiness and all negative states. In this perspective, the present study provides very reliable evidence that MPFC represents specific, but comparably unpleasant, states in others across multiple integrated sources of information.

Further considerations and overall conclusions

Overall, context influenced the networks for face processing in both an *additive* and *multiplicative* fashion. This mirrors partially the *additive* results from Experiment 1, whereby context increases relevant classifications regardless of the displayed face (Figure 2). However, whereas classification results from Experiment 1 could be explained also in terms of response pre-selection, this is not the case for the neuroimaging data, as in the Unpleasantness Rating task response selection occurs along a dimension orthogonal to "pain" and "disgust" categories. Hence, Experiment 2 provides a more stringent evidence that context influences facial processing in *additive* fashion, unveiling also the neural structures that promote specific face categorization (e.g., supramarginal/postcentral for pain). Unfortunately, despite its inherent interpretational advantages, Experiment 2 does not allow us to link directly brain responses with overt interpretation of facial expressions.

In this study, we exploited a dataset of spontaneous dynamic facial expressions, characterized by rubber cannulas connected to the face's nostrils (Figure 1; see Dirupo et al., 2020, 2022, for more details). This might have influenced negatively the plausibility of the experimental set-up, as none of the manipulated contexts involved odorants delivered through tubes. We believe that the cannulas (present constantly in all videos) and potential plausibility considerations had negligible influences on our results, especially considering that participants were not required to explicitly compare faces with previous sentences, which operated instead as a task-irrelevant information.

Finally, although our findings provided converging evidence with prior neuroimaging results (see above), it is unclear how to interpret discrepancies, as some of these studies adopted different approaches, and manipulated contexts as task-relevant information. Future research will need to examine more thoroughly the role played by task demands in the networks mediating contextual-facial integration.

In conclusion, our study is a systematic investigation of the cognitive and neural processes mediating contextual influences on affective face processing. Across two experiments we found that individuals partly classify the expressions based on contextual information, regardless of the facial information displayed. This effect was further supported by evidence that neural structures specifically implicated in pain and disgust contexts, were subsequently reactivated for any expression following said context. Additionally, we found that the perigenual MPFC discriminated between face-context pairings that were consistent (vs. inconsistent) from one another. Overall, our study unveils key neural processes underlying the coding of state-specific information from both face and context, and sheds new light on how they are integrated within the MPFC.

References

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body Cues, Not Facial Expressions, Discriminate Between Intense Positive and Negative Emotions. *Science*, *338*(6111), 1225–1229. https://doi.org/10.1126/science.1224313

Bruneau, E. G., Dufour, N., & Saxe, R. (2013). How We Know It Hurts: Item Analysis of Written Narratives Reveals Distinct Neural Responses to Others' Physical Pain and Emotional Suffering. *PLOS ONE*, *8*(4), e63085. https://doi.org/10.1371/journal.pone.0063085

Bruneau, E. G., Pluta, A., & Saxe, R. (2012). Distinct roles of the 'Shared Pain' and 'Theory of Mind' networks in processing others' emotional suffering. *Neuropsychologia*, *50*(2), 219–231. https://doi.org/10.1016/j.neuropsychologia.2011.11.008

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*.

Corradi-Dell'Acqua, C., Hofstetter, C., Sharvit, G., Hugli, O., & Vuilleumier, P. (2023). Healthcare experience affects pain-specific responses to others' suffering in the anterior insula. *Human Brain Mapping*. https://doi.org/10.1002/HBM.26468

Corradi-Dell'Acqua, C., Hofstetter, C., & Vuilleumier, P. (2014). Cognitive and affective theory of mind share the same local patterns of activity in posterior temporal but not medial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, *9*(8), 1175–1184. https://doi.org/10.1093/scan/nst097

Corradi-Dell'Acqua, C., Ronchi, R., Thomasson, M., Bernati, T., Saj, A., & Vuilleumier, P. (2020). Deficits in cognitive and affective theory of mind relate to dissociated lesion patterns in prefrontal and insular cortex. *Cortex*, *128*, 218–233. https://doi.org/10.1016/j.cortex.2020.03.019

Corradi-Dell'Acqua, C., Tusche, A., Vuilleumier, P., & Singer, T. (2016). Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. *Nature Communications*, *7*, 10904. https://doi.org/10.1038/ncomms10904

De Ruddere, L., Bosmans, M., Crombez, G., & Goubert, L. (2016). Patients Are Socially Excluded When Their Pain Has No Medical Explanation. *The Journal of Pain*, *17*(9), 1028–1035. https://doi.org/10.1016/j.jpain.2016.06.005

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex*, *25*(11), 4596–4609.

https://doi.org/10.1093/cercor/bhv111

Ding, R., Ren, J., Li, S., Zhu, X., Zhang, K., & Luo, W. (2019). Domain-general and domain-preferential neural correlates underlying empathy towards physical pain, emotional situation and emotional faces: An ALE meta-analysis. *Neuropsychologia*, 107286. https://doi.org/10.1016/j.neuropsychologia.2019.107286

Dirupo, G., Corradi-Dell'Acqua, C., Kashef, M., Debbané, M., & Badoud, D. (2020). The role of interoception in understanding others' affect. Dissociation between superficial and detailed appraisal of facial expressions. *Cortex*, *130*, 16–31. https://doi.org/10.1016/j.cortex.2020.05.010

Dirupo, G., Garlasco, P., Chappuis, C., Sharvit, G., & Corradi-Dell'Acqua, C. (2022). State-specific and supraordinal components of facial response to pain. *IEEE Transactions on Affective Computing*, *13*(2), 793–804. https://doi.org/10.1109/TAFFC.2020.2965105

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex*, *26*(8), 3508–3526. https://doi.org/10.1093/cercor/bhw157

Flandin, G., & Friston, K. J. (2019). Analysis of family-wise error rates in statistical parametric mapping using random field theory. *Human Brain Mapping*, *40*(7), 2052–2054. https://doi.org/10.1002/hbm.23839

Gan, X., Zhou, F., Xu, T., Liu, X., Zhang, R., Zheng, Z., Yang, X., Zhou, X., Yu, F., Li, J., Cui, R., Wang, L., Yuan, J., Yao, D., & Becker, B. (2023). Rotten to the core – a neurofunctional signature of subjective core disgust generalizes to socio-moral contexts. *BioRxiv*.

Gan, X., Zhou, X., Li, J., Jiao, G., Jiang, X., Biswal, B., Yao, S., Klugah-Brown, B., & Becker, B. (2022). Common and distinct neurofunctional representations of core and social disgust in the brain: Coordinate-based and network meta-analyses. *Neuroscience & Biobehavioral Reviews*, *135*, 104553. https://doi.org/10.1016/j.neubiorev.2022.104553

Gan, X., Zhou, X., Li, J., Jiao, G., Jiang, X., Biswald, B., Yao, S., Klugah-Brown, B., & Becker, B. (2022). Common and distinct neurofunctional representations of core and social disgust in the brain: An ALE meta-analysis and MACM characterization. *Neuroscience & Biobehavioral Reviews*.

Gu, X., & Han, S. (2007). Neural substrates underlying evaluation of pain in actions depicted in words. *Behavioural Brain Research*, *181*(2), 218–223. https://doi.org/10.1016/j.bbr.2007.04.008

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences*, *16*(5), 701–713. https://doi.org/10.1016/0191-8869(94)90212-7

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Jacoby, N., Bruneau, E., Koster-Hale, J., & Saxe, R. (2016). Localizing Pain Matrix and Theory of Mind networks with both verbal and non-verbal stimuli. *NeuroImage*, *126*, 39–48. https://doi.org/10.1016/j.neuroimage.2015.11.025

Jauniaux, J., Khatibi, A., Rainville, P., & Jackson, P. L. (2019). A meta-analysis of neuroimaging studies on pain empathy: Investigating the role of visual information and observers' perspective. *Social Cognitive and Affective Neuroscience*, *14*(8), 789–813. https://doi.org/10.1093/scan/nsz055

Kasper, L., Bollmann, S., Diaconescu, A. O., Hutton, C., Heinzle, J., Iglesias, S., Hauser, T. U., Sebold, M., Manjaly, Z. M., Pruessmann, K. P., & Stephan, K. E. (2017). The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. *Journal of Neuroscience Methods*, *276*, 56–72. https://doi.org/10.1016/j.jneumeth.2016.10.019

Kunz, M., Peter, J., Huster, S., & Lautenbacher, S. (2013). Pain and Disgust: The Facial Signaling of Two Aversive Bodily Experiences. *PLoS ONE*, *8*(12), e83277. https://doi.org/10.1371/journal.pone.0083277

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Lamm, C., Batson, C. D., & Decety, J. (2007). The Neural Substrate of Human Empathy: Effects of Perspective-taking and Cognitive Appraisal. *Journal of Cognitive Neuroscience*, *19*(1), 42–58. https://doi.org/10.1162/jocn.2007.19.1.42

Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self–other representations in empathy: Evidence from neurotypical function and socio-cognitive disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150083. https://doi.org/10.1098/rstb.2015.0083

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, *62*, 103–134. https://doi.org/10.1146/annurev-psych-120709-145406

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524. https://doi.org/10.3758/BF03195598

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal Representations of Perceived Emotions in the Human Brain. *Journal of Neuroscience*, *30*(30), 10127–10134. https://doi.org/10.1523/JNEUROSCI.2161-10.2010

Righart, R., & Gelder, B. de. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(3), 264–272. https://doi.org/10.3758/CABN.8.3.264

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

Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, *10*(5), 11.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science: A Journal of the American Psychological Society / APS*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Schobert, A.-K., Corradi-Dell'Acqua, C., Frühholz, S., van der Zwaag, W., & Vuilleumier, P. (2018). Functional organization of face processing in the human superior temporal sulcus: A 7T high-resolution fMRI study. *Social Cognitive and Affective Neuroscience*, *13*(1), 102–113. https://doi.org/10.1093/scan/nsx119

Sharvit, G., Lin, E., Vuilleumier, P., & Corradi-Dell'Acqua, C. (2020). Does inappropriate behavior hurt or stink? The interplay between neural representations of somatic experiences and moral decisions. *Science Advances*, *6*(42), eaat4390. https://doi.org/10.1126/sciadv.aat4390

Skerry, A. E., & Saxe, R. (2014). A Common Neural Code for Perceived and Inferred Emotion. *Journal of Neuroscience*, *34*(48), 15997–16008. https://doi.org/10.1523/JNEUROSCI.1676-14.2014

Stewart, S. L. K., Schepman, A., Haigh, M., McHugh, R., & Stewart, A. J. (2019). Affective theory of mind inferences contextually influence the recognition of emotional facial expressions. *Cognition and Emotion*, *33*(2), 272–287. https://doi.org/10.1080/02699931.2018.1450224

Timmers, I., Park, A. L., Fischer, M. D., Kronman, C. A., Heathcote, L. C., Hernandez, J. M., & Simons, L. E. (2018). Is Empathy for Pain Unique in Its Neural Correlates? A Meta-Analysis of Neuroimaging Studies of Empathy. *Frontiers in Behavioral Neuroscience*, *12*. https://doi.org/10.3389/fnbeh.2018.00289

Vrticka, P., Andersson, F., Sander, D., & Vuilleumier, P. (2009). Memory for friends or foes: The social context of past encounters with faces modulates their subsequent neural traces in the brain. *Social Neuroscience*, *4*(5), 384–401. https://doi.org/10.1080/17470910902941793

Vrticka, P., Sander, D., & Vuilleumier, P. (2011). Effects of Emotion Regulation Strategy on Brain Responses to the Valence and Social Content of Visual Scenes. *Neuropsychologia*, *49*(5), 1067–1082.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, *7*(11), 1271–1278. https://doi.org/10.1038/nn1341

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, *142*(5), 472–497. https://doi.org/10.1037/bul0000030

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, *2*(3), 125–141. https://doi.org/10.1089/brain.2012.0073

Wieser, M. J., Gerdes, A. B. M., Greiner, R., Reicherts, P., & Pauli, P. (2012). Tonic pain grabs attention, but leaves the processing of facial expressions intact—Evidence from event-related brain potentials. *Biological Psychology*, *90*(3), 242–248. https://doi.org/10.1016/j.biopsycho.2012.03.019

Zhao, Y., Zhang, L., Rütgen, M., Sladky, R., & Lamm, C. (2021). Neural dynamics between anterior insular cortex and right supramarginal gyrus dissociate genuine affect sharing from perceptual saliency of pretended pain. *eLife*, *10*, e69994. https://doi.org/10.7554/eLife.69994

Zhao, Y., Zhang, L., Rütgen, M., Sladky, R., & Lamm, C. (2022). Effective connectivity reveals distinctive patterns in response to others' genuine affective experience of disgust. *NeuroImage*, *259*, 119404. https://doi.org/10.1016/j.neuroimage.2022.119404

Tables

	Pain vs. Disg.	*Pain vs. Neutral*	*Disg. vs. Neutral*
***Preliminary Analysis*[1]**			
Self-Reported Unpleasantness	$t_{(35)} = 1.27$	$t_{(12)} = \mathbf{15.68}^{***}$	$t_{(10)} = \mathbf{12.53}^{***}$
Observers' Unpleasantness	$t_{(16)} = -0.67$	-	-
Observers' Accuracy	$z = 1.30$	$z = \mathbf{-3.32}^{***}$	$z = \mathbf{-5.47}^{***}$
Experiment 2			
Unpleasantness	$t_{(10)} = -0.88$	$t_{(21)} = \mathbf{7.79}^{***}$	$t_{(10)} = \mathbf{8.36}^{***}$

[1] data from Dirupo et al. (2020).

***Table 1**: Results from preliminary analysis testing dichotomic effects of Expressions (Painful vs. Disgusted, Pain vs. Neutral, Disgusted vs. Neutral) as within-subjects factor. Full details for the development and validation of the video-database is available in Dirupo et al. (2020). For consistency purposes, the table reports also the data from Experiment 2 involving the processing of facial expressions in absence of preceding contexts. The lmer-syntax of the tested models is the following:*

Model 1: Self-Reported Unpleasantness ~ Expressions + (Expressions| Portrayed person).
Model 2 & 4: Observers' Unpleasantness ~ Expressions + (Expressions| Subjects) + (Expressions| Portrayed person).
Model 3: Observers' Accuracy ~ Expressions + (Expressions| Subjects) + (Expressions| Portrayed person).

Models 1-2 & 4 were Linear Mixed Model, and effect significance was calculated using the Satterthwaite approximation of the degrees of freedom. For the analysis of classification accuracy, we employed a Generalized Linear Mixed Model with a binomial distribution and Laplace approximation. For each dependent variable (displayed horizontally), and for each effect of interest (vertically), the table reports the associated t/z-values Significant effects are highlighted.

	Pain vs. Disg.	*Pain vs. Neutral*	*Disg. vs. Neutral*
Sentence Information			
Lexical Frequency	$t_{(26)}$ = 0.64	$t_{(26)}$ = -0.57	$t_{(26)}$ = 0.06
# Characters	$t_{(26)}$ = 1.54	$t_{(26)}$ = 0.92	$t_{(26)}$ = 1.79
Pilot 1			
Pain Rating	$t_{(32)}$ = **7.52*****	$t_{(30)}$ = **8.65*****	$t_{(25)}$ = **2.84****
Disgust Ratings	$t_{(36)}$ = **-8.51*****	$t_{(38)}$ = **3.08****	$t_{(28)}$ = **8.66*****
Unpleasantness Ratings	$t_{(59)}$ = -0.47	$t_{(30)}$ = **7.78*****	$t_{(32)}$ = **8.54*****
Experiment 2			
Unpleasantness Ratings	$t_{(44)}$ = 0.06	$t_{(53)}$ = **15.01*****	$t_{(44)}$ = **15.69*****

Table 2: *Contextual information. Data from 81 French sentences, organized in 27 triplets, matched for lexical frequency and length, but descriptive of a Painful, Disgusted or Neutral context (see Methods). The table reports pairwise t-test comparisons (Painful vs. Disgusted, Pain vs. Neutral, Disgusted vs. Neutral) for Lexical Frequency, and sentence length. Furthermore, the results from Pilot 1, where 24 independent individuals rated each sentence in terms of Pain, Disgust and Unpleasantness are also displayed. Finally, for consistency purposes, the table reports also the data from Experiment 2 involving the rating of contextual sentences without associated facial expressions. In these cases, the analyses were carried out with linear mixed model, with the following lmer syntax:*

Rating ~ Category + (Category | Subjects) + (1 | Sentence)

	Expressions		***Context***		***Expressions*Context***			
	fP – fD	*fN – fD*	*cP – cD*	*cN – cD*	*(fP – fD)* (cP – cD)*	*(fN – fD)* (cP – cD)*	*(fP – fD)* (cN – cD)*	*(fN – fD)* (cN – cD)*
Experiment 1								
Accuracy	*z* = -1.30	*z* = **3.71*****	*z* = **-2.12***	*z* = -0.94	*z* = **2.22***	*z* = 0.54	*z* = 1.48	*z* = 0.44
Pain Resp.	*z* = **4.81*****	*z* = -1.84	*z* = **2.04***	*z* = 0.65	*z* = -0.99	*z* = 0.27	*z* = 0.40	*z* = -0.17
Disgust Resp.	*z* = **-4.40*****	*z* = **-5.63*****	*z* = **-2.06***	*z* = -1.00	*z* = 0.15	*z* = 0.45	*z* = -0.33	*z* = -0.11
Neutral Resp.	*z* = 0.19	*z* = **6.61*****	*z* = -0.24	*z* = -0.99	*z* = 0.60	*z* = 0.02	*z* = 0.61	*z* = 0.88
Reaction-Times	$t_{(25)}$ = 0.44	$t_{(25)}$ = -1.80	$t_{(37)}$ = -0.65	$t_{(29)}$ = 0.04	$t_{(30)}$ = -0.45	$t_{(40)}$ = 0.26	$t_{(31)}$ = -0.83	$t_{(28)}$ = 0.70
Unpleasant.	$t_{(16)}$ = 0.91	$t_{(23)}$ = **-6.16*****	$t_{(22)}$ = 0.97	$t_{(21)}$ = -0.50	$t_{(18)}$ = -1.21	$t_{(24)}$ = -0.60	$t_{(15)}$ = -1.26	$t_{(21)}$ = -1.08
Experiment 2								
Unpleasant.	$t_{(11)}$ = 0.61	-	$t_{(16)}$ = 0.56	-	$t_{(36)}$ = 0.68	-	-	-

Table 3*: Contextual effects on facial processing. Analyses from Experiments 1 & 2 describing the effect of contextual information on the classification (Experiment 1) or unpleasantness rating (Experiments 1-2) of facial expressions. All analyses were carried out through a linear mixed model scheme, with the following lmer syntax:*

*Model 1: Accuracy ~ Expressions*Context + (Expressions*Context | Subjects) + (Expressions*Context | Portrayed Person) + (Expressions | Sentence)*
*Model 2: Pain Responses ~ Expressions*Context + (Expressions*Context | Subjects) + (Expressions+Context | Portrayed Person) + (1 | Sentence)*
*Models 3-4: Disgust/Neutral Responses ~ Expressions*Context + (Expressions*Context | Subjects) + (Expressions*Context | Portrayed Person) + (Expressions | Sentence)*
*Model 5: Reaction Times ~ Expressions*Context + (Expressions*Context | Subjects) + (Expressions+Context | Portrayed Person) + (Expressions | Sentence)*
*Models 6-7: Unpleasantness Ratings ~ Expressions*Context + (Expressions*Context | Subjects) + (Expressions*Context | Portrayed Person) + (Expressions | Sentence)*

Models 1-4 were Generalized Linear Mixed Model with a binomial distribution and Laplace approximation. Instead, models 5-7 were Linear Mixed Model, and effect significance was calculated using the Satterthwaite approximation of the degrees of freedom. For each dependent variable (displayed horizontally), and for each effect of interest (vertically), the table reports the associated t/z-values. For Experiment 2 we implemented a simpler design, underlying a restricted number of fixed effect terms. Significant effects are highlighted.

	SIDE	***Coordinates***			$T_{(25)}$	***Cluster size***
		X	***Y***	***Z***		
Painful – Neutral Expressions						
Fusiform Gyrus	R	44	-44	-20	6.40	
Inferior Occipital Gyrus	R	46	-68	0	8.98	
Middle Temporal Gyrus	R	50	-60	8	9.28	4387***
Superior Temporal Sulcus (*posterior*)	R	48	-20	-8	6.59	
Fusiform Gyrus	L	-42	-42	-24	6.49	
Inferior Occipital Gyrus	L	-58	-68	6	7.41	3356***
Middle Temporal Gyrus	L	-40	-66	2	8.57	
Precentral Gyrus	R	50	4	40	6.30	723***
Inferior Frontal Gyrus (*opercular*)	R	56	16	26	5.79	
Disgusted – Neutral Expressions						
Fusiform Gyrus	R	44	-44	-18	5.38	
Inferior Occipital Gyrus	R	40	-90	4	6.17	2905***
Middle Temporal Gyrus	R	46	-70	-2	9.14	
Fusiform Gyrus	L	-46	-50	-22	6-19	
Inferior Occipital Gyrus	R	-46	-70	0	6.02	2090***
Middle Temporal Gyrus	R	-44	-58	10	6.82	
Painful – Disgust Expressions						
Inferior Frontal Gyrus (*opercular*)	R	52	20	28	5.48	418**
Superior Temporal Gyrus *(anterior)*	R	52	-16	-10	5.07	425**
Superior Temporal Gyrus *(posterior)*	R	50	-40	12	4.25	

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$ family-wise corrected for the whole brain; † $p < 0.05$ family-wise corrected for the bilateral amygdala.

Table 4: *Regions implicated when observing facial expressions in absence of a preceding context. Unless stated otherwise, all clusters survived correction for multiple comparisons at the cluster level. Coordinates (in standard MNI space) refer to maximally activated foci as indicated by the highest t value within an area of activation: x = distance (mm) to the right (+) or the left (−) of the midsagittal line; y = distance anterior (+) or posterior (−) to the vertical plane through the anterior commissure (AC); z = distance above (+) or below (−) the inter-commissural (AC-PC) line. L and R refer to the left and right hemisphere, respectively.*

	SIDE	*Coordinates*			*$T_{(25)}$*	*Cluster size*
		X	*Y*	*Z*		
Painful – Neutral Contexts						
Posterior Insula	R	40	-4	-6	6.05	417**
Frontal Operculum	R	42	10	6	5.14	
Posterior Insula	L	-40	-8	-4	6.98	414**
Frontal Operculum	L	-44	12	2	4.66	
Superior Temporal Sulcus	L	-54	4	-10	4.66	
Supra Marginal Gyrus	L	-64	-28	44	7.41	887***
Middle Temporal Gyrus	L	-54	-60	0	7.49	680***
Lingual Gyrus	L	38	-80	-12	6.35	330**
Cerebellum	R	26	-66	-26	5.59	367**
Disgust – Neutral Contexts						
Superior Temporal Sulcus	L	-56	6	-10	5.43	219*
Frontal Operculum	L	-44	10	0	4.38	
Middle Temporal Gyrus	L	-56	-60	0	5.71	252*
Amygdala	L	-20	-8	-16	3.46	370**
Hippocampus	L	-18	-22	-12	4.66	
Parahippocampal Gyrus	L	-16	-30	-10	4.54	
Supplementary Motor Area	M	-2	8	64	5.85	241*
Inferior Occipital Gyrus	R	34	-86	-10	7.06	458***
Lingual Gyrus	R	8	-86	-8	4.63	
Cerebellum	R	28	-70	-22	5.85	360**
Painful – Disgust Contexts						
Supra Marginal Gyrus	L	-58	-26	34	11.15	1054***
Precentral Gyrus	L	-46	-2	22	6.24	517***
Posterior Insula	L	-38	-10	8	4.18	
Middle Frontal Gyrus	L	-34	32	20	5.60	259*
Disgust – Painful Contexts						
Angular Gyrus	R	56	-56	26	6.27	292**
Angular Gyrus	L	-40	-62	28	4.65	378**
Temporal Pole	L	-50	-6	-32	5.60	498***
Posterior Cingulate Gyrus	M	4	-46	34	4.89	306***
Precuneus	M	-10	-54	10	5.67	550***
Superior Frontal Gyrus	M	-14	62	26	5.98	319**

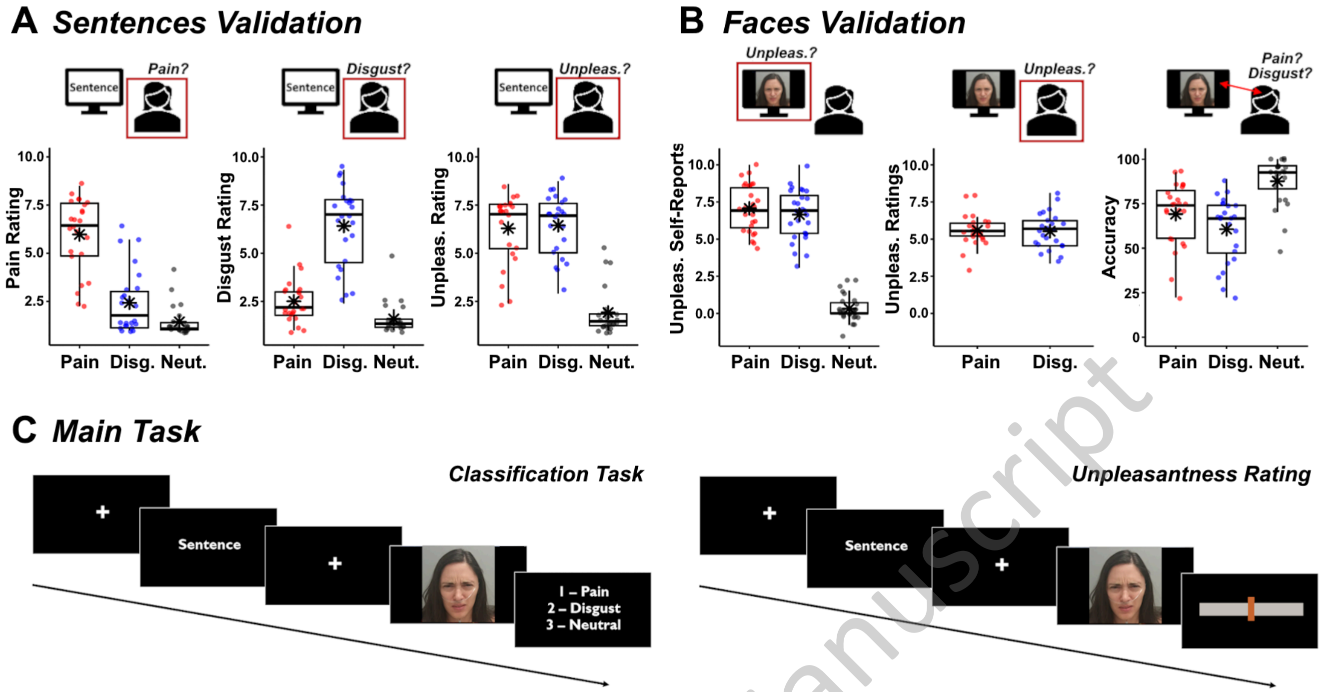
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$ family-wise corrected for the whole brain.

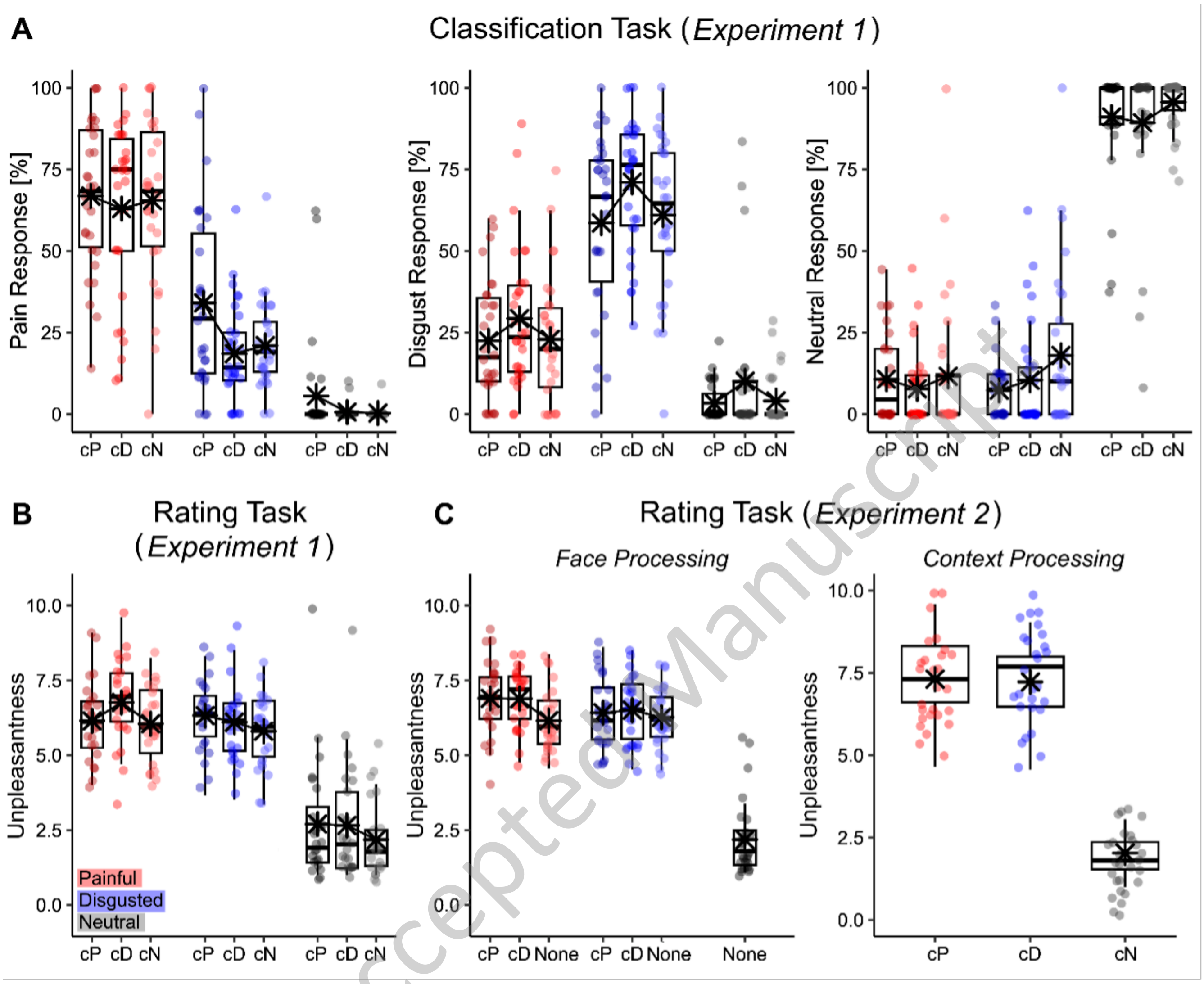
Table 5: *Regions implicated when reading contextual sentences without any associated facial expression. All clusters survived correction for multiple comparisons at the cluster level.*

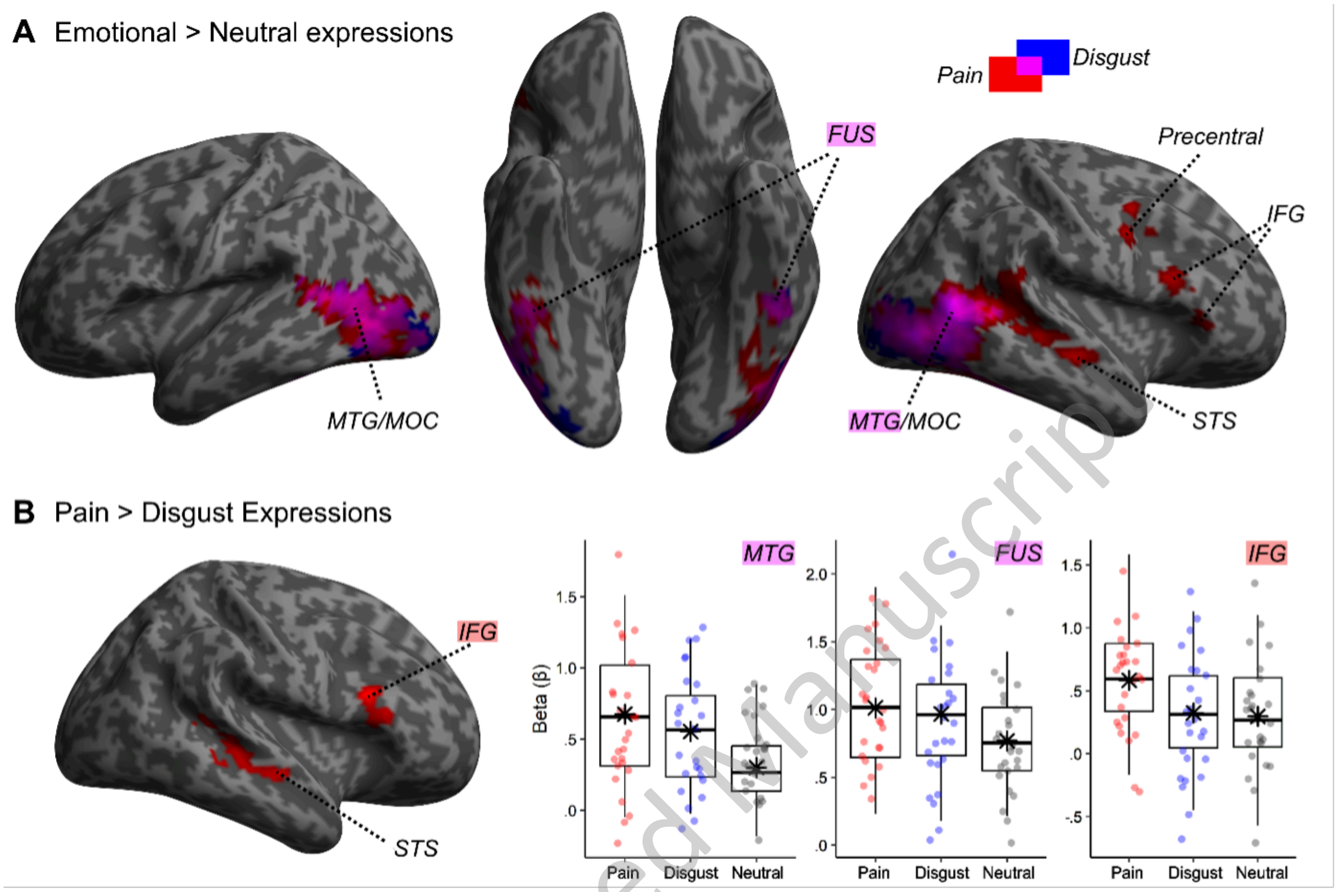
	SIDE	Coordinates			$T_{(25)}$	Cluster size
		X	Y	Z		
Expression following Painful – Disgusting contexts						
Postcentral Gyrus	L	-44	-18	34	4.36	275^{*}
Central Operculum	L	-48	-8	12	5.36	
Posterior Insula	L	-36	-10	12	5.16	
Precuneus	L	-8	-50	68	6.07	1103^{***}
Expression following Disgusting – Painful contexts						
Inferior Frontal Gyrus	L	-50	28	6	6.00	545^{***}
Lateral Orbital Gyrus	L	-40	40	-16	4.20	
Angular Gyrus	L	-52	-54	26	4.41	185
Faces following Coherent - Incoherent context						
Medial Prefrontal Cortex	M	2	60	6	$4.37^{\dagger}$	1

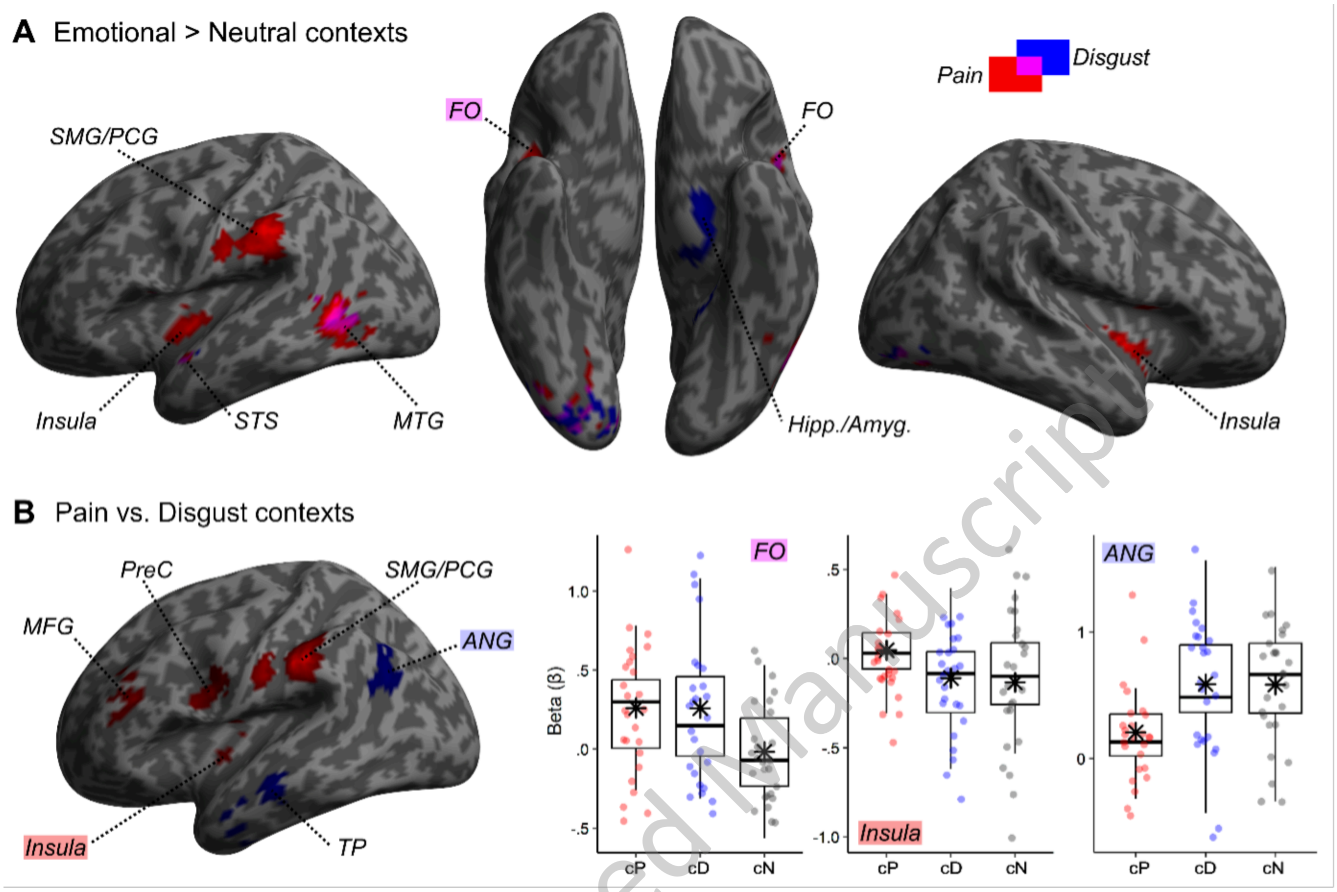
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$ family-wise corrected for the whole brain; † $p < 0.05$ family-wise corrected for the medial prefrontal cortex.

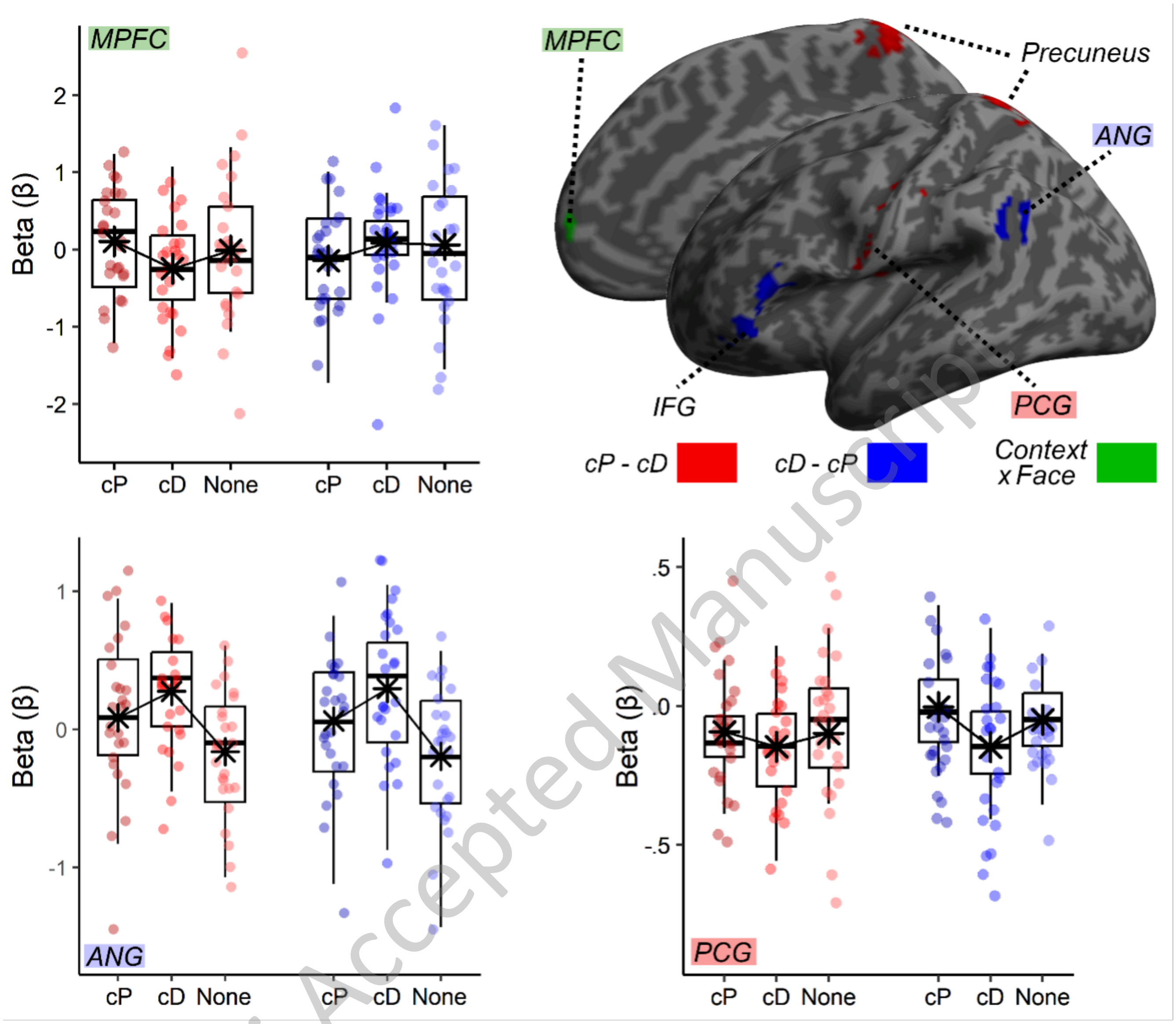
***Table 6**: Brain structures whose response to facial expressions is influenced by the preceding context. Unless stated otherwise, all clusters survived correction for multiple comparisons at the cluster level.*


A Sentences Validation
Sentence
Pain?
Disgust?
Unpleas.?
Pain Rating
Disgust Rating
Unpleas. Rating
10.0
7.5
5.0
2.5
Pain
Disg.
Neut.
B Faces Validation
Unpleas.?
Pain?
Disgust?
Unpleas. Self-Reports
Unpleas. Ratings
Accuracy
0.0
100
75
50
25
0
C Main Task
Classification Task
Unpleasantness Rating
+
Sentence
1 – Pain
2 – Disgust
3 – Neutral


A
Classification Task (Experiment 1)
Pain Response [%]
Disgust Response [%]
Neutral Response [%]
100
75
50
25
0
cP cD cN cP cD cN cP cD cN
B
Rating Task (Experiment 1)
C
Rating Task (Experiment 2)
Face Processing
Context Processing
Unpleasantness
10.0
7.5
5.0
2.5
0.0
Painful
Disgusted
Neutral
cP cD None cP cD None None
cP cD cN


A Emotional > Neutral expressions
Pain
Disgust
FUS
Precentral
IFG
MTG/MOC
MTG/MOC
STS
B Pain > Disgust Expressions
IFG
STS
MTG
FUS
IFG
Beta (β)
1.5
1.0
.5
.0
2.0
1.5
1.0
.5
.0
1.5
1.0
.5
.0
-.5
Pain
Disgust
Neutral
Pain
Disgust
Neutral
Pain
Disgust
Neutral


A Emotional > Neutral contexts
Pain
Disgust
SMG/PCG
FO
FO
Insula
STS
MTG
Hipp./Amyg.
Insula
B Pain vs. Disgust contexts
PreC
SMG/PCG
MFG
ANG
Insula
TP
Beta (β)
1.0
.5
.0
-.5
FO
cP
cD
cN
.5
.0
-.5
-1.0
Insula
cP
cD
cN
ANG
1
0
cP
cD
cN


MPFC
Beta (β)
2
1
0
-1
-2
cP
cD
None
cP
cD
None
MPFC
Precuneus
ANG
IFG
PCG
cP - cD 🟥
cD - cP 🟦
Context x Face 🟩
Beta (β)
1
0
-1
ANG
cP
cD
None
cP
cD
None
Beta (β)
.5
.0
-.5
PCG
cP
cD
None
cP
cD
None