INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Lego L, Cornec C, Corvin S, Massenet M, Papet L, Patural H, Reby D, Jouen F, Mathevon N. 2025 Nonlinear acoustic phenomena tune the adults' facial thermal response to baby cries with the cry amplitude envelope. *J. R. Soc. Interface* **22**: 20250150.
https://doi.org/10.1098/rsif.2025.0150

Received: 20 February 2025
Accepted: 14 August 2025

Subject Category:
Life Sciences–Physics interface

Subject Areas:
medical physics, biomedical engineering

Keywords:
baby cries, nonlinear phenomena, autonomic nervous system, thermal infrared imaging

Author for correspondence:
Nicolas Mathevon
e-mail: mathevon@univ-st-etienne.fr

Nonlinear acoustic phenomena tune the adults' facial thermal response to baby cries with the cry amplitude envelope

Lény Lego[1,3], Clément Cornec[1], Siloé Corvin[1], Mathilde Massenet[1,4,5], Leo Papet[1], Hugues Patural[2], David Reby[1,6], François Jouen[3,7,†] and Nicolas Mathevon[1,3,6,†]

[1]ENES Bioacoustics Research Lab, CRNL, CNRS, Inserm, and [2]SAINBIOSE lab, Inserm, Neonatal and Pediatric Intensive Care Unit, University of Saint-Etienne, Saint-Etienne, France
[3]Ecole Pratique des Hautes Etudes – PSL, CHArt Lab, University Paris-Sciences-Lettres, Paris, France
[4]Department of Communication, UCLA, Los Angeles, CA, USA
[5]Division of Cognitive Science, Lund University, Lund, Sweden
[6]Institut universitaire de France, Paris, France
[7]R2P2 Unit, Raymond Poincaré Hospital, Garches, France

CC, 0000-0002-4110-328X; SC, 0000-0002-9281-6087; MM, 0000-0002-0085-1871; LP, 0000-0001-8916-4666; NM, 0000-0003-0219-6601

Getting caregivers to respond to their pain cries is vital for the human baby. Previous studies have shown that certain features of baby cries—the nonlinear phenomena (NLP)—enable caregivers to assess the pain felt by the baby. However, the extent to which these NLP mobilize the autonomic nervous system of an adult listener remains unexplored. Here, we show that variations in a listener's facial temperature, a marker of the autonomic emotional response, reflect the pain expressed by a baby's cry. Specifically, by conducting listening experiments with cries expressing mild discomfort or acute pain, we demonstrate that NLP modulate the facial thermal response in adult listeners, irrespective of sex and of cry pitch variation. The temporal dynamics of the thermal response is more closely synchronized with the amplitude envelope of the acoustic signal when listening to a cry containing a large level of NLP than when NLP are less prominent. The pain encoded in a baby's cry thus generates a synchronized emotional response in both adult men and women, emphasizing that our ability to decode the information carried by babies' cries integrates an immediate activation of the autonomic nervous system before engaging higher-order cognitive processes.

1. Introduction

The cry is a sound signal enabling the baby to solicit help [1,2]. This signal carries both static information (the baby's individualized vocal signature [3]) and dynamic information (the baby's emotional state [4]). Its main function is to alert the baby's caregivers, encouraging them to provide care [5,6]. As a graded signal, the cry gives the baby the possibility to express a variable amount of discomfort or pain by varying a set of acoustic features [7,8]. When a baby cries because of an uncomfortable situation but is not really in pain, the vocal folds vibrate mostly periodically, producing a sound consisting mainly of a fundamental frequency and its harmonic series. When a baby is in real distress, they forcefully contract their rib cage, resulting in higher-pressure airflow that produces faster and chaotic vibrations of the vocal folds. This results in a more variable pitch (fundamental frequency, f_0) and disharmonious sounds, known as nonlinear phenomena (NLP) [9–

[†]These authors jointly supervised this work.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.7720934.

THE ROYAL SOCIETY
PUBLISHING

11]. As detailed in Corvin *et al*. [8], cries can express various types of NLP: *deterministic chaos* (aperiodic vibration of the vocal folds; hereafter chaos), low *amplitude modulation* (modulation of the fundamental frequency by a much lower frequency, possibly produced by another sound source, such as the ventricular folds), *subharmonics* (resulting from period doubling or tripling in the vocal fold vibration), *vibrato* (resulting from frequency and amplitude modulations of the fundamental frequency), *frequency jumps* (sudden jump in the fundamental frequency) and *vocal fry* (slow and irregular glottal pulses; for further explanations on nonlinear acoustic phenomena, see [12–14]). Previous psychoacoustic experiments have shown that these pain- or distress-induced acoustic traits (pitch variation, NLP) affect adult listeners' perception of the pain expressed by cries [8–11,13]. Notably, auditory tests using synthetic cries where the presence of each type of NLP in the cry was controlled, demonstrated that almost each NLP significantly contributes to an adult listener's perception of pain (namely: chaos, amplitude modulation, subharmonics, vibrato in both parents and non-parents; frequency jumps only in parents; see figure 1 for a representation of these NLP) [8].

The response to babies' cries by caregivers involves both low- and high-level neurophysiological processes, from the detection of stimuli by the peripheral auditory system to the decoding of the information by cortical areas [15]. In order to assess the effect of NLP on the perception of a baby cry, it would be informative to measure physiological processes directly related to the spontaneous response of the autonomic nervous system using methods such as the electrodermal response or heart rate monitoring [16,17]. In the present study, we chose to test a promising non-invasive method: facial thermography, i.e. the measurement of variations in the surface temperature of the face. Thermography is based on a simple principle: the surface temperature of an individual varies as a function of blood flow through the skin, and these variations can be captured remotely by filming with an infra-red camera [18,19]. Thermography already has a wide range of applications in the field of human health, such as monitoring revascularization processes after traumatic accidents [20] and the monitoring of emotion [21]. As the autonomic nervous system is primarily responsible for vasodilation and vasoconstriction of vessels in the skin's dermis, surface temperature is an indicator of its activity [22–29]. The major advantage of thermography lies in its non-invasive nature, since it does not require any sensor to be placed directly on the skin and only requires filming the individual from a fixed distance. The face is usually the target area, since this is where temperature variations linked to emotional expression are the most significant [30]. As facial thermal imaging is performed at a distance, it can be easily deployed in adults and children, as well as in non-human animals [31,32].

Here, we assess whether changes in facial temperature measured by thermography reveal variation in the emotional state of human adult listeners in responses to baby cries. Specifically, we investigate the potential relationship between temporal variations in listeners' facial temperature and the presence of NLP in cries. As summarized in figure 2, we first filmed with a thermal camera the faces of 41 adult men and women listening to baby cries carrying a variable level of NLP. We asked these participants to rate whether the cries were recorded in a situation of simple discomfort, or rather strong pain. We then tested the hypothesis that NLP amplify temporal synchronization between the cry amplitude envelope and the variations in the listener's facial temperature, controlling for the cry pitch variation. As we showed in a previous study that adults of both sexes identify pain in a baby's cry with the same acuity [33], we further tested the hypothesis that the effect of NLP on the thermographic response is the same regardless of the participants' sex. To achieve these objectives, we used a dynamic time warping (DTW) approach to measure the distance between the cry amplitude envelope and the thermal signal.

2. Methods

2.1. Cry stimuli

We selected 23 baby cry recording sequences from the ENES Bioacoustics Research Laboratory recording database (duration of each cry sequence = 6.37 ± 1.27 s). This database includes discomfort cries recorded during a bathing session, and pain cries recorded during a vaccination session at the paediatrician's office. The 23 cry sequences were recorded from 16 different babies (7 girls, 9 boys, aged 60 ± 3 days; see [4] for details). We selected these cries by ensuring the quality of the signals (negligible background noise, minimum sequence duration of 6 s) and because they effectively represent the diversity of NLP that may be present in a cry [8].

In each of the 23 cry sequences selected, we quantified the presence of NLP that have been demonstrated to contribute to non-parent adult listeners' perception of pain [8], namely chaos, amplitude modulation, subharmonics and vibrato, by visually inspecting and annotating spectrograms provided by Praat software (frequency range 0.2–10 kHz and window length 0.05–0.08 s [34]). Chaos is visually identified as broadband noisy sections in the cry's spectrogram, subharmonics as spectral bands often occurring at $f_o/2$ or $f_o/3$, amplitude modulation as spectral bands not harmonically related to fo occurring at $m^*f_o \pm n^*j_o$, and vibrato as rapid modulations in both frequency and amplitude contours (see figure 1 for details).

We measured the absolute duration of each instance of NLP (in seconds), as well as their relative duration to the total duration of the cry (expressed as a percentage, % NLP). This level of NLP ranges from 0% to 76.37%, representing the percentage of time during which at least one phenomenon is present in the cry. When we detected different types of NLP occurring simultaneously, we did not sum the duration of each NLP to get a value for the total NLP time before calculating the percentage of the cry with NLP. It was thus not possible to potentially have more than 100% of NLP in a cry. Besides quantifying NLP, we also measured the cry mean pitch and the pitch variation expressed as the pitch range (= F_0 range), obtained by dividing the maximum by the minimum pitch [35,36]. We used the Soundgen R package to extract pitch contour and measure pitch values [37]. As a precaution, this automatic pitch tracking was manually validated for all stimuli. Table 1 summarizes the

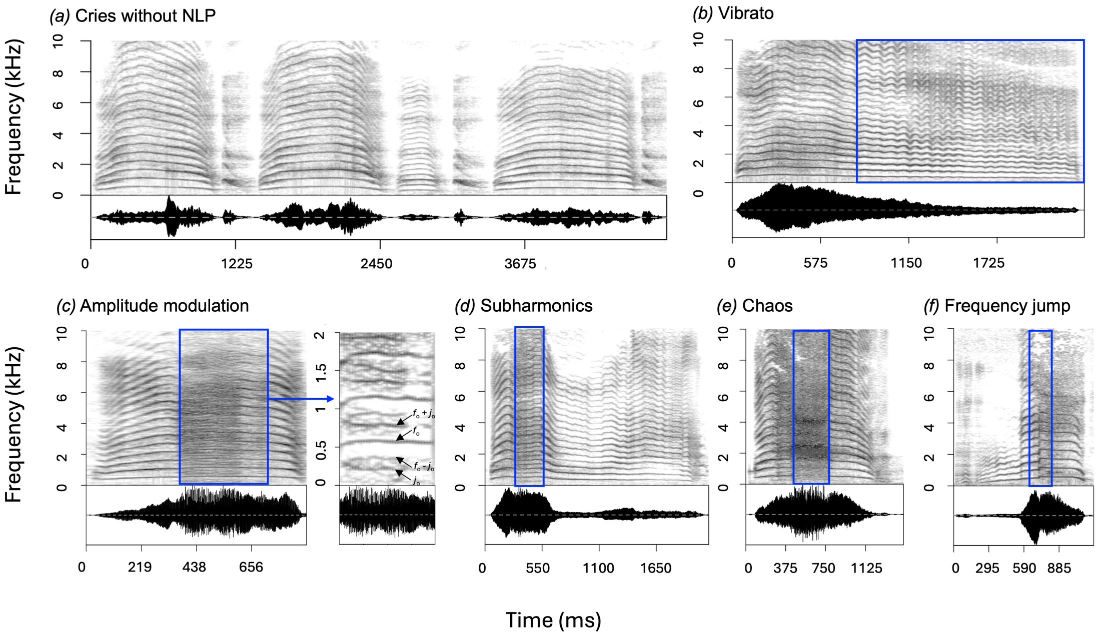

Figure 1. Spectrograms illustrating nonlinear acoustic phenomena (NLP) in human baby cries. Segments of cries containing NLP are outlined in blue. The NLP presented here are those known to contribute to an adult listener's perception of pain (see main text for explanations). Vibrato (*b*) is characterized by rapid frequency fluctuations in the f_0 contour. Amplitude modulation (*c*) manifests as sidebands occurring at $m*f_0 \pm n*j_0$, where j_0 is a lower frequency not harmonically related to f_0, and *m* and *n* are integers. In contrast, subharmonics (*d*) appear as spectral bands harmonically related to f_0, typically at $f_0/2$. Finally, chaos (*e*) corresponds to the broadband, noisy section of the cry, and a frequency jump (*f*) corresponds to a discontinuity in the f_0 contour, such as an upward shift. Spectrograms were produced with the function spectrogram() from the *Soundgen* R package [15].

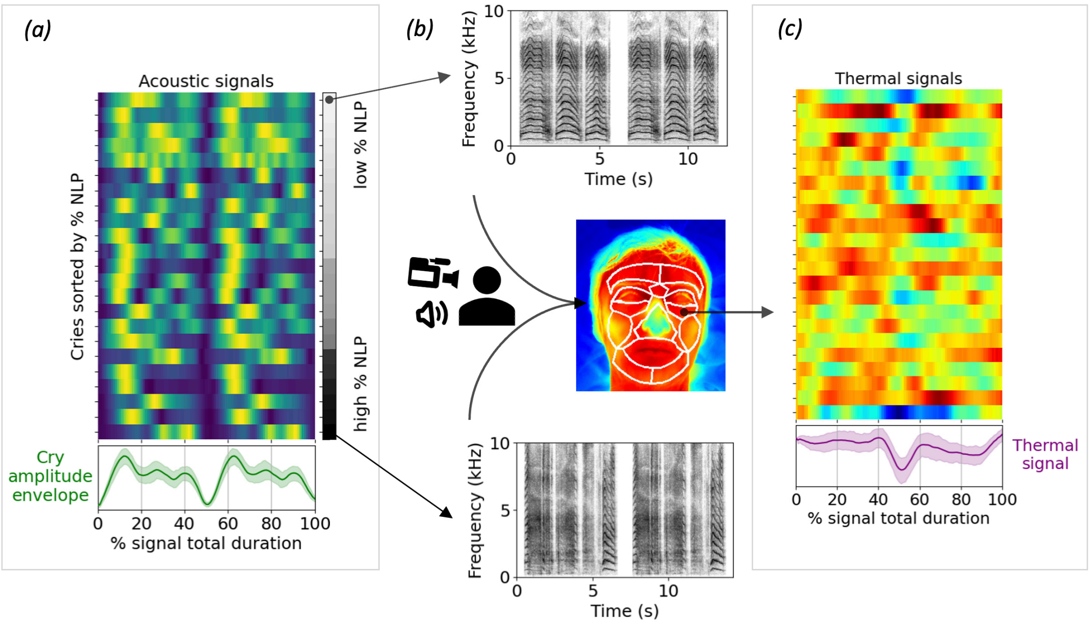

Figure 2. Experimental measurement of facial temperature variations in adults listening to baby cries. (*a*) Each coloured line of the panel represents a cry stimulus consisting of two successive repetitions of the same cry sequence. The colour variations indicate the intensity modulation (envelope) of the cries over time (ranging from blue = silence, to deep yellow = maximum intensity). We used 23 cry stimuli, ordered from top to bottom from the cries with the least NLP (low % NLP) to the cries with the most (high % NLP; see table 1 for details). The lower graph represents the cries' average envelope. Cry duration varied between 3 and 7 s but has been temporarily shaped to align all acoustic stimuli with the corresponding thermal signals (see §2). (*b*) The cry stimuli, illustrated by two examples with different levels of NLP, are emitted by a loudspeaker. The listener is filmed with a thermal camera to measure temperature variations in different areas of the face. (*c*) Each coloured line represents the temporal variation of the thermal signal recorded at a specific area of the face (here, right preorbital face area). The colour scale represents the normalized temperature value (ranging from blue = lowest temperature, to intense red = highest temperature). The lower graph shows the average thermal signal.

characteristics of each cry sequence used to build the stimuli in the experiment. Each cry stimulus consisted of two repetitions of the same cry sequence, separated by 1 s of silence.

Table 1. Quantification of nonlinear acoustic phenomena (NLP) and pitch variation in the cries used as stimuli.

cry ID	NLP duration (s)				total cry duration (s)	mean pitch (Hz)	min pitch (Hz)	max pitch (Hz)	F_0 range	NLP (%)
	vibrato	chaos	subharmonic	amplitude modulation						
1					4.963	470	266	559	2.1	0.00
2	0.530		0.149		7.864	421	171	514	3.0	10.06
3		0.396			5.120	470	357	543	1.5	6.49
4	0.586				5.068	427	368	506	1.4	9.61
5				1.410	7.483	485	354	574	1.6	22.04
6	1.016		0.185	0.061	8.680	464	257	576	2.2	16.83
7			0.046	0.301	8.789	426	297	506	1.7	4.56
8	1.208				7.320	496	280	611	2.2	19.34
9				0.081	7.102	462	303	540	1.8	1.35
10	0.622			0.342	7.211	530	339	873	2.6	18.49
11	1.359		0.209	0.132	4.989	455	203	611	3.0	28.89
12	3.812				8.027	399	239	503	2.1	55.25
13	1.712				5.068	371	216	480	2.2	27.83
14		0.687	0.351		6.034	479	334	543	1.6	14.26
15	2.414	0.703	0.926	1.788	6.034	273	203	402	2.0	75.48
16		2.644	0.060	1.310	6.034	570	478	666	1.4	37.99
17		3.662	0.068	0.283	4.780	472	362	567	1.6	65.07
18	0.160	2.598	0.051	1.683	7.810	924	803	1055	1.3	43.71
19	2.078	2.946	0.317		5.303	456	256	590	2.3	76.37
20	3.249		0.351	3.773	6.034	590	330	672	2.0	62.46
21	2.101	2.584	0.074	1.726	6.034	451	337	656	1.9	68.90
22	0.394	2.929	0.168	1.013	4.833	500	339	602	1.8	60.11
23		2.223		2.577	6.034	404	294	598	2.0	34.49
mean ± s.d.	1.517 ± 1.105	2.137 ± 1.129	0.227 ± 0.238	1.177 ± 1.093	6.375 ± 1.300	478 ± 117	321 ± 126	598 ± 134	2.0 ± 0.5	33.03 ± 25.31

2.2. Experimental procedure

The experiment involved a total of 41 non-parent participants ($n = 21$ men and 20 women, aged 35 ± 13 years old), with little to no experience with babies. The experiments were conducted in an acoustically isolated chamber (TipTopWood©; figure 3). During an experiment, the participant was seated on a chair, facing a screen giving them instructions. The thermal camera (Opris PI450i) was positioned 1 m from their face. A high-fidelity loudspeaker (Triangle Comete 202) played the cry stimuli. The participant did not wear headphones, piercings, nor glasses (except, if necessary, to read instructions) or any object that could alter the measurements. The participant was filmed continuously throughout the experiment (27 frames s^{-1}).

Forty-one participants completed four listening sessions of 16 cries. After deleting the tests unusable due to the non-alignment of the face with the thermal camera causing problems with facial recognition and temperature evaluation, we obtained a total of 2592 thermal signal measurements. Each stimulus was repeated twice with 1 s of silence between the two repetitions (total duration of the whole experiment per participant = approx. 15 min). The order of presentation of the 23 stimuli was pseudo-randomized between participants. At the end of the stimulus, a message appeared on the screen asking the participant to quote the cry as expressing either mere discomfort or acute pain. There was a variable silent phase of random length ranging from 19 to 27 s, between the participant's response to a given cry and the start of the first repetition of a new cry. The duration of silence varied to avoid physiological or cognitive anticipatory effects in listeners. This duration was chosen to be as long as possible to allow a return to the baseline physiological state, while limiting the total duration of the experiment to approximately 15 min. Each listening session consisted of a set of 16 cries (8 vaccines/8 baths) randomly selected from the 23 available. Before the start of the first cry sequence (beginning of the experiment), there was a 1 min pause to take the basal body temperature.

2.3. Thermal signal processing

Thermal images of faces were extracted using OpenCV library [38]. Each thermal image of the participants' faces (27 frames s^{-1}) was first normalized to greyscale, then recoloured using a normalized RGB scale. On these images, we defined four face areas indicative of facial thermal variations due to their density in vascular vessels: left and right preorbital face area and left and right cheeks (electronic supplementary material, figure S1). These areas were based on the 478 facial landmarks estimated by MediaPipe [39]. The temperature of each of these areas was then calculated for each image by averaging the temperature value of each pixel. A thermal signal was thus obtained for each facial area, corresponding to the variation in the area's average temperature over time.

Most of the thermal signals showed a linear temperature drift over the course of the experiment (continuous increase mean slope: 2.53% ± 43.24%), independently of the temporal dynamics of the sound stimuli. This drift, possibly induced by variation in participants' basal stress levels during the experiment, was suppressed by detrending the signal (scipy.signal.detrend Python's function). This detrending process consisted of subtracting the result of a linear least-squares fit from the data. Next, basal temperature differences between participants were corrected by normalizing each thermal signal by its initial value immediately before the onset of the sound stimulus.

2.4. Measurement of the distance between acoustic and thermal signals

We temporally synchronized the acoustic signal (cry amplitude envelope) and the thermal signal as follows. The two signals, acoustic and thermal, were extracted from 3 s before the start of the first repetition of the cry to 3 s after the end of the second repetition. We then resampled the acoustic envelopes so that they had the same sampling frequency as the thermal signals (27 Hz). Both acoustic and thermal signals were denoised using a Savitzky–Golay filter to reduce high-frequency noise (window length: $N_w = 2 \times 27$, polyorder = 2 [40]). As the duration of cry stimuli was not exactly the same across stimuli, we standardized the duration of all signals, sound and thermal. After these resampling and normalization operations, all signals consisted of 101 successive values.

To quantify the link between the temporal dynamics of the acoustic and thermal signals, we measured the distance between the temporal envelopes of the two signals using a DTW algorithm (symmetrical step model P2 constrained by slope [34]). This algorithm gives a DTW distance value (in arbitrary units) that quantifies the difference between the two signals: the smaller the distance between the signals, the more the thermal signal is temporally related to the acoustic signal (see figure 4). In other words, a small DTW distance indicates that the temporal variations of the thermal signal closely follow those of the acoustic signal. Conversely, a larger distance suggests that variations in thermal signal occur independently of temporal variations in the acoustic signal.

The DTW calculation is particularly well suited to assessing the degree of similarity (or distance) between two time series (here the acoustic signal and the thermal signal) that are likely to exhibit phase shifts. Although the similarity between two time series can be assessed by calculating the Euclidean distance between them (or other similar distances), this calculation is particularly sensitive to phase shifts between the series. If it is based on comparing the value of the first time series at time T with the value of the second time series at the same time T, the distance between time series that are out of phase in time but nevertheless similar will be large, in other words the series will be assessed as not being very similar. This problem can be overcome by using DTW. The DTW compares the amplitude of the first signal at time T with the amplitude of the second signal at times $T + 1$ and $T - 1$ (or $T + 2$ and $T - 2$). This ensures that it does not give a low similarity score (high distance) for signals with a similar shape and different phase.

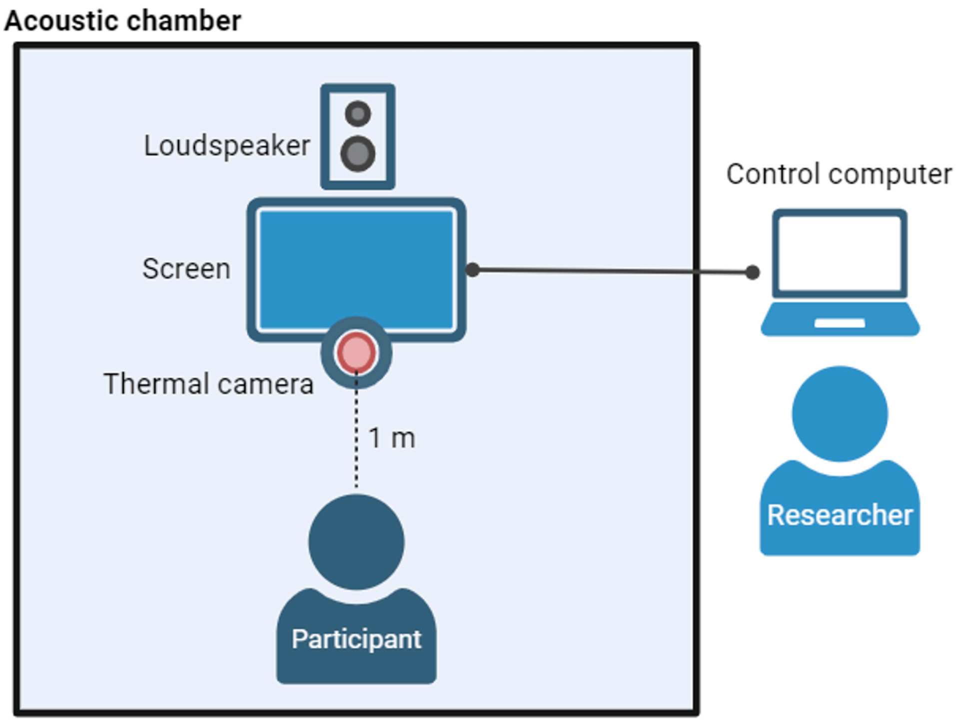

Figure 3. Experimental set-up. Each participant is placed in an acoustic chamber equipped with a loudspeaker, a screen and a thermal camera at a distance of 1 m. Outside the chamber, the researcher monitors the experiment using a computer.

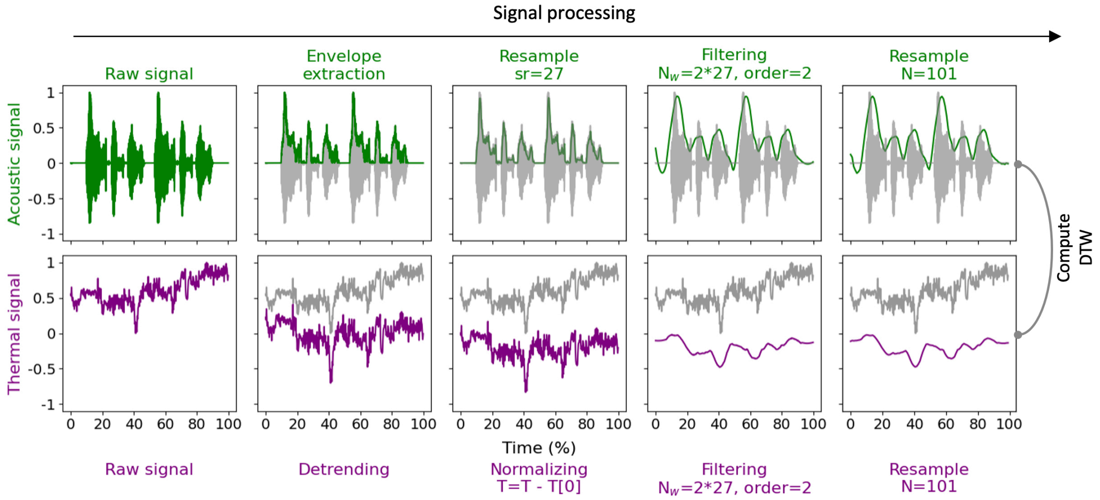

Figure 4. Flow of signal processing steps to quantify the distance between the acoustic signal (baby cry) and the thermal signal (temperature of the listener's face).

2.5. Statistical analysis

We performed statistical analyses using Bayesian mixed models fitted with the 'brms' R package ([41], R v. 4.2.2). The advantages of using the Bayesian approach are manifold, including its high flexibility, quantification of uncertainty in estimates, use of priors for regularization and intuitive interpretation of confidence intervals. Posterior distributions of model parameters were summarized by their median and 95% credibility interval.

To test whether nonlinear acoustic phenomena drive pain ratings by adult listeners, we built the following model: Response ~ NLP × sex + (1|participant) + (1|CryID) + (1|F0_range), where the variable 'Response' corresponds to participants' evaluation of the sound stimulus (i.e. discomfort cry or pain cry) modelled by a logistic function (Bernouilli family) and the fixed factors are the level of NLP in the cry (expressed in percentage of the cry with NLP; 'NLP') and the sex of the adult listener ('sex'). Random factors were the identity of the participant ('participant'), the identity of the cry ('CryID'), and the pitch variation ('F0_range'), measured by dividing the maximum by the minimum cry pitch). We chose to control for pitch variation defined as such rather than for the mean pitch or another pitch value because pain cries are known to have variations with extreme pitch values (sudden incursion of pitch into high frequencies) [8,9,11].

To test whether the presence of nonlinear acoustic phenomena in the cries leads to synchronization of the cry amplitude envelope with variation in the thermal signal, we built the following model: DTW ~ NLP × sex + (1|participant) + (1|CryID) + (1|F0_range), where 'DTW' is the DTW distance between the temporal dynamics of the acoustic and thermal signals. The fixed factors are the level of NLP ('NLP') and the sex of the adult listener ('sex'). Random factors are participant identity

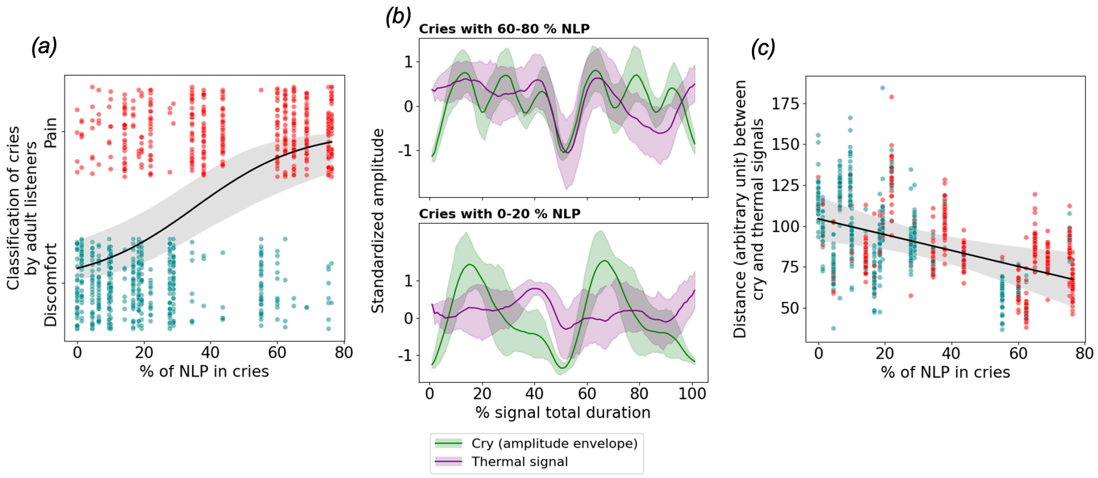

Figure 5. Impact of baby cries' NLP on adult listeners. (*a*) Rating of cries by adult participants according to the level of NLP. The participants were more inclined to classify as expressing pain the cries presenting the most NLP. (*b*) Temporal dynamics of the cry signal (green) and thermal signal (purple) for the 20% of cries showing the most (top) and least (bottom) NLP, respectively. The green curves represent the average stimulus amplitude envelopes with their 95% confidence interval. The purple curves represent the time course of the thermal signals with their 95% confidence interval (measurements taken in the right preorbital face area). Each cry stimulus consisted of two repetitions of the same cry, separated by silence. These curves suggest that the temporal dynamics of the cry stimulus and that of facial temperature are more similar for cries with more NLP. (*c*) Distance between the temporal dynamics of cry and thermal signals. The distance is calculated using a DTW algorithm (see text for details) for the right preorbital face area. The more nonlinear acoustic phenomena the cry expresses, the more related are the acoustic and thermal temporal dynamics.

('participant'), the identity of the cry ('CryID'), and the pitch variation ('F0_range'). This model was applied to different portions of the stimulus, i.e. either the entire stimulus, the first repetition of the cry, or the second repetition of the cry. All these measurements and statistics were carried out on the four areas of the face.

3. Results

As illustrated in figure 5*a*, participants classified cries as either discomfort or pain cries according to their level of NLP. Cries with a high level of NLP were more likely to be classified as pain cries (0.82, 95% CI [0.51, 0.94]) compared with cries with a little or no NLP. This rating was independent of the listener's sex (0.80 [0.45, 0.94] for men; 0.84 [0.55, 0.95] for women; difference between sexes −0.04 [−0.13, 0.03]).

We then found that the percentage of NLP in the cries impacts the DTW distance between the temporal dynamics of the cry (amplitude envelope) and the thermal signal. This is illustrated in figure 5*b*, which shows that the temporal dynamics of the acoustic and thermal signals are more similar for the 20% of cries with the highest level of NLP (top panel) than for the 20% of cries with the lowest level of NLP (bottom panel), this for the right preorbital face area. Figure 5*c* further illustrates this influence of NLP by showing a decrease in the distance between the acoustic signal and the thermal signal as the level of NLP in the cry increases (right preorbital face area: decrease in the acoustic–thermal distance from 0% to 80% NLP = −37.05 [−63.04, −11.59]). This convergence between the temporal dynamics of the two signals when more NLP are present is observed for each of the two repetitions of the cry within the stimulus (first repetition of the cry: −18.62 [−30.90, −5.89]); second repetition of the cry: −18.38 [−31.28, −5.56]). The same results are found in the left preorbital region (−37.48 [−63.26, −11.99]) as well as in the left and right cheeks (and −38.38 [−64.36, −12.35] and −39.14 [−65.00, −13.03], respectively; for details, see electronic supplementary material, table S1 and figure S2). We did not identify any difference between men and women participants (electronic supplementary material, table S2).

4. Discussion

The present study shows that baby cries trigger an autonomic nervous system response resulting in listeners' facial temperature variations. Specifically, our results demonstrate that the level of NLP in a cry modulates the temporal dynamics of the facial thermal response in listeners, independent of their sex. Thermography thus appears as a possible method for investigating physiological effects of baby cries on adult listeners.

It has been established that NLP are reliable markers of the level of distress and/or pain expressed by the baby [8]. The present study's psychoacoustic experiment confirms that adult listeners pay attention to these acoustic markers when judging the degree of distress. The joint analysis of acoustic and thermal signals shows that the presence of these acoustic markers results in a lesser distance between the temporal dynamics of these signals. A cry presenting a large level of NLP thus mobilizes the autonomic nervous system response more effectively than a cry with a low level or no NLP. This increased synchronization between facial thermal variations and the cry signal may reflect a more efficient mobilization of the neurophysiological processes involved in responding effectively to the baby's solicitations.

Our results further reveal that the autonomic nervous system's response, which varies with the pain level expressed in an infant's cry, is consistent across adult listeners regardless of their sex. This observation builds on prior studies showing that men and women are equally proficient in identifying infants through their cries and evaluating the pain expressed [10,33,42,43]. The present research thus strengthens evidence for this absence of sex-based differences in baby cry perception. In the context of human cooperative breeding [44], this shared responsiveness to infant cries is a compelling and expected trait. Despite its potential societal implications, this phenomenon remains nevertheless largely unrecognized by the general public.

To the best of our knowledge, the present study is the first to investigate the facial thermal response of adults to the listening to baby cries. While our results sound interesting and new, this study remains rather preliminary and raises a number of questions, both in terms of interpreting the results and from a methodological point of view. The first point concerns the question of the respective impact of each type of NLP on the thermal response. Indeed, as reported in §1, NLP encompasses various acoustic phenomena, and it would be valuable to assess the contribution of each to eliciting a thermal response. While this is a clear objective, it cannot be reliably achieved in this study. Indeed, we utilized natural baby cries, which contain varying proportions of different types of NLP. Consequently, it is not possible to isolate the effect of any specific NLP. Capturing the thermal response to cries based on each NLP independently of the others is thus not possible here. A good strategy to address this gap would require the use of synthetic cry replicas, allowing precise control over the presence and quantity of each NLP type. At the ENES Bioacoustics Research Laboratory, we have begun such investigations (though without a thermal approach) [8].

A second point, which follows the same line of reasoning, is that we did not test the influence of mean pitch and pitch variation on the thermal response of listeners. The cry stimuli we used here, selected to represent the diversity of NLP, did not allow for rigorous testing of this pitch influence. It is likely that pain cries with significant pitch variations over time could also elicit a thermal response in listeners.

A third point concerns the participants. We only tested people with no or little experience of babies. This choice was made to increase the homogeneity of the sample, given that it was impossible for logistical reasons to test a larger cohort of participants. It should be remembered that prior experience with babies is a major factor driving the ability of adult listeners to assess baby cries. In an earlier study, we showed that parents of young children were able to distinguish between cries of discomfort and cries of pain without difficulty, whereas inexperienced listeners largely failed to do so [33]. It is therefore possible that experienced people can more acutely detect the presence of NLP in cries. Variations in their facial temperature might be even more dependent on the level of NLP contained in the cry. Conversely, it is possible that experienced people may be less easily emotionally affected by babies' cries, and so the variations in their facial temperature may be much more modest. Another important factor probably leading to variability between participants could be their degree of empathy. In a recent study, we showed that the mobilization of the brain connectome involved when listening to baby cries did indeed depend on the degree of empathy of the listeners [45]. This could arise from distinct mechanisms. One possibility is that prior experience leads to desensitization or habituation to distress cues, resulting in a dampened autonomic response over time. Alternatively, in a cooperative breeding context, individuals who are *inherently* less autonomically reactive or sensitive to aversive stimuli like intense crying might be more likely to engage in and persist with caregiving roles; this would represent a selection effect rather than acquired habituation. Intriguingly, this intersects with potential trade-offs involving other motivational systems, such as disease avoidance. A recent study by Gustafsson [46] reported that priming participants with cues of infant infection reduced tenderness towards cute babies, but only among non-carers. This aligns with both possibilities: experienced carers might be desensitized not only to distress but also potentially to pathogen cues associated with caregiving, or individuals inherently less prone to disgust or pathogen avoidance might be more predisposed to invest in childcare. If such traits correlate with autonomic reactivity, it could influence the thermal responses observed.

These considerations have potentially significant implications for interpreting our findings and future work. If, as our study might imply (though not directly tested here), baby cries elicit a stronger or more synchronized autonomic response primarily in non-caregivers, the functional meaning of this response needs careful consideration. Does the heightened facial temperature variation reflect increased irritation and aversion, potentially acting as a barrier to caregiving for those not habituated or temperamentally suited? Or does it signal heightened alertness, attention and perhaps even sympathy, representing a mobilization of resources to potentially respond, as suggested by related work on empathy and brain responses to cries [45] and the general link between emotion and facial thermal signals [25]? Future studies directly comparing experienced caregivers and non-caregivers, potentially combining thermography with subjective ratings of irritation/sympathy and other physiological measures, are crucial to disentangle these competing hypotheses regarding habituation, individual predisposition, motivational trade-offs and the ultimate meaning of the autonomic response to baby cries.

The last point of attention is a technical one. Facial thermography, while non-invasive and promising for capturing autonomic responses, is sensitive to various potential confounds. Factors such as ambient room temperature fluctuations, although mitigated by conducting experiments in a controlled chamber, can influence skin temperature. More significantly, there are inherent individual differences in baseline facial temperature and, crucially, in vascular reactivity. While our method of normalizing each thermal signal by its initial value immediately before the stimulus onset effectively corrects for baseline temperature differences between trials and participants, and detrending addressed slow linear drifts, these steps do not fully eliminate the influence of individual differences in vascular response patterns or potential nonlinear drifts unrelated to the stimuli. Consequently, some variability in the thermal responses observed may stem from this physiological heterogeneity among participants, independent of the cry characteristics. Moreover, the participants' immediate psychological or physiological state (e.g. stress, fatigue) could potentially modulate their autonomic response. Furthermore, analysing the dynamics of the facial thermal response as a function of the temporal dynamics of a cry is no easy task. In view of the results obtained, the method we have chosen appears promising. However, it relies on resampling and filtering the signals, both sound and thermal, to work on the same time scale. These manipulations, although controlled, are likely to mask subtle phenomena present in what we consider noise, and to create undesirable artefacts that have not yet been highlighted. In short, while we feel that we have made a significant advance here, we also believe that much remains to be done before the thermal imaging tool can be used routinely to monitor participants' response to baby cries.

Ethics. All participants gave written informed consent to the study approved by the French Ethics Committee CPP-EST IV (no. ID RCB 2020-A02993-36) and received a 10€ compensation for their participation.

Data accessibility. Data used to produce this analysis are available at Zenodo [47]. All other data are available in the manuscript or the electronic supplementary material.

Supplementary material is available online [48].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. L.L.: formal analysis, investigation, methodology, writing—original draft; C.C.: investigation; S.C.: investigation, methodology, resources, writing—review and editing; M.M.: data curation, methodology; L.P.: methodology; H.P.: conceptualization; D.R.: conceptualization, writing—review and editing; F.J.: conceptualization, formal analysis, funding acquisition, investigation, methodology, validation, writing—review and editing; N.M.: conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. Funding was provided by the ANR (BABYCRY project and SCREAM project), the Mécénat des Mutuelles AXA (N.M.), the Lyon IDEX Fellowship (D.R.), the LABEX CELYA (France), the University of Saint-Etienne (France), the Centre National de la Recherche Scientifique (CNRS, France), the Institut National de la Santé et de la Recherche Médicale (Inserm, France), and the Institut Universitaire de France (N.M. and D.R.). S.C. was a recipient of a PhD grant from the Fondation pour la Recherche Médicale (FRM; code ECO202006011648).

Acknowledgements. We are grateful to the babies, their parents and the participants.

References

1. Bornstein MH, Tamis-LeMonda CS, Tal J, Ludemann P, Toda S, Rahn CW, Pêcheux MG, Azuma H, Vardi D. 1992 Maternal responsiveness to infants in three societies: the United States, France, and Japan. *Child Dev*. **63**, 808–821. (doi:10.1111/j.1467-8624.1992.tb01663.x)
2. Mathevon N. 2025 *The intimate world of babies' cries: the best ways to understand and calm your baby*. Saint-Etienne, France: independently published.
3. Lockhart-Bouron M *et al*. 2023 Infant cries convey both stable and dynamic information about age and identity. *Commun. Psychol.* **1**, 1–15. (doi:10.1038/s44271-023-00022-z)
4. Koutseff A, Reby D, Martin O, Levrero F, Patural H, Mathevon N. 2018 The acoustic space of pain: cries as indicators of distress recovering dynamics in pre-verbal infants. *Bioacoustics* **27**, 313–325. (doi:10.1080/09524622.2017.1344931)
5. LaGasse LL, Neal AR, Lester BM. 2005 Assessment of infant cry: acoustic cry analysis and parental perception. *Ment. Retard. Dev. Disabil. Res. Rev*. **11**, 83–93. (doi:10.1002/mrdd.20050)
6. Soltis J. 2004 The developmental mechanisms and the signal functions of early infant crying. *Behav. Brain Sci*. **27**, 477–490. (doi:10.1017/s0140525x0442010x)
7. Cornec C, Mathevon N, Pisanski K, Entani D, Monghiemo C, Bola B, Planas-Bielsa V, Reby D, Levréro F. 2024 Human infant cries communicate distress and elicit sex stereotypes: cross cultural evidence. *Evol. Hum. Behav*. **45**, 48–57. (doi:10.1016/j.evolhumbehav.2023.08.004)
8. Corvin S, Massenet M, Hardy A, Patural H, Peyron R, Fauchon C, Mathevon N. 2025 Nonlinear acoustic phenomena affect the perception of pain in human baby cries. *Phil. Trans. R. Soc. B* **380**, 20240023. (doi:10.1098/rstb.2024.0023)
9. Laguna A *et al*. 2023 How can cry acoustics associate newborns' distress levels with neurophysiological and behavioral signals? *Front. Neurosci*. **17**, 1266873. (doi:10.3389/fnins.2023.1266873)
10. Corvin S, Fauchon C, Patural H, Peyron R, Reby D, Theunissen F, Mathevon N. 2024 Pain cues override identity cues in baby cries. *iScience* **27**, 110375. (doi:10.1016/j.isci.2024.110375)
11. Bellieni CV, Sisto R, Cordelli DM, Buonocore G. 2004 Cry features reflect pain intensity in term newborns: an alarm threshold. *Pediatr. Res*. **55**, 142–146. (doi:10.1203/01.pdr.0000099793.99608.cb)
12. Massenet M, Mathevon N, Anikin A, Briefer EF, Fitch WT, Reby D. 2025 Nonlinear phenomena in vertebrate vocalizations: mechanisms and communicative functions. *Phil. Trans. R. Soc. B* **380**, 20240002. (doi:10.1098/rstb.2024.0002)
13. Massenet M, Pisanski K, Reynaud K, Mathevon N, Reby D, Anikin A. 2025 Acoustic context and dynamics of nonlinear phenomena in mammalian calls: the case of puppy whines. *Phil. Trans. R. Soc. B* **380**, 20240022. (doi:10.1098/rstb.2024.0022)
14. Anikin A, Herbst CT. 2025 How to analyse and manipulate nonlinear phenomena in voice recordings. *Phil. Trans. R. Soc. B.* **380**, 20240003. (doi:10.1098/rstb.2024.0003)
15. Witteman J, Van IJzendoorn MH, Rilling JK, Bos PA, Schiller NO, Bakermans-Kranenburg MJ. 2019 Towards a neural model of infant cry perception. *Neurosci. Biobehav. Rev*. **99**, 23–32. (doi:10.1016/j.neubiorev.2019.01.026)
16. Amin R, Faghih RT. 2022 Physiological characterization of electrodermal activity enables scalable near real-time autonomic nervous system activation inference. *PLoS Comput. Biol*. **18**, e1010275. (doi:10.1371/journal.pcbi.1010275)
17. Ghiasi S, Greco A, Barbieri R, Scilingo EP, Valenza G. 2020 Assessing autonomic function from electrodermal activity and heart rate variability during cold-pressor test and emotional challenge. *Sci. Rep*. **10**, 5406. (doi:10.1038/s41598-020-62225-2)
18. Mccafferty DJ. 2007 The value of infrared thermography for research on mammals: previous applications and future directions. *Mammal Rev*. **37**, 207–223. (doi:10.1111/j.1365-2907.2007.00111.x)
19. Tattersall GJ. 2016 Infrared thermography: a non-invasive window into thermal physiology. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **202**, 78–98. (doi:10.1016/j.cbpa.2016.02.022)
20. Ramirez-GarciaLuna JL, Bartlett R, Arriaga-Caballero JE, Fraser RDJ, Saiko G. 2022 Infrared thermography in wound care, surgery, and sports medicine: a review. *Front. Physiol*. **13**, 838528. (doi:10.3389/fphys.2022.838528)

21. Ioannou S, Morris P, Terry S, Baker M, Gallese V, Reddy V. 2016 Sympathy crying: insights from infrared thermal imaging on a female sample. *PLoS One* **11**, e0162749. (doi:10.1371/journal.pone.0162749)
22. Baker CJ, Kasper DL. 1976 Correlation of maternal antibody deficiency with susceptibility to neonatal Group B streptococcal infection. *N. Engl. J. Med*. **294**, 753–756. (doi:10.1056/nejm197604012941404)
23. Clay-Warner J, Robinson DT. 2015 Infrared thermography as a measure of emotion response. *Emot. Rev*. **7**, 157–162. (doi:10.1177/1754073914554783)
24. Ermatinger FA, Brügger RK, Burkart JM. 2019 The use of infrared thermography to investigate emotions in common marmosets. *Physiol. Behav*. **211**, 112672. (doi:10.1016/j.physbeh.2019.112672)
25. Ioannou S, Gallese V, Merla A. 2014 Thermal infrared imaging in psychophysiology: potentialities and limits. *Psychophysiology* **51**, 951–963. (doi:10.1111/psyp.12243)
26. Jarlier S, Grandjean D, Delplanque S, N'Diaye K, Cayeux I, Velazco MI, Sander D, Vuilleumier P, Scherer KR. 2011 Thermal analysis of facial muscles contractions. *IEEE Trans. Affect. Comput*. **2**, 2–9. (doi:10.1109/t-affc.2011.3)
27. Marqués-Sánchez P, Liébana-Presa C, Benítez-Andrades JA, Gundín-Gallego R, Álvarez-Barrio L, Rodríguez-Gonzálvez P. 2020 Thermal infrared imaging to evaluate emotional competences in nursing students: a first approach through a case study. *Sensors* **20**, 2502. (doi:10.3390/s20092502)
28. Robinson DT, Clay-Warner J, Moore CD, Everett T, Watts A, Tucker TN, Thai C. 2012 Toward an unobtrusive measure of emotion during interaction: thermal imaging techniques. In *Biosociology and neurosociology* (eds W Kalkhoff, SR Thye, EJ Lawler), pp. 225–266. Leeds, UK: Emerald Group Publishing Limited. (doi:10.1108/s0882-6145(2012)0000029011)
29. Sonkusare S, Breakspear M, Guo C. 2019 Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci*. **23**, 699–714. (doi:10.1016/j.tics.2019.05.004)
30. Fernández J, Albayay J, Gálvez-García G, Iborra O, Huertas C, Gómez-Milán E, Caballo VE. 2024 Facial infrared thermography as an index of social anxiety. *Anxiety Stress Coping* **37**, 114–126. (doi:10.1080/10615806.2023.2199209)
31. Travain T, Valsecchi P. 2021 Infrared thermography in the study of animals' emotional responses: a critical review. *Animals* **11**, 2510. (doi:10.3390/ani11092510)
32. Chotard H, Ioannou S, Davila-Ross M. 2018 Infrared thermal imaging: positive and negative emotions modify the skin temperatures of monkey and ape faces. *Am. J. Primatol*. **80**, e22863. (doi:10.1002/ajp.22863)
33. Corvin S, Fauchon C, Peyron R, Reby D, Mathevon N. 2022 Adults learn to identify pain in babies' cries. *Curr. Biol*. **32**, R824–R825. (doi:10.1016/j.cub.2022.06.076)
34. Boersma P. 2001 Praat, a system for doing phonetics by computer. *Glot Int.* **5**, 341–345.
35. Romanow CA, Riede T, Lingle S. 2025 Vocal characteristics of distress and reproductive vocalizations in North American wapiti. *Curr. Zool.* zoaf006. (doi:10.1093/cz/zoaf006)
36. Titze I, Riede T, Mau T. 2016 Predicting achievable fundamental frequency ranges in vocalization across species. *PLoS Comput. Biol*. **12**, e1004907. (doi:10.1371/journal.pcbi.1004907)
37. Anikin A. 2019 Soundgen: an open-source tool for synthesizing nonverbal vocalizations. *Behav. Res. Methods* **51**, 778–792. (doi:10.3758/s13428-018-1095-7)
38. OpenCV DT. 2015 *Camera calibration and 3D reconstruction*. See https://docs.opencv.org/4.x/d9/d0c/group__calib3d.html.
39. Lugaresi C *et al*. 2019 MediaPipe: a framework for building perception pipelines. *arXiv.* (doi:10.48550/arXiv.1906.08172)
40. Savitzky A, Golay MJE. 1964 Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639. (doi:10.1021/ac60214a047)
41. Bürkner PC. 2017 brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 01–28. (doi:10.18637/jss.v080.i01)
42. Gustafsson E, Levréro F, Reby D, Mathevon N. 2013 Fathers are just as good as mothers at recognizing the cries of their baby. *Nat. Commun*. **4**, 1698. (doi:10.1038/ncomms2713)
43. Reby D, Levréro F, Gustafsson E, Mathevon N. 2016 Sex stereotypes influence adults' perception of babies' cries. *BMC Psychol*. **4**, 19. (doi:10.1186/s40359-016-0123-6)
44. Hrdy SB. 2024 *Father time: a natural history of men and babies*, pp. 1–424. Princeton, NJ: Princeton University Press.
45. Fauchon C, Corvin S, Faillenot I, Patural H, Reby D, Peyron R, Mathevon N. In Press. Parenting and empathy capabilities drive brain response to pain cues in baby cries submit.
46. Gustafsson E. 2024 Child care and disease avoidance: exploring motivational systems trade-offs. *Lett. Evol. Behav. Sci.* **15**, 52–59. (doi:10.5178/lebs.2024.122)
47. Lego L, Cornec C, Jouen F, Mathevon N. 2025 Data for 'Nonlinear acoustic phenomena tune the adults' facial thermal response to baby cries with the cry amplitude envelope' (Journal of the Royal Society Interface, 2025). Zenodo. (doi:10.5281/ZENODO.15341516)
48. Corvin S, Massenet M, Hardy A, Patural H, Peyron R, Fauchon C *et al*. 2025 Supplementary material from: Nonlinear acoustic phenomena affect the perception of pain in human baby cries. Figshare. (doi:10.6084/m9.figshare.c.7720934)