Authors for correspondence:
Maria Serena Panasiti
e-mail: salvatoremaria.aglioti@uniroma1.it
Salvatore Maria Aglioti
e-mail: mariaserena.panasiti@uniroma1.it

THE ROYAL SOCIETY
PUBLISHING

Thermal facial reactivity patterns predict social categorization bias triggered by unconscious and conscious emotional stimuli

Giorgia Ponsi[1,2,†], Maria Serena Panasiti[1,2,†], Giulia Rizza[1,2] and Salvatore Maria Aglioti[1,2]

[1]Department of Psychology, University of Rome 'Sapienza', Via dei Marsi, 78-00185 Rome, Italy
[2]IRCCS Fondazione Santa Lucia, Via Ardeatina, 306-00142 Rome, Italy

GP, 0000-0002-7188-8102; MSP, 0000-0002-3149-8855; SMA, 0000-0001-8175-7563

Members of highly social species decode, interpret, and react to the emotion of a conspecific depending on whether the other belongs to the same (ingroup) or different (outgroup) social group. While studies indicate that consciously perceived emotional stimuli drive social categorization, information about how implicit emotional stimuli and specific physiological signatures affect social categorization is lacking. We addressed this issue by exploring whether subliminal and supraliminal affective priming can influence the categorization of neutral faces as ingroup versus outgroup. Functional infrared thermal imaging was used to investigate whether the effect of affective priming on the categorization decision was moderated by the activation of the sympathetic nervous system (SNS). During the subliminal condition, we found that stronger SNS activation after positive or negative affective primes induced ingroup and outgroup face categorization, respectively. The exact opposite pattern (i.e. outgroup after positive and ingroup after negative primes) was observed in the supraliminal condition. We also found that misattribution effects were stronger in people with low emotional awareness, suggesting that this trait moderates how one recognizes SNS signals and employs them for unrelated decisions. Our results allow the remarkable implication that low-level affective reactions coupled with sympathetic activation may bias social categorization.

1. Background

Emotions are short-lived and intense states of the mind and body triggered by salient stimuli that can be internal or external, negative or positive. Emotion states have subjective, physiological, and behavioural components [1]. Visceral bodily changes represent a major part of emotional responses and act as non-conscious information that guides human behaviour [2]. Autonomic nervous system (ANS) indexes of electrodermal, cardiovascular, and respiratory reactivity are gold standard measures in emotion science [3]. Functional infrared thermal imaging (fITI) is an emerging tool for the study of human emotional and social behaviour that allows for the recording of skin temperature by tracking non-invasively emitted infrared heat with high thermal resolution and a significant noise reduction [4]. A variety of affective responses typically focusing on facial blood flow changes in adult humans, infants, and non-human primates have been investigated and classified [5] by using fITI (e.g. fear [6], joy [7], stress [8], and empathy [9]). Levine *et al.* [6], for example, reported rapid (300 ms) periorbital warming and cheek cooling after a fright response caused by a sudden, strong noise; this startle-induced thermal signature is mediated by the sympathetic nervous system (SNS). More complex emotions are driven by ostracism [10], guilt [11], social [12] and gaze contact [13],

and are accompanied by facial temperature increase (especially in the periorbital, nose, forehead, and lip regions). Also, volitional deception is associated with stress-induced temperature change in the periorbital region [14,15].

Experimental evidence indicates that processing the affective content of a stimulus does not necessarily require conscious awareness [16]. Non-conscious affective stimuli (i.e. implicitly conditioned ones or those presented below the threshold for conscious perception [17]) are, in fact, able to evoke emotional reactions [18], facial micro-expressions [19], and neural responses [20]. They are also able to influence perception of sensory noise [21] and complex behaviours [22] like aesthetic and social preferences [23,24]. In addition, unconscious affective stimuli play a key role in (mis)attribution processes. When individuals are not aware of their own affective reactions, they tend to misattribute those reactions and make unrelated judgements and decisions [25].

This phenomenon was recently investigated with the affective misattribution procedure (AMP, [26]), a paradigm in which participants are generally presented with subliminal affective primes followed by neutral objects such as Chinese pictographs, whose visual pleasantness has to be judged. Data repeatedly show that attractiveness judgements are affected by the primes according to a valence-driven rule [26], and that employing AMP as an implicit measure reveals social impressions [27] and empathic resonance abilities [28,29].

Individuals are continuously exposed to affectively charged cues. Subtle information such as non-verbal behaviour [30], olfactory cues [31], and perception of pain—be it physical [32] or social [33]—all influence perceptions and decisions, especially in the social domain. Thus, socially evoked emotional responses need to be constantly regulated (i.e. modified in quality, intensity, and duration) at both explicit and implicit levels in order to avoid interference with the individual's functioning [34].

Social group membership plays an important role in interpreting the emotions of others. Experimental findings have shown that we are more evolutionarily prepared to empathize [35] and to respond to the emotions of those we identify as being in our group (ingroup) than those we do not (outgroup) [36]. Interestingly, facial markers of emotional reactions like Schadenfreude (i.e. the pleasure derived from others' misfortune) are particularly strong towards outgroups [37]. Since our emotional decoding of others is strongly affected by what group they belong to, it could be that the reverse is also true, namely that the social categorization mechanism is influenced by the affective state of the perceiver. Miller et al. [38] investigated this inverse relationship, showing that threat-related cues produced a bias towards outgroup categorization, especially in individuals vulnerable to interpersonal threats. Furthermore, Krosch & Amodio [39] showed that priming economic scarcity altered perception of race and that this triggered disparity in money allocation. Here, we tested whether subliminal and supraliminal affective priming influences the attribution of emotionally neutral faces that lacked any perceptual or reputational ties to a given social group. Our experimental design helped us to actively manipulate the role of the prime's affective valence (positive, negative, and neutral), distinguishing it from the role of the participant's visual awareness (absent in the subliminal block and present in the supraliminal one) in influencing social categorization decisions. Previous research

on the AMP [26] has typically focused on abstract visual stimuli like Chinese pictographs (see [27] for an exception). Despite their undeniable usefulness as non-informative stimuli, Chinese pictographs have poor ecological validity because they are not a target of prejudice in the social world. For example, Heerdink et al. [40] asked participants to guess whether the presented Chinese pictograms represented words related to rejection or acceptance. They found that primed anger enhanced rejection judgements while primed happiness caused higher acceptance ones. Furthermore, recent research shows that sympathetic and neural activity caused by threatening stimuli predicts decreased likability of neutral faces, even in the absence of visual awareness [41,42].

We expand previous knowledge by (i) focusing on neutral faces as a potential target for affective misattribution processes and (ii) investigating how affective information influences group coding mechanisms. Our participants were presented with subliminal and supraliminal affective primes before having to categorize faces that lacked any group-related cue as ingroup or outgroup (figure 1). During the experimental task, we also measured participants' autonomic activity through fITI (see electronic supplementary material, figure S1) in order to obtain a reliable index of emotional processing and to verify whether ANS reactivity has a role in modulating social decisions. Additionally, we investigated how individual differences in trait emotional awareness (EA) determine social categorization and physiological reactivity.

2. Material and methods

(a) Participants

Thirty-three Italian participants (10 male; age range 20–38 years, $M = 25.00$, s.d. $= 3.92$) from the University of Rome 'Sapienza' voluntarily took part in the experiment. Sample size is similar to that of other studies employing affective priming paradigms together with physiological recordings [20,28] and measuring facial–thermal correlates of emotions [7,10–13]. All participants were healthy, naive to the purposes of the study, had normal or corrected-to-normal vision, and signed an informed consent form. Participants received compensation of €15 for their participation.

(b) Experimental stimuli

The affective visual stimuli (positive, neutral, and negative valence) used for the affective priming task were taken from the International Affective Picture System (IAPS) [43]. We selected 186 stimuli; half of them were employed in the subliminal block ($n = 93$, $n = 31$ for each valence) and the other half in the supraliminal one ($n = 93$, $n = 31$ for each valence). The affective stimuli's assignment to the subliminal and supraliminal block was counterbalanced among participants. Affective stimuli were selected according to the norms indicated in the IAPS technical report [43] (see electronic supplementary material, Materials and Methods, Stimulus Material and table S1); in particular, stimuli were selected in such a way that (i) positive, neutral, and negative stimuli significantly differed among each other in valence ratings (positive > neutral > negative) and (ii) positive and negative stimuli were equally arousing while neutral ones were significantly less arousing (neutral < positive = negative). The neutral face stimuli (size: 400×477 pixels) employed in the social categorization task were taken
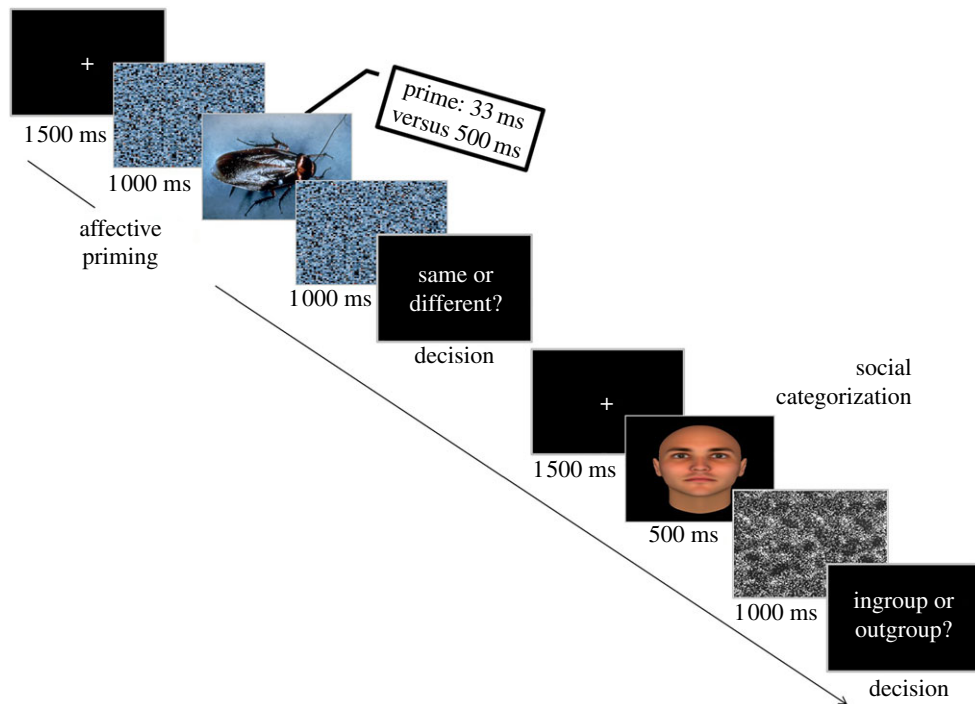
Figure 1. Timeline of the affective priming and social categorization task. Affective stimuli were taken from the IAPS [43]; neutral faces were taken from three validated sets [44–46].

from three validated face sets developed by a trustworthiness computer model and generated using FaceGen 3.1 software [44–46]. We selected 150 Caucasian male faces which were neutral on the trustworthiness dimension (corresponding to 0 s.d. in the trustworthiness level) because we were interested in faces that lacked any (positive or negative) affiliation or reputation cue.

In the social categorization task, participants had to classify the faces as belonging to Italian or Romanian nationality (it was explicitly stated that Romanians were different from Gypsies [47]). We chose the Romanian group as the outgroup because, like Italians, they exhibit a variety of different phenotypic features. Thus, participants could not rely on physical appearance in order to evaluate faces as belonging to a specific group. Before performing the study, we validated the selected stimuli in order to prevent perceptual facial features from creating bias in group categorization towards either Italians or Romanians. Specifically, we recruited 50 Italian participants (12 males; age range 20–36 years, $M = 25.28$, s.d. $= 4.67$) who completed an online survey having to do with faces. The aim of this survey was to determine which faces would be judged as either Italian or Romanian at chance level (and could thus be considered 'neutral' for the social categorization task). A two-sided binomial test (α-level $= 0.05$) revealed that 57 face stimuli were classified as Italian or Romanian significantly (i.e. $p < 0.05$) above the chance level. These stimuli were discarded. The remaining 93 face stimuli were used for the social categorization task.

(c) Procedure
(i) Questionnaire
Prior to the experimental session, participants completed the 'Lack of emotional awareness' subscale of the difficulties in emotion regulation strategies (DERS) [48,49], a self-report questionnaire that individuates specific facets of emotion regulation. We administered this specific subscale because we hypothesized that emotional awareness (and not general emotion regulation abilities) might play a key role in affective misattribution mechanisms. The 'Lack of emotional awareness' subscale consists of reverse coded items that measure the tendency to attend to

and acknowledge one's own emotions (e.g. *I am attentive to my feelings*, *I pay attention to how I feel*, and *I care about what I am feeling*). Scores of this subscale ranged from 3 to 12 ($M = 5.82$, s.d. $= 2.17$, Me $= 6$), with higher scores suggesting a lack of awareness or inattention to emotional responses.

(ii) Task and design
In the affective priming task, the prime's visual presentation could be subliminal (i.e. inaccessible to visual awareness) or supraliminal (i.e. accessible to visual awareness) depending on its duration (33 ms—as in [20]—for the subliminal block and 500 ms—as in [50]—for the supraliminal one). We employed the forward and backward masking technique [17] in which the prime is both preceded and followed by visual masks created by scrambling the prime itself (i.e. the masks are composed of randomly generated 35×35 squares and characterized by the same contrast and luminance values of the associated prime). In the subsequent social categorization task, participants were asked to categorize the neutral face as ingroup (i.e. Italian) or outgroup (i.e. Romanian).

Each trial had the following sequence of elements: (i) fixation cross (1 500 ms), (ii) scrambled mask (1 000 ms), (iii) prime (33 ms in the subliminal block; 500 ms in the supraliminal block), (iv) scrambled mask (1 000 ms), (v) same/different recognition task (*Same or different?*), (vi) fixation cross (1 500 ms), (vii) face stimulus (500 ms), (viii) visual noise (1 000 ms), and (ix) social categorization task (*Italian or Romanian?*) (figure 1). Each trial was concluded by a final intertrial interval (ITI) of 4 000 ms constituted by a fixation cross. All 186 affective prime stimuli were presented once; within each block, they were presented in a fully randomized order. All 93 neutral face stimuli were presented twice (once in the subliminal block and once in the supraliminal one).

Participants were then presented with a surprise memory recognition task consisting of the affective stimuli of the social categorization task (see electronic supplementary material, Results, Memory Recognition Task).

Thus, the entire experiment consisted of a subliminal block, a supraliminal block, and a final surprise recognition memory task.

The order of the tasks was fixed. In keeping with [20], the subliminal block preceded the supraliminal one to avoid participants being differentially biased by a search of hidden images in the subliminal condition. Participants were not informed about the masking procedure, and we only wanted them to be aware of the affective primes during the supraliminal block.

We created a cover story in order to justify the employment of the masks, telling participants that the experiment aimed to investigate the effects of cognitive load on face categorization (see electronic supplementary material, Procedure, Task and design, Cover story). Employing this task, we also ensured that participants were paying attention to the screen during the prime presentation. Participants were asked to look at the fixation cross once it appeared on the screen. The stimuli were presented using E-Prime Professional software v. 2.0.

(iii) Physiological recordings

While acclimatizing in the experimental room (for details, see electronic supplementary material, Procedure, Physiological recordings), participants completed the 'Lack of emotional awareness' subscale of the DERS (see the Questionnaire section). Face temperature (see electronic supplementary material, figure S1) during the affective priming and social categorization task was measured by means of fITI. This contact-free technique allows for skin temperature to be recorded at any distance (greater than 0.4 m) by tracking changes in temperature with high thermal (< 20 mK $= \leq 0.02\,^{\circ}$C) and temporal resolution. The digital infrared FLIR$^{©}$ camera SC3000 employed for the thermal recording was set at 10 Hz (10 frames/sec; temporal resolution: 100 ms) and situated at eye-level 1 m away from the participant in a controlled climate experimental room ($23 \pm 1\,^{\circ}$C). To avoid motion-related artefacts, participants' heads were kept still by a home-made headrest.

(iv) Data reduction and analyses

Thermal data were recorded through the ThermaCAM Researcher Professional 2.8 SR-3 software. To extract thermal information offline, we employed the motion-tracking software OTACS-V1 (Open Thermal Action Coding System) [51], which makes it possible to follow a specific region of interest (ROI) on the face reliably over time. To ensure a reliable positioning and sizing of the ROIs, we employed a square with the largest possible area that did not touch the eyelids (selection criteria as in [15]). We selected left and right periorbital regions as ROIs (see electronic supplementary material, figure S1), as temperature increase in these face regions represents one of the most reliable thermal prints of affective response [6,12,14]. The increase is caused by increased blood flow to the skin as a consequence of heart rate acceleration mediated by the SNS [13].

We filtered thermal data (low-pass filter at 0.01 Hz and high-pass filter at 4.5 Hz) through a home-made MATLAB script. The thermal data from all participants were segmented at 1 500 ms before prime onset until 2 500 ms after prime onset and baseline corrected ($-1 500$ to 0 ms) using the Brain Vision Analyzer v. 1.5 (Brain products GmbH, Munich, Germany) software. Since we did not expect any difference in the lateralization of the thermal signal, we pooled left and right periorbital temperature data into a single channel. We were interested in analysing the thermal signal prior to the social categorization response, so physiological and motion artefacts were identified and rejected using visual inspection in the epoch of interest ($-1 500 + 1 600$) which did not include the presentation of the face to be categorized. The thermal data were constructed by exporting artefact-free epochs (mean number of trials per subject $= 35.55$) averaged in 100 ms time intervals and belonging to six different experimental conditions from 31 participants (two participants had no artefact-free epochs).

3. Results

(a) Periorbital temperature as a predictor of social categorization behaviour

We explored whether periorbital face temperature recorded during subliminal or supraliminal processing of affective stimuli might be a predictor of social categorization behaviour. Single trials of temperature data were pooled into 500 ms time intervals and analysed after 100 ms from the onset of the primes. We were able to detect all early temperature modulations because the most rapid face temperature changes take place after 300 ms from the stimulus presentation [6]. Owing to the different lengths of the epochs, we modelled thermal data separately for subliminal (time windows: 100–600 and 600–1 100 ms) and supraliminal (time windows: 100–600, 600–1 100, and 1 100–1 600 ms) primes.

(b) Temperature subliminal model

Through the R package lme4 v. 1.1-5 [52], we performed a multilevel mixed log-linear regression analysis, a statistical method belonging to the family of linear mixed models (or mixed effects models, [53]; see also electronic supplementary material, Procedure, Single trial generalized linear mixed models). We used categorization behaviour (i.e. labelling the face as ingroup versus outgroup) as a dependent variable and valence (positive, neutral, and negative) as a categorical predictor. As continuous predictors, we used temperature in 100–600 ms and temperature in 600–1 100 ms time windows. The valence × temperature in 100–600 ms and valence × temperature in 600–1 100 ms interactions were also present in the model (see electronic supplementary material, Results, Models formulas, Temperature subliminal model). As suggested by guidelines [54,55], the higher-order significant interaction was also modelled as a random slope over participants, but this model did not explain an additional significant portion of variance with respect to the previous one ($\chi^2 = 0.26$, $p = 0.99$).

We found a main effect of valence ($\chi^2 = 7.66$, bootstrap-$p = 0.001$) qualified by a significant two-way interaction between valence and temperature in the 600–1 100 ms time window ($\chi^2 = 5.91$, bootstrap-$p = 0.02$, figure 2a). Post hoc analysis showed that higher periorbital temperatures led to significant bias in face categorization towards outgroup in the negative with respect to the positive subliminal condition ($b = -77.11$, s.e. $= 32.36$, $z = -2.38$, $p = 0.02$; figure 2a). The subliminal neutral condition did not significantly differ from the subliminal positive ($b = -29.94$, s.e. $= 33.99$, $z = -0.88$, $p = 0.38$; figure 2a) or negative conditions ($b = -47.16$, s.e. $= 33.76$, $z = -1.39$, $p = 0.16$; figure 2a) in biasing social categorization of faces. Thus, in line with the affective misattribution hypothesis [26], data show that the increase in periorbital temperature (which reflects an engagement of the SNS) in the emotional subliminal condition predicts a categorization behaviour that is congruent with the valence of the induction (i.e. outgroup after the negative induction and ingroup after the positive one). In particular, these results show that the affective misattribution of the emotional prime on the categorization decision is moderated by SNS activity (figure 2a).
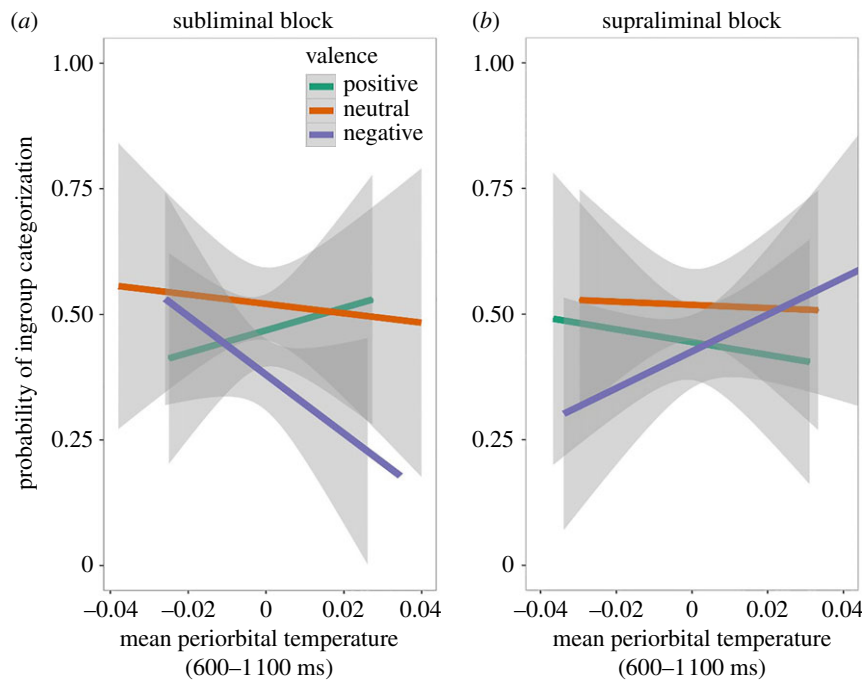
Figure 2. (a) (subliminal block, $n = 31$): predicted probability to categorize a face as ingroup when considering the two-way interaction between negative valence and mean periorbital temperature in the $600-1\,100$ ms time interval after the prime onset. (b) (supraliminal block, $n = 31$): predicted probability to categorize a face as ingroup when considering the two-way interaction between negative valence and mean periorbital temperature in the $600-1\,100$ ms time window after the prime onset. The shaded bands represent 95% CIs.

(c) Temperature supraliminal model

We performed a multilevel mixed log-linear regression analysis with categorization behaviour (i.e. labelling the face as ingroup versus outgroup) as a dependent variable, valence (positive, neutral, and negative) as a categorical predictor, and temperature in the $100-600$, $600-1\,100$, and $1\,100-1\,600$ ms time windows as continuous predictors. The interactions valence $\times$ temperature in the $100-600$ ms, $600-1\,100$ ms, and $1\,100-1\,600$ ms time windows were also present in the model (see electronic supplementary material, Results, Models formulas, Temperature supraliminal model). The higher-order significant interaction was additionally modelled as a random slope over participants, but this model did not explain any additional significant portion of variance with respect to the previous one ($\chi^2 = 0.61$, $p = 0.99$).

We found a significant two-way interaction between valence and temperature in the $600-1\,100$ ms time window ($\chi^2 = 6.99$, bootstrap-$p = 0.05$; figure 2b). *Post hoc* analysis showed that the periorbital temperature significantly predicted ingroup versus outgroup categorization after affective induction with negative respect to the positive ($b = 82.34$, s.e. $= 34.78$, $z = 2.37$, $p = 0.02$; figure 2b). The supraliminal neutral condition did not significantly differ from the supraliminal positive one in biasing social categorization of faces ($b = -55.33$, s.e. $= 37.34$, $z = -1.48$, $p = 0.13$; figure 2b), nor did it differ from the supraliminal negative condition ($b = 27.00$, s.e. $= 38.59$, $z = 0.70$, $p = 0.48$; figure 2b). Thus, data from the supraliminal block indicate that increase in periorbital temperature (which reflects an engagement of the SNS) after the presentation of emotional primes predicted a tendency of face categorization which was incongruent with the valence of the induction (i.e. ingroup after the negative and outgroup after the positive induction). This result pattern suggests that the regulation of one's own affective reactions when primes are supraliminally presented plays a major role in social categorization. We submit that SNS

activation triggered by a supraliminal affective induction might allow the conscious appraisal of the emotional stimuli and a subsequent regulatory process that leads to a categorization which is incongruent with the valence of the priming.

(d) Individual differences

To investigate the role of individual differences in trait-EA, we modelled the categorization response (dependent variable) using predictors resulting from the previous analyses as the significant variables (i.e. negative/positive valence and mean periorbital temperature in the $600-1\,100$ ms time window) in both blocks (subliminal and supraliminal) along with the scores of the 'Lack of emotional awareness' subscale of the DERS [48,49]. All the reciprocal interactions were also present in the model (see electronic supplementary material, Results, Models formulas, Individual differences, Main analysis). The within highest-order significant interaction was additionally modelled as random slopes over participants, but this model did not explain the additional significant portion of variance with respect to the previous one ($\chi^2 = 4.57$, $p = 0.99$).

We found a significant four-way interaction among valence, block, temperature, and trait-EA ($\chi^2 = 5.86$, bootstrap-$p = 0.001$, figure 3). To test the possibility that participants with low- versus high-EA differed regarding their basic SNS responses, we ran a model in which EA was the predictor and the temperature in the $600-1\,100$ ms time window was the dependent variable. We found that the temperature did not significantly change according to EA scores ($\chi^2 = 0.98$, $p = 0.32$). We conducted the analysis separately for low- ($n = 16$) and high-EA participants ($n = 15$) in order to enhance our test of the four-way interaction. All the reciprocal interactions were also present in the models (see electronic supplementary material, Results, Models formulas, Individual differences, Low- and high-EA
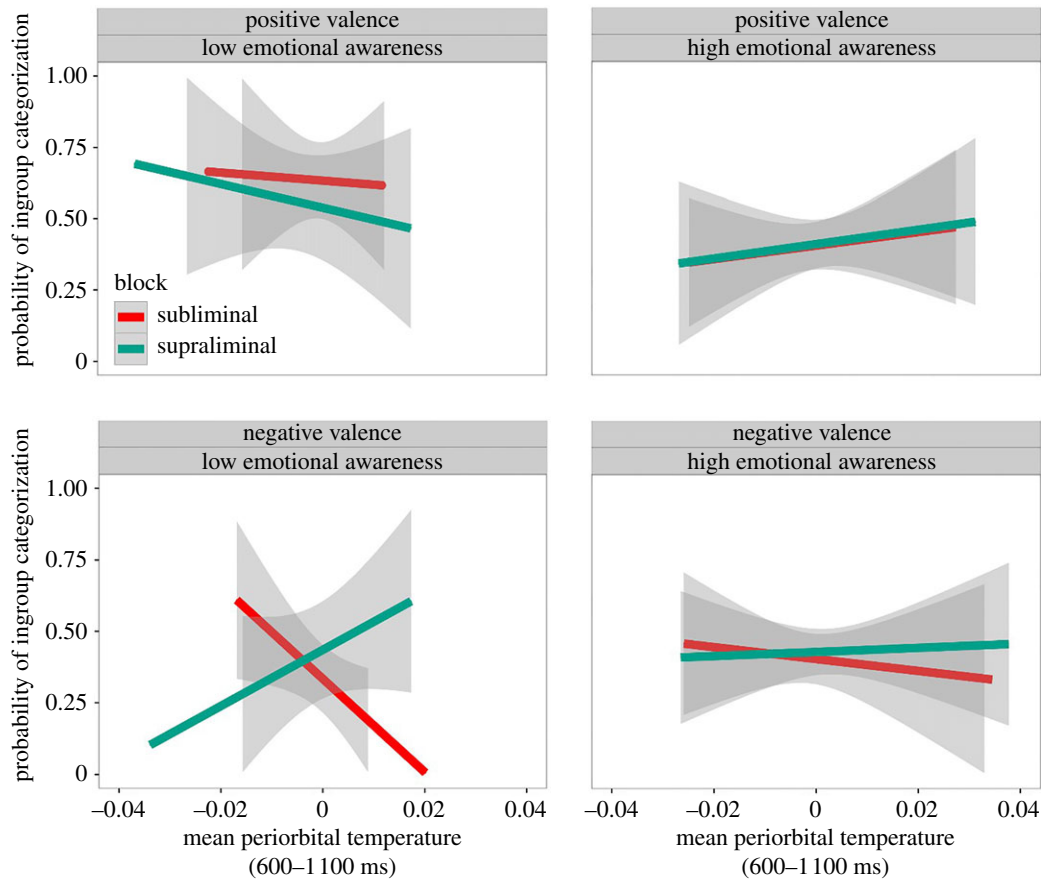
Figure 3. Predicted probability of categorizing a face as ingroup or outgroup when considering the four-way interaction ($n = 31$) among valence (negative, positive), trait-EA, block (subliminal versus supraliminal) and mean periorbital temperature in the second time window (600–1 100 ms after the prime onset). The shaded bands represent 95% CIs.

participants). The highest-order significant interaction was additionally modelled as random slopes over participants, but these models did not explain any additional significant portion of variance with respect to the previous ones both for low-EA ($\chi^2 = 5.30$, $p = 0.98$) and high-EA ($\chi^2 = 3.62$, $p = 0.99$) participants. We found a main effect of valence ($\chi^2 = 10.04$, bootstrap-$p = 0.001$) in the low-EA subsample which was qualified by a three-way interaction in trend towards significance between valence, temperature in the 600–1 100 ms time window, and Block ($\chi^2 = 3.36$, bootstrap-$p = 0.06$, figure 3a); the same main effect ($\chi^2 = 0.00$, bootstrap-$p = 0.99$) and three-way interaction ($\chi^2 = 1.03$, bootstrap-$p = 0.32$) were not significant for high-EA participants (figure 3b). *Post hoc* analysis showed that the opposite effect of subliminal versus supraliminal induction on categorization behaviour was present in low-EA participants after the negative prime ($b = 48.18$, s.e. $= 23.49$, $z = 2.05$, $p = 0.04$) but not the positive one ($b = -14.82$, s.e. $= 24.50$, $z = -0.60$, $p = 0.54$; figure 3a). In low-EA participants, in other words, valence-congruent effects in the subliminal condition and valence-incongruent effects in the supraliminal one are moderated by sympathetic activation, but only after the negative primes.

4. Discussion

We combined affective priming with high-sensitivity thermal imaging to investigate whether periorbital temperature recorded while processing subliminal versus supraliminal emotional visual primes would predict the social

categorization of emotionally neutral faces that lack any perceptual or reputational cues.

We found that the affective misattribution effect was moderated by the activation of the SNS. Previous fITI studies have tested the conscious processing of emotional stimuli and shown that a large variety of primary lower-level [6–8] and secondary higher-level [9–11] emotions elicit skin temperature responses. However, to the best of our knowledge, our study is the first to demonstrate that facial temperature changes are contingent upon subliminal processing of affective stimuli. Importantly, we found that while activation of the SNS after a negative subliminal induction leads to an outgroup categorization, activation following positive primes causes the opposite behavioural outcome.

These results can be interpreted according to various theoretical frameworks. The affective misattribution hypothesis [26] states that people tend to misattribute priming-induced affective valence to neutral-unrelated stimuli. In our study, positive and negative valence leads to misattribute a neutral face to an ingroup and outgroup category, respectively. This misattribution pattern is congruent with the so-called intergroup bias [56], according to which people tend to evaluate more positively those they perceive as belonging to their same social group than those perceived as from another group. According to the affect-as-information theory [25], positive affect signals that the environment is safe, and thus that no additional cognitive resources are required to deal with it. Conversely, negative affect acts as a cue for incoming threats and induces a more accurate and detailed way of thinking (but see [57,58] for a detailed discussion about the differential effects of emotions

on executive processing), often leading to conservative choices in taking risks [59,60]. In these terms, the negative valence-induced SNS activation in our subliminal condition might have triggered a more analytical way of thinking that leads to a conservative choice in terms of social categorization (i.e. labelling the face as outgroup), while the positive valence-induced SNS activation might have caused the opposite. This interpretation is in line with the results of previous studies, showing that participants with high ingroup over-exclusion effect [47,61] (i.e. the tendency to classify individuals as outgroup in order to protect one's own group integrity) were also the ones who based their group membership decisions on perceived competence (a trait that does not depend on individual control [56]). Moreover, Miller et al. [38] showed consistently across six experiments that cues signalling potential threat induce a bias towards outgroup categorization.

In addition, our results are in keeping with previous research reporting that (i) sympathetic and neural reactivity to threatening stimuli predicts decreased likability of neutral faces, even when not consciously perceived [41,42], (ii) people tend to prefer faces presented after a positive subliminal affective prime compared to a negative one [27], (iii) bodily states affect the expression of race-threat stereotypes [62], and (iv) the effect of negative or positive priming on the subsequent misattribution is mediated by amygdala or nucleus accumbens activation, respectively [63], suggesting that different neural activations contribute to opposite behavioural outcomes.

Interestingly, we found that supraliminal presentation of negative primes brought about exactly the opposite social categorization pattern, i.e. the negative valence-induced SNS activation predicted an ingroup categorization while the positive valence-induced sympathetic activation led to the opposite choice. Studies indicate that affective primes can lead to assimilation or contrast effects [64,65]. Assimilation effects refer to situations in which judgements about neutral stimuli are consistent with the valence of the prime. Contrast effects meanwhile refer to situations in which judgements about neutral stimuli are inconsistent with the valence of the primes. Our paradigm thus shows that affect-induced SNS activation leads to assimilation effects in the subliminal condition and contrast effects in the supraliminal one.

Importantly, we found that assimilation and contrast effects are moderated by SNS activation and, specifically, that both the subliminal assimilation and supraliminal contrast effects depend on the amplitude of signals coming from the SNS. The stronger these signals are, the more participants tend to misattribute the valence in the subliminal condition, or regulate the effect of the valence in the supraliminal one. We submit that when affect-induced SNS activation is caused by a supraliminal induction, participants consciously appraise the valence of the prime and engage in emotion regulation processes consequently. This engagement eventually results in a categorization decision that is incongruent with the valence of the prime. Considering that temperature and valence of the stimuli were randomly distributed across participants, we are quite confident that the significance of our interactions was not driven by sequence, learning, or fatigue effects.

Previous studies have shown that conscious awareness allows participants to process information more flexibly and to resist the congruent reaction induced by priming stimuli that results in overcompensation or contrast effects [64,65].

Interestingly, deliberate emotion regulation is associated with increased cardiovascular activity [66,67]. Moreover, misattribution bias is predicted by amygdala activity when people are unaware of the emotional induction and it instead reduced when participants consciously engaged in emotion regulation processes [42]. Our results are in line with the idea that there are two different emotional systems involved in subliminal and supraliminal processing: one that responds automatically (subliminal) and one that does so reflectively (supraliminal) [68].

A further important result of the present study comes from our exploration of the emotional awareness trait. We found that when the prime's valence is negative, individuals with low emotional awareness (i) have stronger SNS-induced assimilation effects in the subliminal condition and (ii) stronger SNS-induced contrast effects in the supraliminal condition. Not surprisingly, we found this modulation only after the negative induction, a pattern that could be explained in various ways. First, the 'lack of emotional awareness subscale' of the DERS focuses mainly on upsetting emotions. It is thus reasonable to believe that individual differences in this scale relate exclusively to the negative emotional condition of our paradigm. Furthermore, patients with impaired emotion regulation ability (e.g. anxiety disorders) have more active amygdala and insula regions during negative emotional processing. As these structures are linked to negative emotional responses, such patients seem to be more sensitive to negative information [69].

These results expand studies showing that individual differences in emotional awareness are reflected in one's physiological [68,70] and neural [71] activity during emotional processing, and that they can influence proneness to aggression [72]. It has been consistently shown that high emotional awareness is associated with greater engagement of the dorsal anterior cingulate cortex (ACC) (a brain region implicated in the allocation of attention to emotional information) during the processing of highly arousing emotional stimuli [71]. Also relevant is that emotional awareness plays a role in physiological reactivity dysfunction [73,74]: increased basal hypothalamic-pituitary-adrenal axis (HPA) activity [75] is found in alexithymic individuals who typically exhibit reduced interoceptive awareness and deficits in the perception of emotional signals. More generally, research in political psychology provides evidence for a relationship between emotional reactivity, emotion regulation, and prejudiced attitudes. Conservatives, for example, display greater skin conductance than liberals when facing threatening stimuli [76], suggesting that heightened physiological reactivity could play a role in conservatives' prejudiced attitudes. Moreover, engaging in emotion regulation (i.e. reappraisal) successfully attenuates the relationship between disgust sensitivity and support of conservative policies [77].

Comprehensively considered, our study provides novel insights into how a higher-order process like social categorization is significantly biased by subliminal versus supraliminal affective induction. Importantly, we show that the effect of affective priming on social categorization is moderated by the activation of the SNS system as indexed by periorbital temperature. Specifically, SNS activation caused either assimilation or contrast effects in the subliminal and supraliminal conditions, respectively. We also provide a new insight into how social categorization behaviour and autonomic reactivity are modulated by individual differences in emotional awareness. This

suggests that, depending on the awareness level (subliminal or supraliminal induction), SNS activation can oppositely influence human proclivity to automatically categorize others along the 'us' versus 'them' dichotomy.

References

1. Kringelbach M, Phillips H. 2014 *Emotion. Pleasure and pain in the brain*. London, UK: Oxford University Press.

2. Bechara A, Damasio H, Tranel D, Damasio AR. 1997 Deciding advantageously before knowing the advantageous strategy. *Science* **275**, 1293–1295. (doi:10.1126/science.275.5304.1293)

3. Kreibig SD. 2010 Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* **84**, 394–421. (doi:10.1016/j.biopsycho.2010.03.010)

4. Ioannou S, Gallese V, Merla A. 2014 Thermal infrared imaging in psychophysiology: potentialities and limits. *Psychophysiology* **51**, 951–963. (doi:10.1111/psyp.12243)

5. Cruz-Albarrán IA, Benítez-Rangel JP, Osornio-Ríos RA, Morales-Hernández LA. 2017 Human emotions detection based on a smart-thermal system of thermographic images. *Infrared Phys. Technol.* **81**, 250–261. (doi:10.1016/j.infrared.2017.01.002)

6. Levine JA, Pavlidis I, Cooper M. 2001 The face of fear. *Lancet* **357**, 1757. (doi:10.1016/S0140-6736(00)04936-9)

7. Nakanishi R, Imai-Matsumura K. 2008 Facial skin temperature decreases in infants with joyful expression. *Infant Behav. Dev.* **31**, 137–144. (doi:10.1016/j.infbeh.2007.09.001)

8. Nakayama K, Goto S, Kuraoka K, Nakamura K. 2005 Decrease in nasal temperature of rhesus monkeys (*Macaca mulatta*) in negative emotional state. *Physiol. Behav.* **84**, 783–790. (doi:10.1016/j.physbeh.2005.03.009)

9. Salazar-López E *et al.* 2015 The mental and subjective skin: emotion, empathy, feelings and thermography. *Conscious. Cogn.* **34**, 149–162. (doi:10.1016/j.concog.2015.04.003)

10. Paolini D, Alparone FR, Cardone D, van Beest I, Merla A. 2016 'The face of ostracism': the impact of the social categorization on the thermal facial responses of the target and the observer. *Acta Psychol.* **163**, 65–73. (doi:10.1016/j.actpsy.2015.11.001)

11. Ioannou S *et al.* 2013 The autonomic signature of guilt in children: a thermal infrared imaging study. *PLoS ONE* **8**, e79440. (doi:10.1371/journal.pone.0079440)

12. Hahn AC, Whitehead RD, Albrecht M, Lefevre CE, Perrett DI. 2012 Hot or not? Thermal reactions to social contact. *Biol. Lett.* **8**, 864–867. (doi:10.1098/rsbl.2012.0338)

13. Ioannou S, Morris P, Mercer H, Baker M, Gallese V, Reddy V. 2014 Proximity and gaze influences facial temperature: a thermal infrared imaging study. *Front. Psychol.* **5**, 1–12. (doi:10.3389/fpsyg.2014.00845)

14. Pavlidis I, Eberhardt NL, Levine JA. 2002 Seeing through the face of deception. *Nature* **415**, 35. (doi:10.1038/415035a)

15. Panasiti MS, Cardone D, Pavone EF, Mancini A, Aglioti SM. 2016 Thermal signatures of voluntary deception in ecological conditions. *Sci. Rep.* **6**, 35174. (doi:10.1038/srep35174)

16. Tamietto M, de Gelder B. 2010 Neural bases of the non-conscious perception of emotional signals. *Nat. Rev. Neurosci.* **11**, 697–709. (doi:10.1038/nrn2889)

17. Kouider S, Dehaene S. 2007 Levels of processing during non-conscious perception: a critical review of visual masking. *Phil. Trans. R. Soc. B* **362**, 857–875. (doi:10.1098/rstb.2007.2093)

18. Tamietto M, Castelli L, Vighetti S, Perozzo P, Geminiani G, Weiskrantz L, de Gelder B. 2009 Unseen facial and bodily expressions trigger fast emotional reactions. *Proc. Natl Acad. Sci. USA* **106**, 17 661–17 666. (doi:10.1073/pnas.0908994106)

19. Dimberg U, Thunberg M, Elmehed K. 2000 Unconscious facial reactions to emotional facial expressions. *Psychol. Sci.* **11**, 86–89. (doi:10.1111/1467-9280.00221)

20. Gläscher J, Adolphs R. 2003 Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *J. Neurosci.* **23**, 10 274–10 282.

21. Allen M, Frank D, Samuel Schwarzkopf D, Fardo F, Winston JS, Hauser TU, Rees G. 2016 Unexpected arousal modulates the influence of sensory noise on confidence. *eLife* **5**, 1–17. (doi:10.7554/eLife.18103)

22. Winkielman P, Berridge KC, Wilbarger JL. 2005 Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Pers. Soc. Psychol. Bull.* **31**, 121–135. (doi:10.1177/0146167204271309)

23. Era V, Candidi M, Aglioti SM. 2015 Subliminal presentation of emotionally negative versus positive primes increases the perceived beauty of target stimuli. *Exp. Brain Res.* **233**, 3271–3281. (doi:10.1007/s00221-015-4395-5)

24. Panasiti MS, Puzzo I, Chakrabarti B. 2016 Autistic traits moderate the impact of reward learning on social behaviour. *Autism Res.* **9**, 471–479. (doi:10.1002/aur.1523)

25. Schwarz N, Clore GL. 1983 Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol.* **45**, 513–523. (doi:10.1037/0022-3514.45.3.513)

26. Payne BK, Cheng CM, Govorun O, Stewart BD. 2005 An inkblot for attitudes: affect misattribution as implicit measurement. *J. Pers. Soc. Psychol.* **89**, 277–293. (doi:10.1037/0022-3514.89.3.277)

27. Anderson E, Siegel E, White D, Barrett LF. 2012 Out of sight but not out of mind: unseen affective faces influence evaluations and social impressions. *Emotion* **12**, 1210–1221. (doi:10.1037/a0027514)

28. Chiesa PA, Liuzza MT, Acciarino A, Aglioti SM. 2015 Subliminal perception of others' physical pain and pleasure. *Exp. Brain Res.* **233**, 2373–2382. (doi:10.1007/s00221-015-4307-8)

29. Chiesa PA, Liuzza MT, Macaluso E, Aglioti SM. 2017 Brain activity induced by implicit processing of others' pain and pleasure. *Hum. Brain Mapp.* **38**.

30. Jacob H, Bruck C, Domin M, Lotze M, Wildgruber D. 2014 I can't keep your face and voice out of my head: neural correlates of an attentional bias toward nonverbal emotional cues. *Cereb. Cortex* **24**, 1460–1473. (doi:10.1093/cercor/bhs417)

31. Li W, Moallem I, Paller KA, Gottfried JA. 2007 Subliminal smells can guide social preferences. *Psychol. Sci.* **18**, 1044–1049. (doi:10.1111/j.1467-9280.2007.02023.x)

32. Mancini A, Betti V, Panasiti MS, Pavone EF, Aglioti SM. 2011 Suffering makes you egoist: acute pain increases acceptance rates and reduces fairness during a bilateral ultimatum game. *PLoS ONE* **6**, e26008. (doi:10.1371/journal.pone.0026008)

33. Mancini A, Betti V, Panasiti MS, Pavone EF, Aglioti SM. 2014 Perceiving monetary loss as due to inequity reduces behavioral and cortical responses to pain. *Eur. J. Neurosci.* **40**, 2378–2388. (doi:10.1111/ejn.12582)

34. Gyurak A, Gross JJ, Etkin A. 2011 Explicit and implicit emotion regulation: a dual-process framework. *Cogn. Emotion* **25**, 400–412. (doi:10.1080/02699931.2010.544160)

8

rspb.royalsocietypublishing.org   *Proc. R. Soc. B* **284**: 20170908

35. Tarrant M, Dazeley S, Cottom T. 2009 Social categorization and empathy for outgroup members. *Br. J. Soc. Psychol.* **48**, 427–446. (doi:10.1348/014466608X373589)

36. Brown LM, Bradley MM, Lang PJ. 2006 Affective reactions to pictures of ingroup and outgroup members. *Biol. Psychol.* **71**, 303–311. (doi:10.1016/j.biopsycho.2005.06.003)

37. Cikara M, Fiske ST. 2012 Stereotypes and Schadenfreude: affective and physiological markers of pleasure at outgroup misfortunes. *Soc. Psychol. Pers. Sci.* **3**, 63–71. (doi:10.1177/1948550611409245)

38. Miller SL, Maner JK, Becker DV. 2010 Self-protective biases in group categorization: threat cues shape the psychological boundary between 'us' and 'them'. *J. Pers. Soc. Psychol.* **99**, 62–77. (doi:10.1037/a0018086)

39. Krosch AR, Amodio DM. 2014 Economic scarcity alters the perception of race. *Proc. Natl Acad. Sci. USA* **111**, 9079–9084. (doi:10.1073/pnas.1404448111)

40. Heerdink MW, van Kleef GA, Homan AC, Fischer AH. 2014 Emotional expressions as social signals of rejection and acceptance: evidence from the affect misattribution paradigm. *J. Exp. Soc. Psychol.* **56**, 60–68. (doi:10.1016/j.jesp.2014.09.004)

41. Lapate RC, Rokers B, Li T, Davidson RJ. 2014 Nonconscious emotional activation colors first impressions: a regulatory role for conscious awareness. *Psychol. Sci.* **25**, 349–357. (doi:10.1177/0956797613503175)

42. Lapate RC, Rokers B, Tromp DPM, Orfali NS, Oler JA, Doran ST, Adluru N, Alexander AL, Davidson RJ. 2016 Awareness of emotional stimuli determines the behavioral consequences of amygdala activation and amygdala-prefrontal connectivity. *Sci. Rep.* **6**, 25826. (doi:10.1038/srep25826)

43. Lang PJ, Bradley MM, Cuthbert BN. 2008 *International affective picture system (IAPS): affective ratings of pictures and instruction manual*. Technical Report A-8. Gaineville, FL: University of Florida.

44. Oosterhof NN, Todorov A. 2008 The functional basis of face evaluation. *Proc. Natl Acad. Sci. USA* **105**, 11 087–11 092. (doi:10.1073/pnas.0805664105)

45. Todorov A, Dotsch R, Porter JM, Oosterhof NN, Falvello VB. 2013 Validation of data-driven computational models of social perception of faces. *Emotion* **13**, 724–738. (doi:10.1037/a0032335)

46. Todorov A, Oosterhof NN. 2011 Modeling social perception of faces. *Signal Process. Mag. IEEE* **28**, 117–122. (doi:10.1109/MSP.2010.940006)

47. Ponsi G, Panasiti MS, Scandola M, Aglioti SM. 2016 Influence of warmth and competence on the promotion of safe in-group selection: SCM and social categorization of faces. *Q. J. Exp. Psychol.* **69**, 1464–1479. (doi:10.1080/17470218.2015.1084339)

48. Gratz KL, Roemer L. 2004 Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *J. Psychopathol. Behav. Assess.* **26**, 41–54. (doi:10.1023/B:JOBA.0000007455.08539.94)

49. Sighinolfi C, Norcini A, Rocco L. 2010 Difficulties in emotion regulation scale (DERS): the Italian translation and adaptation. *Psicoterapia Cognitiva Comportamentale* **16**, 141–170.

50. Williams LM, Liddell BJ, Kemp AH, Bryant RA, Meares RA, Peduto AS, Gordon E. 2006 Amygdala–prefrontal dissociation of subliminal and supraliminal fear. *Hum. Brain Mapp.* **27**, 652–661. (doi:10.1002/hbm.20208)

51. Zhou Y, Tsiamyrtzis P, Lindner P, Timofeyev I, Pavlidis I. 2013 Spatiotemporal smoothing as a basis for facial tissue tracking in thermal imaging. *IEEE Trans. Biomed. Eng.* **60**, 1280–1289. (doi:10.1109/TBME.2012.2232927)

52. Bates D, Maechler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48. (doi:10.18637/jss.v067.i01)

53. Pinheiro JC, Bates DM. 2000 *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

54. Jaeger TF. 2008 Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* **59**, 434–446. (doi:10.1016/j.jml.2007.11.007. Categorical)

55. Barr DJ. 2013 Random effects structure for testing interactions in linear mixed-effects models. *Front. Psychol.* **4**, 328. (doi:10.3389/fpsyg.2013.00328)

56. Cuddy AJC, Fiske ST, Glick P. 2008 Warmth and competence as universal dimensions of social perception: the stereotype content model and the BIAS map. In *Advances in experimental social psychology* (ed. MP Zanna), pp. 61–150. San Diego, CA: Academic Press.

57. Pessoa L. 2009 How do emotion and motivation direct executive control? *Trends Cogn. Sci.* **13**, 160–166. (doi:10.1016/j.tics.2009.01.006)

58. Mobbs D, Hagan CC, Dalgleish T, Silston B, Prévost C. 2015 The ecology of human fear: survival optimization and the nervous system. *Front. Neurosci.* **9**, 1–22. (doi:10.3389/fnins.2015.00055)

59. Yuen KSL, Lee TMC. 2003 Could mood state affect risk-taking decisions? *J. Affect. Disord.* **75**, 11–18. (doi:10.1016/S0165-0327(02)00022-8)

60. Schwager S, Rothermund K. 2013 Motivation and affective processing biases in risky decision making: a counter-regulation account. *J. Econ. Psychol.* **38**, 111–126. (doi:10.1016/j.joep.2012.08.005)

61. Leyens J-P, Yzerbyt VY. 1992 The ingroup overexclusion effect: impact of valence and confirmation on stereotypical information search. *Eur. J. Soc. Psychol.* **22**, 549–569. (doi:10.1002/ejsp.2420220604)

62. Azevedo RT, Garfinkel SN, Critchley HD, Tsakiris M. 2017 Cardiac afferent activity modulates the expression of racial stereotypes. *Nat. Commun.* **8**, 13854. (doi:10.1038/ncomms13854)

63. Suslow T, Kugel H, Ohrmann P, Stuhrmann A, Grotegerd D, Redlich R, Bauer J, Dannlowski U. 2013 Neural correlates of affective priming effects based on masked facial emotion: an fMRI study. *Psychiatry Res. Neuroimaging* **211**, 239–245. (doi:10.1016/j.pscychresns.2012.09.008)

64. Glaser J, Banaji MR. 1999 When fair is foul and foul is fair: reverse priming in automatic evaluation. *J. Pers. Soc. Psychol.* **77**, 669–687. (doi:10.1037/0022-3514.77.4.669)

65. Lombardi WJ, Higgins ET, Bargh JA. 1987 The role of consciousness in priming effects on categorization: assimilation versus contrast as a function of awareness of the priming task. *Pers. Soc. Psychol. Bull.* **13**, 411–429. (doi:10.1177/0146167287133009)

66. Butler EA, Egloff B, Wilhelm FH, Smith NC, Erickson EA, Gross JJ. 2003 The social consequences of expressive suppression. *Emotion* **3**, 48–67. (doi:10.1037/1528-3542.3.1.48)

67. Richards JM, Gross JJ. 1999 Composure at any cost? The cognitive consequences of emotion suppression. *Pers. Soc. Psychol. Bull.* **25**, 1033–1044. (doi:10.1177/01461672992511010)

68. Evers C, Hopp H, Gross JJ, Fischer AH, Manstead ASR, Mauss IB. 2014 Emotion response coherence: a dual-process perspective. *Biol. Psychol.* **98**, 43–49. (doi:10.1016/j.biopsycho.2013.11.003)

69. Etkin A, Wager TD. 2007 Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry* **164**, 1476–1488. (doi:10.1176/appi.ajp.2007.07030504)

70. Berna G, Ott L, Nandrino J-L. 2014 Effects of emotion regulation difficulties on the tonic and phasic cardiac autonomic response. *PLoS ONE* **9**, e102971. (doi:10.1371/journal.pone.0102971)

71. McRae K, Reiman EM, Fort CL, Chen K, Lane RD. 2008 Association between trait emotional awareness and dorsal anterior cingulate activity during emotion is arousal-dependent. *NeuroImage* **41**, 648–655. (doi:10.1016/j.neuroimage.2008.02.030)

72. Berkowitz L, Lepinski J, Angulo E. 1969 Awareness of own anger level and subsequent aggression. *J. Pers. Soc. Psychol.* **11**, 293–300. (doi:10.1037/h0027037)

73. Kanbara K, Fukunaga M. 2016 Links among emotional awareness, somatic awareness and autonomic homeostatic processing. *BioPsychoSoc. Med.* **10**, 16. (doi:10.1186/s13030-016-0059-3)

74. Lane RD, Schwartz GE. 1987 Levels of emotional awareness: a cognitive-developmental theory and its application to psychopathology. *Am. J. Psychiatry* **144**, 133–143. (doi:10.1176/ajp.144.2.133)

75. de Timary P, Roy E, Luminet O, Fillée C, Mikolajczak M. 2008 Relationship between alexithymia, alexithymia factors and salivary cortisol in men exposed to a social stress test. *Psychoneuroendocrinology* **33**, 1160–1164. (doi:10.1016/j.psyneuen.2008.06.005)

76. Oxley DR, Smith KB, Alford JR, Hibbing MV, Miller JL, Scalora M, Hatemi PK, Hibbing JR. 2008 Political attitudes vary with physiological traits. *Science* **321**, 1667–1670. (doi:10.1126/science.1157627)

77. Lee JJ, Sohn Y, Fowler JH. 2013 Emotion regulation as the foundation of political attitudes: does reappraisal decrease support for conservative policies? *PLoS ONE* **8**, e83143. (doi:10.1371/journal.pone.0083143)