Title: The use of infra-red thermography to study emotions of wild crested macaques (*Macaca nigra*) in their natural habitat.

Berthier, Juliette M., Wheeler, Brandon C., Qomariah, Indira N., Perwitasari-Farajallah, Dyah, Newton-Fisher, Nicholas E.

Physiology & Behavior. ISSN 0031-9384.

Abstract

Infra-red thermography (IRT) has been validated across a range of taxa as a non-invasive method to quantify physiological changes (variation in skin temperature) that serve as a proxy for emotional arousal in humans and other animals. While its effectiveness has been demonstrated in captivity, its application in wild animals under natural conditions remains underexplored. This study validates IRT in wild crested macaques (*Macaca nigra*), observed in their natural habitat, to assess its potential for studying short-term emotions as proximate mechanisms of behavior of wild animals. Seventy-five playback experiments and predator model presentations were conducted on 14 adult females, exposing them to stimuli including affiliative grunts, contact calls, aggressive screams, alarm calls, and python models. The objectives were to: (1) evaluate whether environmental and physical factors, such as atmospheric conditions and camera-subject distance, impact thermal measurements; (2) determine whether facial skin temperature variations reliably indicate emotional arousal and valence in response to ecologically relevant stimuli; and (3) explore correlations between temperature changes and behavioral indicators of arousal, such as physical activity and self-directed behaviors. Results showed that: (1) environmental and physical factors did not significantly affect thermal measurements when following our protocol; (2) facial skin temperature changes reliably reflected the intensity of emotional arousal elicited by different stimuli; and (3) temperature variation was associated with behavioral indicators of arousal and anxiety.

These findings establish IRT as a reliable tool for investigating emotion-driven skin temperature changes in wild primates, providing valuable insights into the role of short-term emotions in shaping behavior.

Key-words: Emotional Arousal, Field Experiments, Infra-Red Thermography, Natural Environment, Self-Directed Behaviors, Wild Primates.

1. Introduction

Infra-red thermography (IRT) has emerged as a promising, non-invasive and contact-free method for assessing emotional states in animals. By capturing infrared radiation emitted from an object [1,2], IRT allows for accurate measurement of temperature changes in peripheral body regions. These fluctuations in response to affective situations are linked to changes in subcutaneous blood flow regulated by the sympathetic branch of the autonomic nervous system [3]. Typically, negative emotions (e.g., fear and anxiety) trigger peripheral vasoconstriction, redirecting blood to vital organs and leading to a decrease in skin temperature, followed by stress-induced hyperthermia (SIH), characterized by a rise in core body temperature [4,5]. These changes are detectable using IRT camera, making it a valuable tool for studying emotional responses of a wide range of mammals and birds [6,7].

Numerous studies have successfully employed IRT to assess emotional responses in various domesticated and captive animals (horses: *Equus caballus* [8–10]; sheep: *Ovis aries* [11]; hens: *Gallus domesticus* [12]; cows: *Bos taurus* [13,14]; see also [15,16] for reviews). The method has proven effective in evaluating SIH activation across diverse species, including birds (blue tits: *Cyanistes caeruleus* [17]; house sparrows: *Passer domesticus* [18,19]; herring gulls: *Larus argentatus* [20]; little auk: *Alle alle* [21]) and mammals (rats: *Rattus norvegicus* [22], mice: *Mus musculus*: [23–25]; seals: *Phoca vitulina* [26]), with the magnitude of skin temperature variation often reflecting the intensity of the stressor (hybrid rabbits: *Oryctolagus* sp. [27], rhesus macaques: *Macaca mulatta* [28]; dogs: *Canis familiaris* [29]; hens [12], mice [23]). Additionally, temperature changes frequently correlate with other

physiological markers, such as heart rate or heart rate variability [9,18,29,30,31], reinforcing the connection between thermal responses and autonomic regulation.

In humans, IRT has been applied to study facial temperature changes in response to different emotional states including laughing, lying, stress, fear, guilt, sexual pleasure, love and empathy [32–36], with distinct thermal patterns observed across multiple regions of the face [37], underscoring the need to use a multi-region approach to differentiate between emotional states. IRT has also been successfully applied to non-human primates including chimpanzees (*Pan troglodytes)* [30], rhesus macaques [38], Western lowland gorillas (*Gorilla gorilla gorilla*) [39], common marmosets (*Callithrix jacchus*) [40,41], and others [42], revealing significant decreases in nasal temperature in response to stressful stimuli. However, thermal responses to positive stimuli have been less extensively studied, but some studies found increases in nose temperature during positive emotional states such as during grooming [31] or the presentation of preferred food [40].

IRT has often been combined with behavioral observations to enhance the reliability and precision of emotional state assessments such as spatial displacements, movements of the body or body parts, and piloerection [40,41,43]. Self-directed behaviors (SDBs), such as scratching, self-grooming and yawning, are commonly used as behavioral proxies for uncertainty and anxiety in animals, especially primates [44,45]. These SDBs are more frequent in stressful situations, highlighting their role as indicators of negative emotional states [46–51]. Despite the promise of combining IRT with behavioral observations, little research has explicitly examined the relationship between surface temperature variation and the expression of SDBs or physical activities, particularly in wild settings.

While IRT has demonstrated its utility in studying emotions in captive animals, its application in wild settings remains underexplored (but see [52–54]). The complexities of wild environments, such as uncontrollable atmospheric conditions, weather, and subject-camera distance, pose challenges for

78 accurate thermal measurement [55–60]. Nevertheless, pioneering work investigating emotional
79 responses in wild chimpanzees (*P. troglodytes schweinfurthii*) in Uganda, suggest that IRT can be
80 effectively applied in natural settings when rigorous protocols are followed to mitigate environmental
81 influences [61].
82
83 This study aims to validate IRT as a method for studying short-term emotional responses in wild
84 mammals and birds, with habituated crested macaques (*Macaca nigra*) in their natural environment
85 as a model species. To do so, we performed playback experiments and predator model presentations
86 on adult females, using four types of vocalizations – affiliative grunts, contact calls, aggressive screams,
87 and alarm calls – and two model pythons. The first objective is to establish a rigorous protocol of data
88 collection and evaluate whether environmental and physical factors—such as atmospheric and
89 environmental conditions, camera-subject distance, and angle of view—affect thermal measurements
90 obtained with this method. We hypothesize that stringent methodological controls will minimize these
91 effects, supporting IRT as a reliable tool for examining emotionally induced skin temperature variations
92 in the wild. The second objective examines whether variation in facial temperature reflect degrees of
93 emotional arousal and valence (in the framework of affective neuroscience, emotional arousal refers
94 to the intensity of an emotional response, while emotional valence pertains to its positive or negative
95 quality [62,63]) in response to different stimuli (affiliative, neutral, aggressive, and predation-related).
96 We predict positive stimuli (i.e., affiliative grunts) will increase facial temperature, whereas negative
97 stimuli (i.e., aggressive screams, alarm calls, and predator models) will produce graded decreases in
98 temperature linked to anxiety or fear levels (see **Fig. 1**). The last objective explores correlations
99 between emotionally-mediated facial temperature variation and behavioral indicators of arousal and
100 anxiety, such as spatial displacement frequency (hereafter named '*physical activity*') and the
101 expression of SDBs. We expect that lower facial temperatures will align with increased physical activity
102 and higher SDB rates, particularly in response to negative stimuli.
103

2. Method

2.1. Study site and animals

The data presented in this study were collected on wild adult female crested macaques belonging to one group between March and December 2021 at the Duasudara Nature Reserve (formerly known as Tangkoko Nature Reserve), North Sulawesi, Indonesia, in collaboration with the Macaca Nigra Project (MNP) [65]. The crested macaque is one of seven macaque species endemic to Sulawesi Island, Indonesia [66]. Its geographic distribution is limited to northern Sulawesi, the island of Manado Tua, and Talise Island. Additionally, some individuals were introduced to Bacan, Halmahera, and Kasiruta Islands in the North Maluku province, where they have established self-sustaining populations outside their native range [67].

The Duasudara Nature Reserve has an area of 8,867 hectares and an elevation from sea level up to 1,350 meters. Annual rainfall ranges from 1,550 to 2,400 mm. Temperatures are relatively constant between 21.7 and 34.4°C. The vegetation comprises lowland tropical rainforest, categorized into primary and secondary forest, and shrubland. Primary forest features a semi-closed canopy of 40-50-meter-tall buttressed trees, numerous lianas, and sparse understory. In contrast, secondary forest has an open canopy with cut-over areas, dense mid-layer, and abundant forest floor vegetation. For further habitat details of the site, see [68].

Crested macaques are medium-sized, predominantly terrestrial primates, spending most of their diurnal time traveling, foraging, and feeding on the ground [69]. Females exhibit a tolerant social style and engage frequently in social interactions on the ground [70], facilitating thermal data collection. Three groups of wild macaques have been followed daily since 2006 by the MNP. All the adult males and females are individually recognized and habituated to human presence. Fifteen subjects were randomly chosen among the 19 adult females present in the group named PB1B (19-25 adult females, 8-11 adult males, around 70 individuals in total) at the beginning of data collection to take part of the experiments

(ages estimated between 6 and 16 years old). However, due to the Covid-19 pandemic that caused substantial delays in data collection, we were not able to conduct all the experiments on the 15 subjects as initially planned, reducing the sample size to 14. **Table S1** provides more details about the subjects and their involvement in the different field experiments.

This sample size was based on practical constraints inherent to field studies of wild primates, where ethical considerations and logistical challenges limit subject availability and experimental repetitions, and at the same time was considered sufficient to provide a meaningful representation of the population under study [71]. Moreover, previous studies investigating emotional responses in non-human primates using infrared thermography have drawn reliable conclusions with similar or smaller sample sizes (e.g., [30, 40]). To account for individual variability, we used repeated measures for each subject and applied linear mixed-effects models [72], a statistical approach well-suited to handle within-subject variation and unbalanced data (see *section 2.8* for more details). This method allowed for robust analyses while respecting ethical obligations toward minimally invasive research in a natural environment.

2.2. Stimuli used to generate different contexts

In their natural habitat, the main predator of crested macaques is the reticulated python (*Python reticulatus*). The detection of a python is assumed to generate a highly negative emotional response, perceived as highly threatening to the survival of the subjects. In this study, two distinct natural-size python models were randomly used in the experiments. The first model was home-made using python skin-like fabric that was hand-sewn and filled with cotton stuffing. The second model consisted of a stuffed toy that had been customized through fabric painting and the addition of pieces of python skin-like fabric to enhance its resemblance to an authentic python (**Fig. S1**).

We used the following vocalizations during playback experiments: (1) Contact calls, defined as a "*brief and tonal sound with wide frequency modulation; given when individuals are several to ten meters apart*" ([73; p. 205] & **Fig. S2a**). There are assumed to be neutral and should not modify the emotional state of

the subjects significantly; (2) Affiliative grunt, defined as "*a grunting sound, often emitted in series, and uttered in affiliative interactions and contacts*" ([73; p. 207] & **Fig. S2c**), often emitted together with affiliative facial expressions commonly produced to initiate or maintain an affiliative interaction [74]. Grunts are assumed to generate a positive emotional response associated with an affiliation context and to have a calming effect on the subjects; (3) Screams (**Fig. S2d**) are noisy and repeated vocalizations uttered during aggressive interactions and often accompanied by aggressive facial displays that act as threat or protest in agonistic contexts [73]. Screams are assumed to be aversive, generating a negative emotional response linked to a context of aggression, but not perceived as particularly threatening by the subjects; (4) When detecting a predator (particularly pythons but also vipers, raptors, dogs and humans – *personal observations*), the macaques produce series of alarm calls. This vocalization is defined as "a *brief sound composed by a tonal unit followed by a harsh unit and sounds like a bark"* ([73; p. 205] & **Fig. S2b**)*,* and, uttered repeatedly, leads other groupmates to approach and engage in mobbing behavior ([71]; *personal observations*). It is assumed to generate a highly negative emotional response associated with a context of predation risk and to be perceived as highly threatening for the survival by the subjects.

2.3. Audio recordings and playback preparation

Naturally-occurring contact calls, screams and affiliative grunts were recorded opportunistically on the 19 females with a Marantz PMD660 audio recorder (16 bits, PCM, frequency response: 20–20 000 Hz+3.0 dB, sampling rate: 44.1 kHz) and a Sennheiser K6-ME66 directional microphone (Wedemark, Germany, frequency response: 40–20 000 Hz+2.5 dB). The recordings were made at a distance of 3-5 m from the target animal when the background noise was low. For each recording, the ID of the caller, date, time, and general context were recorded. Alarm calls were captured during model python experiments (see below).

Preparation of the calls was performed with Audacity® 2.3.2 (Audacity Team, 1999-2018). The stimuli used for the playback were prepared by extracting a series of contact calls, affiliative grunts, screams or alarm calls from the best recordings of each caller. In some cases, the vocalizations of the same individual were used several times on different subjects and different stimuli from different recordings were created to avoid pseudoreplication. To maintain consistency between the stimuli, we selected only call series with a high signal-to-noise ratio. Background noise at the onset and offset of the recordings were faded to avoid sudden amplitude increase and decrease during the experimental trials. We standardized the duration of playback stimuli to ensure comparability across vocalization types while preserving the natural structure of call sequences. Rather than artificially truncating bouts, we retained complete call sequences to avoid producing unnatural stimuli that might be perceived differently by the subjects. On average, playbacks of grunts lasted 9.5 ± 5 s and contained 21 calls, screams lasted 8 ± 5 s with 20 calls, alarm calls lasted 13 ± 2 s with 17 calls, and contact calls lasted 5 ± 3 s, strictly including three calls. These parameters ensured that playback stimuli were salient for the subjects while minimizing uncontrolled variability.

Once edited, the calls were stored as '.wav' files in a smartphone. All the calls were played back at a similar amplitude through a BOSE speaker amplified by a Lepy LP-2024A+ digital amplifier powered with a X-MOOVE portable power bank able to generate a power supply of 220V/50Hz. The playbacks were run remotely with a remote connected to the smartphone by Bluetooth. The appropriate amplitude of the calls was judged in the forest beforehand without macaques present, by an experienced human listener situated 10 m away from the speaker to match naturally-occurring vocalizations.

2.4. Experimental protocols

2.4.1. Playback of contact calls and python model presentation:

Playback trials were conducted between 8 am and 4 pm to avoid disrupting the sleep patterns of the macaques [75]. Three experimenters worked in parallel: Experimenter 1 ensured experimental

207 conditions were met, recorded the subject's thermal reactions with a FLIR T560 infrared thermographic
208 video camera (temperature sensitivity: 0.04 °C; resolution: 640×480 pixels; sampling rate: 30 fps) and
209 documented behavioral responses via vocal comments and a GoPro (Hero7 White) affixed to the
210 shoulder. Vocalizations, particularly alarm calls, were captured using a Marantz PMD660 audio recorder
211 carried by Experimenter 1. Experimenter 2 managed the playback equipment, while Experimenter 3
212 handled the python model.
213
214 **Step 1 - Recording Baseline and Setting Up (Pre-Phase):** When a subject was isolated from the rest of
215 the group, out of sight and at least 20 m away from any groupmate, and environmental conditions
216 permitted thermal measurements (no direct sunlight, dense vegetation, wind, or rain – see below),
217 Experimenter 1 recorded the baseline temperature for at least 30 seconds before the exposure to the
218 stimulus. Meanwhile, Experimenters 2 and 3 discreetly positioned themselves behind big trees, at
219 approximately 10 m from the subject (mean = 11 m, min = 8 m, max = 16 m) and a few meters from one
220 another. Experimenter 2 ensured playback volume and speaker orientation were correct [76].
221
222 **Step 2 - playback of contact calls (priming phase):** If all the conditions were still met after the recording
223 of the baseline, Experimenter 1 played the contact calls through the remotely connected speaker.
224 Experimenter 1 recorded the thermal and behavioral reaction of the subject. This step was initially made
225 to simulate the presence of a specific audience in proximity of the python model for the purpose of
226 other research questions (manuscript in *prep*) but the emotional and the behavioral responses of
227 subjects to contact calls was analyzed in detail in this study.
228
229 **Step 3 - presentation of the python model (post-phase and recovery):** A few seconds after the playback
230 began, Experimenter 3, still out of sight of the subject, presented the model by slowly moving the python
231 head to be visible above a root of the tree and manually making small movements to catch the attention
232 of the subject. The subjects detected the python model within 1 s to 1 min post-playback (average: 22

s). Detection was indicated by a notable shift in the subject's behavior, including heightened body tension and focused gaze in the direction of the python location. The model was concealed promptly after the first alarm call to standardize exposure among the different subjects, therefore minimizing the variation in the production of the alarm calls potentially caused by a longer exposure. Once out of sight behind the root, the python model was quickly hidden in the backpack of Experimenter 3 and Experimenters 2 and 3 discreetly left the area to not be seen by the subject or other individuals approaching to inspect the site. Experimenter 1 recorded the reaction of the subject during at least one minute (post-phase) and up to 3 minutes (recovery) after the subject had detected the python.

A total of 14 successful model presentations were conducted on nine subjects between July 13 and December 9, 2021. On average, 11.5 days separated two trials made on different subjects (min: 1, max: 34). Once, we conducted two experiments the same day on two different subjects but 1 h 14 min separated the two trials and the group had travelled 130 m away from the location of the previous presentation. This situation naturally occurred with real pythons during the period of data collection and, thus should not have disturbed the subjects more than might naturally occur on rare occasions. On average, 47 days separated two trials made on the same subject (min: 7, max: 104). All females reacted to the python model by uttering a series of alarm calls and inspecting or monitoring the python area. In most cases (except for two trials made on the same subject), at least one group-mate approached and monitored the area, but none of them produced any alarm calls, presumably because the model python was by then out of sight.

2.4.2. Playback of grunts, screams and alarm calls

The playbacks of grunts, screams and alarm calls were carried out following the same experimental protocol as for the playbacks of contact calls (refer to *steps 1 and 2* of the previous section). The grunts and the screams were played back at a distance of 10 ± 5 m (grunts: mean = 8 m, min = 5 m, max = 15 m; screams: mean = 11 m, min = 5 m, max = 15 m) and the alarm calls at a distance of 15 ± 5 m (mean =

16 m, min = 10 m, max = 20 m). To facilitate the video-coding of the subjects looking direction and to be sure that they reacted to the playback and not to events occurring in the group, the stimulus was played from a direction roughly perpendicular to the group's direction and from an angle of at least 35° to the subject's body orientation (as previously performed by Micheletta et al. [71]; see **Fig. S3**).

The playback trials of grunts and screams were conducted from March 22 to November 12, 2021 on 12 subjects, resulting in 23 and 24 successful experiments respectively (**Table 1S**). Despite the frequent occurrence of these vocalizations in crested macaques, we limited experiments to one per day per subject, with one to two experiments conducted daily for the same vocalization type. On average 41 days (min: 16, max: 91) separated two trials made on the same subject for grunts and 54 days (min: 13, max: 160) for screams.

Alarm call playback trials occurred between October 15 and December 9, 2021, with 13 successful trials on eight subjects (**Table 1S**). A maximum of two trials were conducted per day on different subjects, ensuring a minimum travelling distance of 150 m and a 1-hour gap between experiments (as previously performed by Micheletta et al. [71]). On average, 26 days (min: 12, max: 47) separated trials on the same subject.

To control for habituation effects, we randomized the order of stimulus presentations across trials. However, because alarm call recordings were made during these model presentations, playbacks were conducted more often later in the data collection period. Despite this, experimental trials remained balanced across subjects, and no evidence suggested that responses diminished over time.

2.5. Thermal measurements and extraction

In many IRT studies on primates, only the nasal region is examined, although other regions of interest (ROIs) can provide valuable insights [38,61,77]. Research has shown differential temperature responses

285 across ROIs to the same stimulus [37,38,61]. Thus, to capture nuanced emotional changes, we extracted
286 temperatures from three ROIs: nose tip, nose bridge, and upper lip (**Fig. 2**). Dividing the nasal region
287 helps reduce bias from respiration, especially during emotionally charged situations. The upper lip,
288 unaffected by breathing and highly vascularized [78], is expected to offer a more reliable measure of
289 arousal, particularly in fearful contexts [37,42].
290
291 We extracted thermal data using a customized MATLAB® (2018b) script developed by Ermatinger et al.
292 [40]. Thermal data extraction was performed by a single observer who was not blind to the experimental
293 condition. To minimize bias, we applied a standardized frame selection protocol, ensuring consistency
294 across all trials.
295
296 We manually positioned an ellipse over the ROIs on the three best frames (see below) for each one-
297 second interval and automatically exported the temperature data to a spreadsheet for further analysis.
298 The mean temperature of the three frames was calculated for each one-second interval to reduce
299 measurement bias. We extracted the minimum temperature for nasal ROIs and the maximum
300 temperature for the upper lip to obtain more accurate measurements, as the pixels surrounding these
301 ROIs are often hotter and colder respectively.
302
303 Appropriate frames were selected based on the following criteria: (1) The subject's face was in the focus
304 of the camera; (2) The subject's head was oriented straight towards the camera (maximum tilt angle of
305 approximately 50°); The ROIs were (3) not covered by vegetation or (4) not blurred due to subject or
306 camera movements (i.e., we did not select frames in which the subject was walking, running or climbing).
307 When more than three frames meet all the criteria for a given one-second interval, the first three frames
308 were selected. A respective interval was omitted if it did not contain useful frames. Over the 75 trials,
309 we extracted a total of 19,735 thermal data points (nose tip: 6,613; nose bridge: 6,634; upper lip: 6,488).

Since our focus was on temperature changes relative to baseline rather than absolute values, we calculated a centered temperature measure for each ROI. For each trial, we subtracted the mean baseline temperature (30 seconds before stimulus onset) from all one-second interval measurements. This approach mitigates inter-individual differences, minimizes potential environmental bias, and captures both the direction (increase or decrease) and amplitude of facial temperature changes from the baseline, providing insights into the valence and intensity of emotional arousal.

2.6. Protocol to minimize the effects of environmental and physical factors on thermal measurements

The emissivity (i.e., the effectiveness of a material in emitting energy as thermal radiation) was constantly set on 0.98, the emissivity value detected for human skin [79] and used in previous thermal imaging studies on primates, including black-skinned primates (e.g., [42,61]).

We did not record thermal data during rainy and windy episodes, or when subjects were directly exposed to sunlight to avoid any measurement bias that these factors may cause [57] (**Fig. 3**). Subjects were strictly observed under dense canopy cover, shielding them from direct solar radiation. Following rainfall, data collection resumed only after an hour to allow vegetation and monkeys to dry and for atmospheric humidity to return to normal levels. We noted if measurements were taken while vegetation was dry or wet (but not dripping) and recorded weather conditions (sunny or cloudy) to control for their potential impact on environmental radiation that could affect thermal readings. Additionally, we categorized the environment based on vegetation characteristics (i.e., sea front, open forest, dense forest) for later analysis.

The atmospheric temperature and relative humidity were logged in the IRT camera before each experiment, and recorded throughout the trials using a separate automatic logger (OMEGA OM-HL-SP-TH), with measurements taken every 3 s. Although we did not expect important fluctuations of these parameters due to the brevity of the trials (under 5 min), this interval allowed us to monitor any major fluctuations. The visual inspection of these data revealed no unexpected variation in atmospheric

temperature or humidity within individual trials. This stability allowed us to proceed without excluding any trials based on these factors. The average air temperature during trials was 28.3°C (ranging from 23.6°C to 31.8°C), and the relative humidity averaged 86.8% (ranging from 69.6% to 99.3%).

We did not record thermal data on subjects located arboreally more than 3 m above the ground due to the obtuse angle of view, especially when the monkeys looked upward. During data extraction, we also discarded frames where the monkey was looking considerably up or down, even if the ROIs were visible, to control for vertical axis distortion. Additionally, the posture of the monkey (sitting, standing still, or moving) was noted to control for small movements and the natural angle differences between sitting (head slightly upward) and standing postures (head straighter) as illustrated in **Fig. S4**.

During trials, Experimenter 1 maintained a consistent distance between the thermal sensor and the subject to minimize measurement bias, estimating distance by eye after a week of self-training with a laser device. For each one-second interval, a finer distance assessment was estimated using an indirect method. In addition to experimental trials, 165 IRT images of adult females were taken at various distances and angles, with precise laser measurements of the distance. The pictures were taken when the target was looking in the opposite direction and the laser was pointed on the shoulder of the monkey to avoid any risk of glare from the laser. The images were processed using FLIR® software ResearchIR MAX (4.44.11.35), where an ellipse was drawn around the monkey's body (excluding the head and the limbs) to obtain the number of pixels. This data was used to create a reference curve correlating pixel count to distance, allowing more accurate distance estimations during video data extraction using the same method.

Finally, to ensure that distance did not affect the thermal measurements, while accounting for environmental factors, we conducted a control experiment after each trial. We took IRT images of a flat target wrapped in black insulation tape of known emissivity ($\varepsilon = 1$), positioned at about 35 cm above the ground (the average height of a female macaque) and at distances of 1 m, 2 m, and 3 m of the camera,

362 measured with a laser device (**Fig. S5**). The central pixel temperature was extracted, and we statistically
363 tested whether distance impacted the temperature readings (see below). **Table S2** provides a summary
364 of the environmental and physical factors (and clear definitions where applicable) considered in this
365 study to control for their potential effects on the thermal measurements.
366
367 *2.7. Behavioral expressions of arousal and anxiety levels: extraction and coding*
368 Behavioral data were extracted from the GoPro videos using BORIS® v7.12.2 allowing frame-by-frame
369 analysis. Prior to and throughout each experimental trial (at least 30 seconds before stimulus onset and
370 one minute after), we recorded the occurrence of SDBs, i.e., scratching, self-grooming, yawning and body
371 shaking (as defined in **Table S3**), and the subjects' physical activity. Although these SDBs have been
372 established as indicators of anxiety [44-51], some studies suggest that specific SDBs may be stronger
373 indicators of arousal (e.g., scratching has been more consistently linked to heightened emotional states
374 in primates [45]). However, we did not differentiate their relative importance in our statistical analyses
375 due to the limited sample size, especially in regards to the rarer behaviors such as yawning and body
376 shaking in the field. Physical activity was measured by counting transitions between postures or spatial
377 displacements, such as sitting, standing, walking, running, and climbing up and down a tree (e.g., the
378 sequence 'walking-standing-sitting-walking' counts as three transitions). SDBs and physical activity rates
379 per minute were calculated separately for the pre- and post-phases, with rate changes determined by
380 subtracting pre-phase from post-phase rates for each trial.
381
382 *2.8. Statistical analysis*
383 All statistical analyses were conducted in R 4.3.1 [80], and significance was set at $\alpha = 0.05$. We used
384 Linear Mixed-effects Models (LMM; function '*lmer*' of the '*lme4*' package) [72] to evaluate the effects of
385 multiple covariates on facial temperature variation and behavioral responses, and Cumulative Link
386 Mixed Models for ordinal variables (CLMM; function '*clmm*' of the '*ordinal*' package) [81] when
387 appropriate. Focal ID was entered as a random effect in all models to account for repeated measures on

388 the same individuals. Model assumptions were assessed by visually inspecting residual plots for
389 normality and homogeneity. No violations of these assumptions were detected. Statistics of the models
390 were given by the function '*summary*' and '*anova*' and significance of the LMMs were obtained with the
391 function '*Anova*' of the '*car*' package [82]. We computed post-hoc pairwise contrasts when necessary,
392 using the '*emmeans*' package [83], adjusted with the Tukey method for multiple comparisons [84]. In
393 this study, model selection was not performed as we aimed to assess the combined effects of all
394 covariates rather than isolate individual predictors. Given the limited sample size, we did not include
395 first-order interactions between covariate fixed effects to avoid overfitting and ensure robust parameter
396 estimates. The same post hoc procedures as in previous studies were applied to maintain consistency
397 and comparability with existing literature (e.g., [54]).
398
399 We first investigated whether the environmental and physical factors had an impact on the thermal
400 measurements obtained following our protocol. To this end, we analyzed one-second interval thermal
401 data recorded during the pre-phase. Outliers (data points outside 2 standard deviations) were removed
402 using the '*boxplot.stats*' R function to avoid bias. For each ROI, we built a LMM with thermal
403 measurements centered around the mean trial temperature (baseline) as the dependent variable. Fixed
404 covariates included atmospheric temperature, relative humidity, camera-to-subject distance, subject
405 posture, weather, vegetation type, and vegetation dryness.
406
407 The distance between the thermal sensor of the camera and the subject has been shown to have
408 significant effects on the thermal measurements in previous studies (e.g., [60]). We thus made a control
409 (as described above; **Fig. S5**), by measuring the temperature of a black target placed at exactly 1 m, 2 m
410 and 3 m from the camera just after each experimental trial to have the same environmental factors as
411 during the trials. To determine if the distance impacted the IRT measurements, we performed a
412 '*between-subject factor'* ANOVA (for paired data) using the '*aov*' R function fitted with the temperature

of the black target as dependent variable, the distance as explicative variable, and the sample/distance as the 'error' component of the function.

To investigate factors linked to changes in facial skin temperature in response to different stimuli, we built a series of LMMs. The one-second interval temperature changes from the three ROIs were included as dependent variables in separate analyses. Fixed effects included the experimental condition (i.e., the five stimuli), changes in physical activity and SDB rates between the pre- and post-phases, while time and trial number were added as covariates to account for repeated and time-series measurements. The subjects' reproductive status and rank were included to account for potential influences on emotional responses, though these results will be discussed elsewhere (manuscript *in prep.*). To further corroborate the findings of this analysis with an even more conservative approach, and to assess the robustness of our results, we performed similar statistical analysis on a reduced dataset in which the one-second interval temperature changes were reduced to the mean temperature change recorded during 1 min post-stimulus.

To determine whether the variation of the facial skin temperature was a good predictor of the strength of emotional arousal elicited by different stimuli, we conducted a series of CLMMs. The stimuli conditions were ordered by assumed emotional arousal intensity, from lowest to highest (see **Fig. 1**). The one-second interval temperature change of (a) the nose tip, (b) the nose bridge, or (c) the upper lip was included as the dependent variables, while the condition was set as the fixed effect. Time and trial number were added as covariate fixed effects to account for repeated and time-series measurements.
Finally, to help in the interpretation of the results regarding the link between facial skin temperature variation and behavioral indicators of arousal and anxiety levels, we investigated whether the experimental condition influenced changes in activity and SDB rates. Two LMMs were fitted with (a) the change in activity rate, and (b) the change of SDB rate between the pre- and post-phase as the dependent variables, with 'condition' as a fixed effect.

439

440 *2.9. Ethical statement*

441 Ethical approval was provided by the School of Anthropology and Conservation of the University of Kent.
442 Permission to conduct the research was granted by the Indonesian Ministry of Research and Technology
443 (permit numbers: 463/E5/E5.4/SIP/2019 and 58/E5/E5.4/SIP.EXT/2020), the Indonesian Ministry of
444 Environment and Forestry (Permit number: 03/BKSDA.Sulut/TU/SIMAKSI/1/2020) and the Macaca Nigra
445 Project. The study animals were habituated to researchers, and all data were collected using non-
446 invasive, contact free methods only. To minimize disturbance, predator model presentations and alarm
447 call playbacks were conducted at a frequency comparable to natural encounters with real pythons.
448 Moreover, predator model presentations lasted only a few seconds, with the model being removed
449 immediately after the first alarm call was uttered, minimizing the exposure duration. The methodology
450 aligns with previous research using brief exposure to alarm calls or predator stimuli in a way that is
451 expected to cause the sorts of short-term stress responses that the subjects regularly experience under
452 natural conditions, but not long-term behavioral disturbances [71, 85].

453

454 **3. Results**

455 *3.1. Thermal measurements and environmental/physical factors*

456 The first aim of this study was to determine the potential effects of environmental and physical factors
457 on the thermal measurements, despite the strict protocol that we followed during data collection and
458 extraction to avoid measurement biases. The results of the LMM analysis showed that none of the
459 environmental or physical factors had a significant effect on the thermal measurements made on the
460 macaques before the stimuli exposure (**Table 1**).

461

462 **Table 1.** Results of the LMMs including the environmental and physical factors and performed on the temperature
463 deviation from the baseline recorded during the pre-phase of the three ROIs.

	Nose tip (n = 594)				
Random effect: Subject ID (N =14)	**Min**	**Max**	**Mean**	**SD**	
	-0.345	0.332	-0.004	0.131	
Fixed effect	**Sum ²**	**Mean ²**	***F***	***X²*(df)**	***P***
Posture *(3 levels)*	0.013	0.006	0.38	0.80 (2)	0.66
Distance *(m)*	0.012	0.012	0.74	1.39 (1)	0.23
Air temperature *(°C)*	0.009	0.009	0.56	0.41 (1)	0.51
Relative humidity *(%)*	0.00005	0.00005	0.003	0.04 (1)	0.82
Weather *(2 levels)*	0.003	0.003	0.21	0.01 (1)	0.91
Environment *(3 levels)*	0.045	0.022	1.33	2.42 (2)	0.29
Vegetation dryness *(2 levels)*	0.0003	0.0003	0.01	0.01 (1)	0.89

	Nose bridge (n = 623)				
Random effect: Subject ID (N =14)	**Min**	**Max**	**Mean**	**SD**	
	-0.288	0.289	0.0007	0.110	
Fixed effect	**Sum ²**	**Mean ²**	***F***	***X²*(df)**	***P***
Posture *(3 levels)*	0.012	0.006	0.50	1.04 (2)	0.59
Distance *(m)*	0.021	0.021	1.75	1.68 (1)	0.19
Air temperature *(°C)*	0.002	0.002	0.19	0.10 (1)	0.74
Relative humidity *(%)*	0.00004	0.00004	0.003	0.001 (1)	0.96
Weather *(2 levels)*	0.001	0.001	0.10	0.08 (1)	0.77
Environment *(3 levels)*	0.0001	0.00006	0.005	0.01 (2)	0.99
Vegetation dryness *(2 levels)*	0.0004	0.0004	0.03	0.03 (1)	0.84

	Upper lip (n = 614)				
Random effect: Subject ID (N =14)	**Min**	**Max**	**Mean**	**SD**	
	-0.249	0.247	0.0014	0.090	
Fixed effect	**Sum ²**	**Mean ²**	***F***	***X²*(df)**	***P***
Posture *(3 levels)*	0.005	0.002	0.32	0.29 (2)	0.86
Distance *(m)*	0.001	0.001	0.19	0.03 (1)	0.85
Air temperature *(°C)*	0.003	0.003	0.46	0.45 (1)	0.50
Relative humidity *(%)*	0.0003	0.0003	0.04	0.08 (1)	0.76
Weather *(2 levels)*	0.004	0.004	0.53	0.33 (1)	0.56
Environment *(3 levels)*	0.004	0.002	0.26	0.51 (2)	0.77
Vegetation dryness *(2 levels)*	0.0004	0.0004	0.05	0.05 (1)	0.81

464

465

466 In contrast to previous studies (e.g. [60,54]), the distance and the air temperature were not significantly
467 associated with the temperature variation of the three ROIs in the absence of any stimuli (Distance -
468 Nose tip: t = 1.18, P = 0.238; Nose bridge: t = 1.29, P = 0.194; Upper lip: t = -0.18, P = 0.851; Air

Temperature – Nose tip: t = -0.64, P = 0.518; Nose bridge: t = -0.33, P = 0.741; Upper lip: t = -0.67, P = 0.502).

Similarly, the analyses performed on the temperature of the black target at different distances showed that the temperature measurements were not impacted by the distance (ANOVA: n_{Temp} = 221, n_{sample} = 74, $F(1)$ = 1.15, P = 0.319).

3.2. Factors affecting facial skin temperature variation

LMM analysis performed on the temperature changes of the three facial ROIs recorded every second post-stimuli indicated that the condition, the changes of activity rate, and, to a lesser extent, the expression of SDBs were significantly associated with facial skin temperature variation (**Table 2**). To validate these findings, we performed a more conservative analysis using the mean temperature change over the one-minute post-stimulus period. This analysis confirmed the significant association with the condition across all ROIs, but the associations with behavioural expressions of emotional arousal (activity rates) and anxiety (SDB rates) became non-significant (**Table 2**). Further analysis is presented in the following sections.

Table 2. Results of LMM analysis performed on the temperature changes (TC) of the three ROIs (nose tip, nose bridge and upper lip) post-stimuli. '*FULL TC*' = analysis performed on the one-second interval measurements recorded post-stimuli; '*MEAN TC*' = analysis performed on the mean temperature changes recorded post-stimuli. Bold values indicate $P < 0.05$.

Random effect: Subject ID (N = 14)		**Nose tip**			
		Full TC (n = 1198)		*Mean TC (n = 81)*	
Fixed effect		***X²*(df)**	***P***	***X²*(df)**	***P***
f (5)	Condition	134.55 (4)	**< 0.001**	18.00 (4)	**0.001**
c (rate/min)	SDBs change	2.026 (1)	0.154	0.003 (1)	0.954
c (rate/min)	Activity change	35.857 (1)	**< 0.001**	0.41 (1)	0.521

Random effect: Subject ID (N = 14)		**Nose bridge**			
		Full TC (n = 1193)		*Mean TC (n = 81)*	
Fixed effect		***X²*(df)**	***P***	***X²*(df)**	***P***
f (5)	Condition	63.22 (4)	**< 0.001**	11.50 (4)	**0.021**
c (rate/min)	SDBs change	11.75 (1)	**< 0.001**	0.53 (1)	0.464
c (rate/min)	Activity change	61.81 (1)	**< 0.001**	1.33 (1)	0.247

Random effect: Subject ID (N = 14)		**Upper lip**			
		Full TC (n = 1182)		*Mean TC (n = 81)*	
Fixed effect		***X²*(df)**	***P***	***X²*(df)**	***P***
f (5)	Condition	93.65 (4)	**< 0.001**	18.45 (4)	**0.001**
c (rate/min)	SDBs change	0.25 (1)	0.616	1.02 (1)	0.310
c (rate/min)	Activity change	51.30 (1)	**< 0.001**	1.08 (1)	0.298

3.2.1. Facial skin temperature as a marker of emotional arousal

The second aim of this study was to determine if the facial skin temperature variation was a good indicator of the strength of emotional arousal. The LMM analysis above showed that the condition had a significant effect on the temperature change (TC) of the three ROIs (LMM '*Full TC*' - Nose tip: $F = 77.60$, $P < 0.001$; Nose bridge: $F = 58.73$, $P = < 0.001$; Upper lip: $F = 73.34$, $P < 0.001$; **Table 2** and **Fig. 4**).

To corroborate this finding, we conducted a series of CLMMs with the stimuli ordered following their assumed emotional arousal intensity (**see Fig. 1**). This analysis showed that the temperature changes of three ROIs were good indicators of the strength of context arousal. The greater the strength of arousal,

the greater the temperature decrease for the three ROIs (CLMM; Nose tip: n = 1198, Z = -12.77, P < 0.001; Nose bridge: n = 1193, Z = -10.67, P < 0.001; Upper lip: n = 1182, Z = -11.84, P < 0.001; **Fig. 4**).

3.2.2. Facial skin temperature variation and behavioral expressions of arousal and anxiety

The third aim of this study was to determine whether the variation of facial temperature was linked to the change of activity and SDB rates in response to different stimuli as a behavioral indicator of arousal and anxiety. The LMM analysis above showed that the temperature change of the three ROIs was negatively associated with the change of activity rate in response to the different stimuli (**Table 2** – '*Full TC*'). Greater negative changes of the facial temperature were associated with a greater increase of activity rate (LMM '*Full TC*' – Nose tip: t = -5.98, P < 0.001; Nose bridge: t = -7.86, P = < 0.001; Upper lip: t = -7.16, P < 0.001; **Fig. 5**).

This analysis also showed that only the temperature change of the nose bridge was negatively associated with the change of SDB rate in response to the different stimuli (**Table 2** – '*Full TC*'). Greater negative changes of the nose bridge temperature were associated with a greater increase of the SDB rate (LMM '*Full TC*' – Nose bridge: t = -3.42, P < 0.001; **Fig. 6**).

To better interpret the results, we investigated whether the experimental condition influenced changes in activity and SDB rates. The LMM analysis revealed that activity rate one minute post-stimulus was significantly affected by the condition (LMM: n_{obs} = 81, $N_{subject}$ = 14, $X^2(4)$ = 26.77, P < 0.001). Pairwise comparisons indicated that predator detection significantly increased activity rate compared to all other stimuli, except when hearing alarm calls (**Fig. 7a.**). However, the condition did not explain the variation of the change of SDB rates between the pre- and the post-phase (LMM: n_{obs} = 81, $N_{subject}$ = 14, $X^2(4)$ = 5.873, P = 0.2; **Fig. 7b**).

4. Discussion

The results of this study demonstrate that (1) thermal measurements are not biased by environmental or physical factors; (2) facial skin temperature variation reliably indicates the strength of emotional arousal; and (3) temperature in at least one of the three ROIs correlates with behavioral indicators of arousal and anxiety, such as physical activity and SDBs.

4.1. Thermal measurements and environmental/physical factors

The results first demonstrate that none of the environmental or physical factors examined here (see **Table 1**) significantly affected thermal measurements. Particularly, the subject-to-camera distance did not affect the measurements made on the macaque faces, which was corroborated by our results on the black target. Overall, our results indicate that the precautions taken in both fieldwork and thermal data extraction effectively minimized any potential biases, ensuring accurate measurements of facial temperature. These findings demonstrate that IRT is a reliable tool for studying surface skin temperature variation in wild, habituated, medium-sized, terrestrial animals, even in complex environments like tropical rainforests. While applicable to other mammals and birds, careful planning is essential. In this study, thermal data were collected when the macaques were on the ground and under dense canopy cover, minimizing distortions from sunlight and viewing angles [57,59]. Additionally, the primates studied here are medium-sized and habituated, ensuring optimal subject-to-camera distance and unobstructed views of ROIs [60] without behavioral disturbances. This approach may thus pose challenges for small, arboreal, or unhabituated species, especially in open environments, requiring tailored solutions.

4.2. Variation in facial skin temperature as a reliable marker of emotional arousal

Second, we found good evidence that variation in facial temperature is a reliable indicator of the strength of emotional arousal in crested macaques. Indeed, when exposed to different stimuli eliciting different levels of arousal (from a calming effect to a very stressful effect), the facial temperature of the macaques decreased significantly when arousal increased. The two independent analyses made on the

553 one-second interval measurements and the mean temperature both yielded statistically significant
554 results, strengthening the robustness and the accuracy of this finding. Thus, this physiological marker of
555 emotional arousal provides a reliable tool with which to investigate the role of emotions in the daily life
556 of wild primates when they cope with different social and environmental stimuli in their natural habitat.
557 In this study, we focused on the overall temperature variation of key facial ROIs to assess emotional
558 arousal in response to social vocalizations and predator encounters. While a finer-grained analysis of
559 temperature fluctuations across different ROIs and stimuli could provide additional insights into the
560 dynamics of emotional arousal and their effects on responsiveness and call production, such analyses
561 are beyond the scope of this study and will be explored in future work (manuscript *in prep.*).
562
563 The results were less clear regarding the valence of the stimuli. We did not observe distinct patterns of
564 facial temperature variation in response to affiliative grunts, which are expected to have a calming,
565 positive effect [74, 86]. Contrary to our expectations, temperature increases compared to baseline were
566 not significant, and changes in the nose tip and upper lip temperatures were not statistically different
567 from those observed in response to neutral contact calls. This might have been due to inter-individual
568 variation in subject perception of the stimuli, with some interpreted grunts as affiliative, and so calming,
569 while others found them stressful. A more detailed analysis, considering factors such as the caller's
570 identity, dominance rank, and the strength of the relationship between the subject and caller, is needed.
571 Additionally, the absence of significant thermal responses to positive stimuli might be attributed to the
572 specific nature of the affiliative grunts used in this study. While these vocalizations, coupled with
573 affiliative facial expressions, are common signals promoting socio-positive interactions in crested
574 macaques [74], their acoustic properties may not be inherently arousing enough to elicit measurable
575 changes in peripheral temperature. Prior research indicates that tactile affiliative interactions, such as
576 grooming, tend to evoke stronger physiological responses, including increased nasal skin temperature
577 [31], likely due to their direct and rewarding nature. To address this limitation, future studies should
578 explore the emotional impact of a broader array of positively valanced stimuli, including play calls or

social cues linked to resource acquisition, such as the discovery of valuable food [40, 41, 87]. These more intense stimuli may provide a more comprehensive understanding of the role of positive emotional arousal in primate communication and behavior. Expanding the range of stimuli will help clarify whether the lack of response observed here is methodological or reflective of the underlying emotional processes.

4.3. Facial skin temperature variation and behavioral expressions of arousal and anxiety

This study shows that facial temperature variation is linked to changes in activity rate, with greater decreases in temperature corresponding to higher increases in activity. While it could be argued that increased physical activity, through enhanced breathing and air intake, might cool the facial skin, previous studies (e.g., [40]) suggest activity alone cannot fully explain temperature variation during emotional arousal. Furthermore, the upper lip, unaffected by breathing, exhibited similar temperature patterns as nasal regions, supporting the idea that these temperature changes reflect emotional responses rather than just physical activity.

Moreover, the change in activity rate was influenced by the type of stimuli, particularly by the level of emotional arousal they elicited. Stimuli associated with higher emotional arousal, such as alarm calls and predator detection, triggered a greater increase in activity rate, while contact calls had a calming effect, though differences with grunts and screams were not statistically significant. These findings suggest that changes in activity rate are a reliable indicator of emotional arousal. The combined analysis of facial skin temperature and activity rate provides complementary measures of emotional arousal, reinforcing the conclusion that facial temperature variation is a reliable indicator of emotional excitement in crested macaques, and likely in other mammal and bird species.

Regarding SDBs, which are considered reliable indicators of uncertainty and anxiety in animals (e.g., [45]), our results revealed that greater negative changes in nose bridge temperature were linked to

increased SDB rates, suggesting that nose bridge temperature might indicate anxiety levels. However, this finding should be interpreted cautiously, as the same pattern was not observed in other ROIs. Additionally, SDB rate changes were not influenced by the nature of the stimuli, indicating that emotional arousal and anxiety-related SDBs may be independent phenomena depending on context. The high inter-individual variability, particularly for affiliative grunts and screams, suggests these stimuli might be perceived differently across subjects, highlighting potential ambiguity in their emotional significance. Future research should explore these ambiguities by studying naturally occurring interactions and incorporating more social variables. It is also possible that additional factors, such as locomotion and social context, impact the relationship between temperature changes and SDBs. While our study carefully controlled for sudden spatial displacements by excluding frames recorded during significant movement (i.e., when the subject was walking, running or climbing), more subtle changes in posture or facial movement may still influence facial skin temperature measurements. Additionally, the social context of the trials, including the proximity of other groupmates, may contribute to variation in SDB expression despite the fact that all trials were conducted when subjects were isolated. Future research would benefit from the inclusion of control groups and a broader set of behavioral and physiological stress indicators to disentangle the interplay between physiological responses and social or physical factors.

4.4. Individual variability and future directions

Our study focuses on the general emotional responses of wild crested macaques to social signals and predator presence, but it is highly plausible that individual differences may influence these responses. Primate behavior and physiology are often shaped by social rank, relationship strength, personality, and previous experiences [41, 71, 88]. These factors may lead to variation in thermal and behavioral responses to social stimuli and the presence of predator. While the primary aim of the current paper was to establish the viability of IRT in a naturalistic setting, a detailed exploration of how social attributes, such as dominance rank and grooming preferences, modulate emotional and behavioral

631 responses is important part of the broader research using IRT to understand macaque behavior [89].
632 Future research should continue exploring how temperament, past experiences, and hierarchical
633 positions interact with social signaling contexts to shape emotional responses, providing a more
634 nuanced understanding of emotional processes in wild primates.
635
636 *4.5. Conclusion*
637 To summarize, the results presented here validate the use of infrared thermography to study surface
638 skin temperature variation in wild crested macaques in their natural habitat, despite harsh and mostly
639 uncontrollable conditions. In this regard, we have provided measures to minimize the risk of causing
640 biases in the thermal measurements that can be useful for researchers eager to add an emotional
641 component to their research on naturally occurring interactions in wild animals. This study also
642 demonstrates that the variation of the facial skin temperature in response to different stimuli is a
643 reliable indicator of the strength of emotional arousal and could also inform us about the anxiety level
644 experienced by the monkeys. Overall, this validation confirms the robustness of the protocol and
645 supports its broader application in studying emotional responses and their behavioral significance in
646 wild animals. Not only is this method valuable for understanding the emotional states of wild non-
647 human primates, but it also holds potential for terrestrial mammals and even birds, whose physiological
648 responses to stimuli could similarly be monitored through IRT. The versatility and non-invasive nature
649 of this approach make it an invaluable tool for ethological research in diverse species and environments.
650 Beyond its relevance to fundamental research on primate behavior, our study highlights the potential
651 of IRT as a non-invasive tool for conservation and welfare monitoring. IRT could provide insights into the
652 emotional states of individuals undergoing rehabilitation, aiding in assessments of their readiness for
653 reintroduction into the wild. Additionally, this method could be applied in long-term conservation
654 programs to monitor the physiological well-being of wild populations in response to environmental
655 changes or human disturbances, ultimately contributing to more effective conservation strategies.
656

Acknowledgments: Funding was provided by the School of Anthropology and Conservation of the University of Kent through a Vice Chancellor's Research Scholarship. We thank the Indonesian Ministry of Research and Technology, the Directorate General of Forest Protection and Nature Conservation, and the Department for the Conservation of Natural Resources for permission to conduct research in the Tangkoko Reserve. We thank the directors of the Macaca Nigra Project (MNP) for granting access to the MNP research station and the MNP staff in Indonesia for their assistance in the field. We particularly thank the field assistants who helped in the experiments: Indra Kaletuang, Casa Kaunang, Joaquim Lambaihang and Meldy Tamengge. We warmly thank Ermatinger and colleagues for kindly sharing their MatLab script for extracting thermal data. Finally, we thank Dr. Chris Dunmore and Prof. Bonaventura Majolo for their constructive comments on earlier versions of this manuscript.

Data statement: Data used in this article will be publicly available under request to the corresponding author if accepted for publication.

References

[1] Speakman, J. R., & Ward, S. (1998). Infrared thermography: principles and applications. *Zoology*, *101*, 224-232.

[2] Kastberger, G., & Stachl, R. (2003). Infrared imaging technology and biological applications. *Behavior Research Methods, Instruments, & Computers*, *35*, 429-439. https://doi.org/10.3758/BF03195520

[3] Vianna, D. M., & Carrive, P. (2005). Changes in cutaneous and body temperature during and after conditioned fear to context in the rat. *European Journal of Neuroscience*, *21*(9), 2505-2512. https://doi.org/10.1111/j.1460-9568.2005.04073.x

[4] Oka, T., Oka, K., & Hori, T. (2001). Mechanisms and mediators of psychological stress-induced rise in core temperature. *Psychosomatic Medicine*, *63*(3), 476-486. https://doi.org/0033-3174/01/6303-0476

682 [5] Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biological*
683 *Psychology*, *84*(3), 394-421. https://doi.org/10.1016/j.biopsycho.2010.03.010
684 [6] Ioannou, S., Gallese, V., & Merla, A. (2014). Thermal infrared imaging in psychophysiology:
685 potentialities and limits. *Psychophysiology*, *51*(10), 951-963. https://doi.org/10.1111/psyp.12243
686 [7] Clay-Warner, J., & Robinson, D. T. (2015). Infrared thermography as a measure of emotion
687 response. *Emotion Review*, *7*(2), 157-162. https://doi.org/10.1177/1754073914554783
688 [8] Dai, F., Cogi, N. H., Heinzl, E. U. L., Dalla Costa, E., Canali, E., & Minero, M. (2015). Validation of a
689 fear test in sport horses using infrared thermography. *Journal of Veterinary Behavior*, *10*(2), 128-
690 136. https://doi.org/10.1016/j.jveb.2014.12.001
691 [9] Bartolomé, E., Sánchez, M. J., Molina, A., Schaefer, A. L., Cervantes, I., & Valera, M. (2013). Using
692 eye temperature and heart rate for stress assessment in young horses competing in jumping
693 competitions and its possible influence on sport performance. *Animal*, *7*(12), 2044-2053.
694 https://doi.org/10.1017/S1751731113001626
695 [10] Fenner, K., Yoon, S., White, P., Starling, M., & McGreevy, P. (2016). The effect of noseband
696 tightening on horses' behavior, eye temperature, and cardiac responses. *PloS One, 11*(5), e0154179.
697 https://doi.org/10.1371/journal.pone.0154179
698 [11] Cannas, S., Palestrini, C., Canali, E., Cozzi, B., Ferri, N., Heinzl, E., Minero, M., Chincarini, M.,
699 Vignola, G., & Dalla Costa, E. (2018). Thermography as a non-invasive measure of stress and fear of
700 humans in sheep. *Animals*, *8*(9), 146. https://doi.org/10.3390/ani8090146
701 [12] Herborn, K. A., Graves, J. L., Jerem, P., Evans, N. P., Nager, R., McCafferty, D. J., & McKeegan, D. E.
702 (2015). Skin temperature reveals the intensity of acute stress. *Physiology & Behavior*, *152*, 225-230.
703 https://doi.org/10.1016/j.physbeh.2015.09.032
704 [13] Stewart, M., Webster, J. R., Verkerk, G. A., Schaefer, A. L., Colyn, J. J., & Stafford, K. J. (2007). Non-
705 invasive measurement of stress in dairy cows using infrared thermography. *Physiology &*
706 *Behavior*, *92*(3), 520-525. https://doi.org/10.1016/j.physbeh.2007.04.034

[14] Stewart, M., Stafford, K. J., Dowling, S. K., Schaefer, A. L., & Webster, J. R. (2008). Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. *Physiology & Behavior*, *93*(4-5), 789-797. https://doi.org/10.1016/j.physbeh.2007.11.044

[15] McManus, C., Tanure, C. B., Peripolli, V., Seixas, L., Fischer, V., Gabbi, A. M., ... & Costa Jr, J. B. G. (2016). Infrared thermography in animal production: an overview. *Computers and Electronics in Agriculture*, *123*, 10-16. https://doi.org/10.1016/j.compag.2016.01.027

[16] Nääs, I. A., Garcia, R. G., & Caldara, F. R. (2020). Infrared thermal image for assessing animal health and welfare. *Journal of Animal Behaviour and Biometeorology*, *2*(3), 66-72. http://dx.doi.org/10.14269/2318-1265/jabb.v2n3p66-72

[17] Jerem, P., Herborn, K., McCafferty, D., McKeegan, D., & Nager, R. (2015). Thermal imaging to study stress non-invasively in unrestrained birds. *Journal of Visualized Experiments*, *105*, e53184. https://doi.org/10.3791/53184

[18] Jerem, P., & Romero, L. M. (2023). It's cool to be stressed: body surface temperatures track sympathetic nervous system activation during acute stress. *Journal of Experimental Biology*, *226*(20). https://doi.org/10.1242/jeb.246552

[19] Ouyang, J. Q., Macaballug, P., Chen, H., Hodach, K., Tang, S., & Francis, J. S. (2021). Infrared thermography is an effective, noninvasive measure of HPA activation. *Stress*, *24*(5), 584-589. https://doi.org/10.1080/10253890.2020.1868431

[20] Di Giovanni, J., Fawcett, T. W., Templeton, C. N., Raghav, S., & Boogert, N. J. (2022). Urban gulls show similar thermographic and behavioral responses to human shouting and conspecific alarm calls. *Frontiers in Ecology and Evolution*, *10*, 891985. https://doi.org/10.3389/fevo.2022.891985

[21] Jakubas, D., Wojczulanis-Jakubas, K., Grissot, A., Devogel, M., Cendrowska, M., & Chastel, O. (2022). Eye region surface temperature and corticosterone response to acute stress in a high-arctic seabird, the little auk. *Animals*, *12*(4), 499. https://doi.org/10.3390/ani12040499

731 [22] Wongsaengchan, C., McCafferty, D. J., Evans, N. P., McKeegan, D. E., & Nager, R. G. (2023). Body
732 surface temperature of rats reveals both magnitude and sex differences in the acute stress
733 response. *Physiology & Behavior*, *264*, 114138. https://doi.org/10.1016/j.physbeh.2023.114138
734 [23] Watanabe, S. (2015). Social factors modulate restraint stress induced hyperthermia in
735 mice. *Brain Research*, *1624*, 134-139. https://doi.org/10.1016/j.brainres.2015.07.019
736 [24] Watanabe, S. (2017). Social inequality aversion in mice: analysis with stress-induced
737 hyperthermia and behavioral preference. *Learning and Motivation*, *59*, 38-46.
738 https://doi.org/10.1016/j.lmot.2017.08.002
739 [25] Blenkuš, U., Gerós, A. F., Carpinteiro, C., Aguiar, P. D. C., Olsson, I. A. S., & Franco, N. H. (2022).
740 Non-Invasive assessment of mild stress-induced hyperthermia by infrared thermography in
741 laboratory mice. *Animals*, *12*(2), 177. https://doi.org/10.3390/ani12020177
742 [26] MacRae, A. M., McGreevy, P., Daros, R. R., & Fraser, D. (2021). Can eye surface temperature be
743 used to indicate a stress response in harbour seals (*Phoca vitulina*)?. *Animal Welfare*, *30*(1), 91-100.
744 https://doi.org/10.7120/09627286.30.1.091
745 [27] Ludwig, N., Gargano, M., Luzi, F., Carenzi, C., & Verga, M. (2007). Applicability of infrared
746 thermography as a non invasive measurements of stress in rabbit. *World Rabbit Science*, *15*(4).
747 https://doi.org/10.4995/wrs.2007.588
748 [28] Kuraoka, K., & Nakamura, K. (2011). The use of nasal skin temperature measurements in studying
749 emotion in macaque monkeys. *Physiology & Behavior*, *102*(3-4), 347-355.
750 https://doi.org/10.1016/j.physbeh.2010.11.029
751 [29] Csoltova, E., Martineau, M., Boissy, A., & Gilbert, C. (2017). Behavioral and physiological reactions
752 in dogs to a veterinary examination: owner-dog interactions improve canine well-being. *Physiology*
753 *& Behavior*, *177*, 270-281. https://doi.org/10.1016/j.physbeh.2017.05.013
754 [30] Kano, F., Hirata, S., Deschner, T., Behringer, V., & Call, J. (2016). Nasal temperature drop in response
755 to a playback of conspecific fights in chimpanzees: a thermo-imaging study. *Physiology &*
756 *Behavior*, *155*, 83-94. https://doi.org/10.1016/j.physbeh.2015.11.029

757 [31] Grandi, L. C., & Heinzl, E. L. (2016). The effect of pleasant touch on nose skin temperature, heart
758 rate and heart rate variability: preliminary results in a male laboratory rhesus monkey. *PeerJ*
759 *Preprints* (No. e2150v1). https://doi.org/10.7287/peerj.preprints.2150v1
760 [32] Merla, A., & Romani, G. L. (2007, August). Thermal signatures of emotional arousal: a functional
761 infrared imaging study. In: *2007 29th Annual International Conference of the IEEE Engineering in*
762 *Medicine and Biology Society*, pp. 247-249, IEEE. https://doi.org/10.1109/IEMBS.2007.4352270
763 [33] Nakanishi, R., & Imai-Matsumura, K. (2008). Facial skin temperature decreases in infants with
764 joyful expression. *Infant Behavior and Development*, *31*(1), 137-144.
765 https://doi.org/10.1016/j.infbeh.2007.09.001
766 [34] Ioannou, S., Ebisch, S., Aureli, T., Bafunno, D., Ioannides, H. A., Cardone, D., Manini, B., Romani, G.
767 L., Gallese, V., & Merla, A. (2013). The autonomic signature of guilt in children: a thermal infrared
768 imaging study. *PloS One*, *8*(11), e79440. https://doi.org/10.1371/journal.pone.0079440
769 [35] Salazar-López, E., Domínguez, E., Ramos, V. J., De la Fuente, J., Meins, A., Iborra, O., Gálvez, G.,
770 Rodríguez-Artacho, M. A., & Gómez-Milán, E. (2015). The mental and subjective skin: emotion,
771 empathy, feelings and thermography. *Consciousness and Cognition*, *34*, 149-162.
772 https://doi.org/10.1016/j.concog.2015.04.003
773 [36] Moliné, A., Gálvez-García, G., Fernández-Gómez, J., De la Fuente, J., Iborra, O., Tornay, F., Mata
774 Martín, L. L., Puertollano, M., & Gómez Milán, E. (2017). The Pinocchio effect and the cold stress
775 test: lies and thermography. *Psychophysiology*, *54*(11), 1621-1631.
776 https://doi.org/10.1111/psyp.12956
777 [37] Cruz-Albarran, I. A., Benitez-Rangel, J. P., Osornio-Rios, R. A., & Morales-Hernandez, L. A. (2017).
778 Human emotions detection based on a smart-thermal system of thermographic images. *Infrared*
779 *Physics & Technology*, *81*, 250-261. https://doi.org/10.1016/j.infrared.2017.01.002
780 [38] Ioannou, S., Chotard, H., & Davila-Ross, M. (2015). No strings attached: physiological monitoring of
781 rhesus monkeys (*Macaca mulatta*) with thermal imaging. *Frontiers in Behavioral Neuroscience*, *9*,
782 160. https://doi.org/10.3389/fnbeh.2015.00160

783 [39] Heintz, M. R., Fuller, G., & Allard, S. (2019). Exploratory investigation of infrared thermography for
784 measuring gorilla emotional responses to interactions with familiar humans. *Animals*, *9*(9), 604.
785 https://doi.org/10.3390/ani9090604
786 [40] Ermatinger, F. A., Brügger, R. K., & Burkart, J. M. (2019). The use of infrared thermography to
787 investigate emotions in common marmosets. *Physiology & Behavior*, *211*, 112672.
788 https://doi.org/10.1016/j.physbeh.2019.112672
789 [41] Brügger, R. K., Willems, E. P., & Burkart, J. M. (2021). Do marmosets understand others'
790 conversations? A thermography approach. *Science Advances*, *7*(6), eabc8790.
791 https://doi.org/10.1126/sciadv.abc8790
792 [42] Chotard, H., Ioannou, S., & Davila-Ross, M. (2018). Infrared thermal imaging: positive and negative
793 emotions modify the skin temperatures of monkey and ape faces. *American Journal of*
794 *Primatology*, *80*(5), e22863. https://doi.org/10.1002/ajp.22863
795 [43] Tamioso, P. R., Rucinque, D. S., Taconeli, C. A., da Silva, G. P., & Molento, C. F. M. (2017). Behavior
796 and body surface temperature as welfare indicators in selected sheep regularly brushed by a
797 familiar observer. *Journal of Veterinary Behavior*, *19*, 27-34.
798 https://doi.org/10.1016/j.jveb.2017.01.004
799 [44] Troisi, A., Schino, G., D'Antoni, M., Pandolfi, N., Aureli, F., & D'Amato, F. R. (1991). Scratching as a
800 behavioral index of anxiety in macaque mothers. *Behavioral and Neural Biology*, *56*(3), 307-313.
801 https://doi.org/10.1016/0163-1047(91)90469-7
802 [45] Maestripieri, D., Schino, G., Aureli, F., & Troisi, A. (1992). A modest proposal: displacement
803 activities as an indicator of emotions in primates. *Animal Behaviour*, *44*(5), 967-979.
804 https://doi.org/10.1016/S0003-3472(05)80592-5
805 [46] de Waal, F. B., & Yoshihara, D. (1983). Reconciliation and redirected affection in rhesus
806 monkeys. *Behaviour*, *85*(3-4), 224-241. http://www.jstor.org/stable/4534264

807 [47] Diezinger, F. T., & Anderson, J. R. (1986). Starting from scratch: A first look at a “displacement
808 activity” in group-living rhesus monkeys. *American Journal of Primatology*, *11*(2), 117-124.
809 https://doi.org/10.1002/ajp.1350110204
810 [48] Troisi, A., & Schino, G. (1987). Environmental and social influences on autogrooming behaviour in a
811 captive group of Java monkeys. *Behaviour*, *100*(1-4), 292-302. http://www.jstor.org/stable/4534585
812 [49] Aureli, F., Van Schaik, C. P., & Van Hooff, J. A. (1989). Functional aspects of reconciliation among
813 captive long-tailed macaques (*Macaca fascicularis*). *American Journal of Primatology*, *19*(1), 39-51.
814 https://doi.org/10.1002/ajp.1350190105
815 [50] Pavani, S., Maestripieri, D., Schino, G., Turillazzi, P. G., & Scucchi, S. (1991). Factors influencing
816 scratching behaviour in long-tailed macaques (*Macaca fascicularis*). *Folia Primatologica*, *57*(1), 34-
817 38. https://doi.org/10.1159/000156561
818 [51] Castles, D. L., Whiten, A., & Aureli, F. (1999). Social anxiety, relationships and self-directed
819 behaviour among wild female olive baboons. *Animal Behaviour*, *58*(6), 1207-1215.
820 https://doi.org/10.1006/anbe.1999.1250
821 [52] Schraft, H. A., & Clark, R. W. (2017). Kangaroo rats change temperature when investigating
822 rattlesnake predators. *Physiology & Behavior*, *173*, 174-178.
823 https://doi.org/10.1016/j.physbeh.2017.02.004
824 [53] Barrault, C., Soldati, A., Hobaiter, C., Mugisha, S., De Moor, D., Zuberbühler, K., & Dezecache, G.
825 (2022). Thermal imaging reveals social monitoring during social feeding in wild
826 chimpanzees. *Philosophical Transactions of the Royal Society B*, *377*(1860), 20210302.
827 https://doi.org/10.1098/rstb.2021.0302
828 [54] de Vevey, M., Bouchard, A., Soldati, A., & Zuberbühler, K. (2022). Thermal imaging reveals
829 audience-dependent effects during cooperation and competition in wild chimpanzees. *Scientific*
830 *Reports*, *12*(1), 2972. https://doi.org/10.1038/s41598-022-07003-y
831 [55] Kaplan, H. (2007). *Practical applications of infrared thermal sensing and imaging equipment*. SPIE
832 press.

833 [56] Cilulko, J., Janiszewski, P., Bogdaszewski, M., & Szczygielska, E. (2013). Infrared thermal imaging in
834 studies of wild animals. *European Journal of Wildlife Research*, *59*, 17-23.
835 https://doi.org/10.1007/s10344-012-0688-1
836 [57] Church, J. S., Hegadoren, P. R., Paetkau, M. J., Miller, C. C., Regev-Shoshani, G., Schaefer, A. L., &
837 Schwartzkopf-Genswein, K. S. (2014). Influence of environmental factors on infrared eye
838 temperature measurements in cattle. *Research in Veterinary Science*, *96*(1), 220-226.
839 https://doi.org/10.1016/j.rvsc.2013.11.006
840 [58] Thompson, C. L., Scheidel, C., Glander, K. E., Williams, S. H., & Vinyard, C. J. (2017). An assessment
841 of skin temperature gradients in a tropical primate using infrared thermography and subcutaneous
842 implants. *Journal of Thermal Biology*, *63*, 49-57. https://doi.org/10.1016/j.jtherbio.2016.11.005
843 [59] Tabh, J. K., Burness, G., Wearing, O. H., Tattersall, G. J., & Mastromonaco, G. F. (2021). Infrared
844 thermography as a technique to measure physiological stress in birds: body region and image angle
845 matter. *Physiological Reports*, *9*(11), e14865. https://doi.org/10.14814/phy2.14865
846 [60] Playà-Montmany, N., & Tattersall, G. J. (2021). Spot size, distance and emissivity errors in field
847 applications of infrared thermography. *Methods in Ecology and Evolution*, *12*(5), 828-840.
848 https://doi.org/10.1111/2041-210X.13563
849 [61] Dezecache, G., Zuberbühler, K., Davila-Ross, M., & Dahl, C. D. (2017). Skin temperature changes in
850 wild chimpanzees upon hearing vocalizations of conspecifics. *Royal Society Open Science*, *4*(1),
851 160816. https://doi.org/10.1098/rsos.160816
852 [62] Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social*
853 *psychology*, *39*(6), 1161. https://doi.org/10.1037/h0077714
854 [63] Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative
855 approach to affective neuroscience, cognitive development, and psychopathology. *Development*
856 *and psychopathology*, *17*(3), 715-734. https://doi.org/10.1017/S0954579405050340
857 [64] Anderson, D. J., & Adolphs, R. (2014). A framework for studying emotions across
858 species. *Cell*, *157*(1), 187-200. http://dx.doi.org/10.1016/j.cell.2014.03.003

[65] Macaca Nigra Project. http://www.macaca-nigra.org/

[66] Fooden, J. A. C. K. (1976). Provisional classification and key to living species of macaques (Primates: *Macaca*). *Folia Primatologica*, *25*(2-3), 225-236. https://doi.org/10.1159/000155715

[67] Neugebauer, E., Lestari, S. D., Kaletuang, I. R., Perwitasari-Farajallah, D., & Widdig, A. (2024). Distribution and Human-Mediated Establishment of Crested Macaque (Macaca nigra) Populations in the North Moluccas, Indonesia. *International Journal of Primatology*, 1-17. https://doi.org/10.1007/s10764-024-00473-z

[68] Rosenbaum, B., O'Brien, T. G., Kinnaird, M., & Supriatna, J. (1998). Population densities of Sulawesi crested black macaques (*Macaca nigra*) on Bacan and Sulawesi, Indonesia: effects of habitat disturbance and hunting. *American Journal of Primatology*, *44*(2), 89-106. https://doi.org/10.1002/(SICI)1098-2345(1998)44:2<89::AID-AJP1>3.0.CO;2-S

[69] O'Brien, T. G., & Kinnaird, M. F. (1997). Behavior, diet, and movements of the Sulawesi crested black macaque (*Macaca nigra*). *International Journal of Primatology*, *18*, 321-351. https://doi.org/10.1023/A:1026330332061

[70] Duboscq, J., Micheletta, J., Agil, M., Hodges, K., Thierry, B., & Engelhardt, A. (2013). Social tolerance in wild female crested macaques (*Macaca nigra*) in Tangkoko-Batuangus Nature Reserve, Sulawesi, Indonesia. *American Journal of Primatology*, *75*(4), 361-375. https://doi.org/10.1002/ajp.22114

[71] Micheletta, J., Waller, B. M., Panggur, M. R., Neumann, C., Duboscq, J., Agil, M., & Engelhardt, A. (2012). Social bonds affect anti-predator behaviour in a tolerant species of macaque, *Macaca nigra*. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1744), 4042-4050. https://doi.org/10.1098/rspb.2012.1470

[72] Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, *12*(1), 2.

883 [73] Thierry, B., Bynum, E. L., Baker, S., Kinnaird, M. F., Matsumara, S., Muroyama, Y., O'Bien, T. G., Petit,
884 O. and Watanabe, K. (2000). The Social Repertoire of Sulawesi Macaques. *Primate Research. 16*,
885 203-226. https://doi.org/10.2354/psj.16.203
886 [74] Micheletta, J., Engelhardt, A., Matthews, L. E. E., Agil, M., & Waller, B. M. (2013). Multicomponent
887 and multimodal lipsmacking in crested macaques (*Macaca nigra*). *American Journal of*
888 *Primatology*, *75*(7), 763-773. https://doi.org/10.1002/ajp.22105
889 [75] Qomariah, I. N., Perwitasari-Farajallah, D. Y. A. H., Iskandar, E., & Berthier, J. M. (2023).
890 Characteristics of sleeping sites and sleeping trees selected by Celebes crested macaques (*Macaca*
891 *nigra*) in Tangkoko, North Sulawesi, Indonesia. *Biodiversitas Journal of Biological Diversity*, *24*(11).
892 https://doi.org/10.13057/biodiv/d241142
893 [76] Fischer, J., Noser, R., & Hammerschmidt, K. (2013). Bioacoustic field research: a primer to acoustic
894 analyses and playback experiments with primates. *American journal of primatology*, *75*(7), 643-
895 663. https://doi.org/10.1002/ajp.22153
896 [77] Nakayama, K., Goto, S., Kuraoka, K., & Nakamura, K. (2005). Decrease in nasal temperature of
897 rhesus monkeys (*Macaca mulatta*) in negative emotional state. *Physiology & Behavior*, *84*(5), 783-
898 790. https://doi.org/10.1016/j.physbeh.2005.03.009
899 [78] Castelli, W. A., & Huelke, D. F. (1965). The arterial system of the head and neck of the rhesus
900 monkey with emphasis on the external carotid system. *American Journal of Anatomy*, *116*(1), 149-
901 169. https://doi.org/10.1002/aja.1001160108
902 [79] Steketee, J. (1973). Spectral emissivity of skin and pericardium. *Physics in Medicine &*
903 *Biology*, *18*(5), 686. https://doi.org/10.1088/0031-9155/18/5/307
904 [80] R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for
905 Statistical Computing, Vienna, Austria. https://www.R-project.org/
906 [81] Christensen, R. H. B., & Christensen, M. R. H. B. (2015). Package 'ordinal'. *Stand*, *19*(2016).
907 [82] Fox, J., Friendly, G. G., Graves, S., Heiberger, R., Monette, G., Nilsson, H., ... & Suggests, M. A. S. S.
908 (2007). The car package. *R Foundation for Statistical computing*, *1109*, 1431.

[83] Lenth, R., & Lenth, M. R. (2018). Package 'lsmeans'. *The American Statistician*, *34*(4), 216-221.

[84] Keselman, H. J., & Rogan, J. C. (1977). The Tukey multiple comparison test: 1953–1976. *Psychological Bulletin*, *84*(5), 1050. https://doi.org/10.1037/0033-2909.84.5.1050

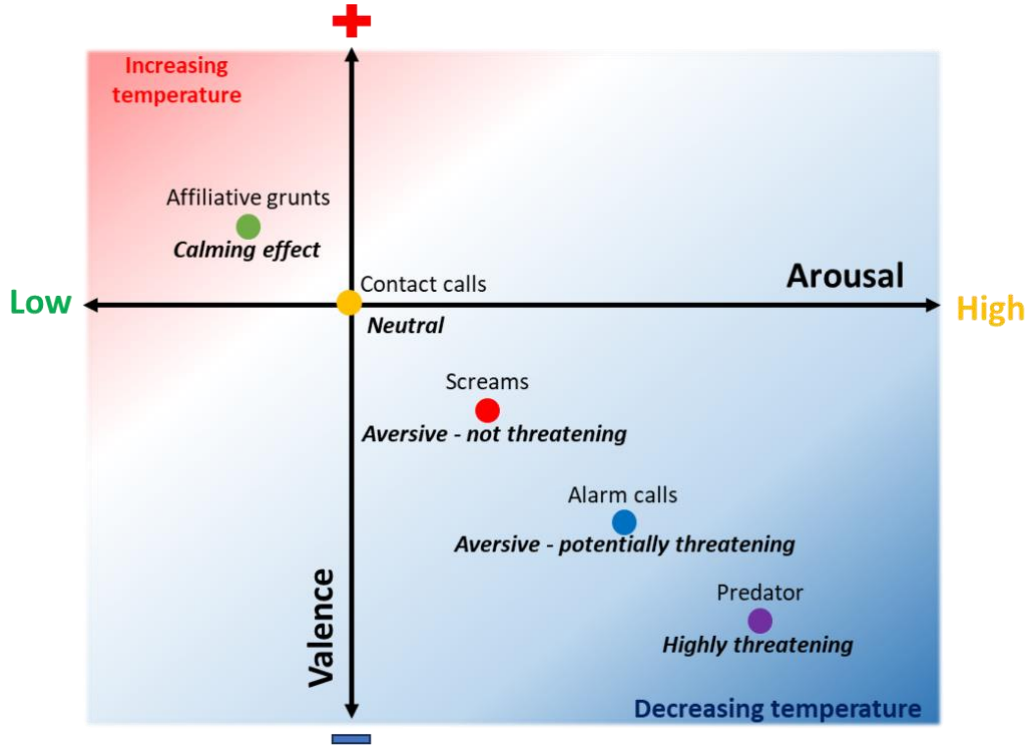
[85] Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, *22*(2), 142-146.

[86] Silk, J. B., Seyfarth, R. M., & Cheney, D. L. (2016). Strategic use of affiliative vocalizations by wild female baboons. *PLoS One*, *11*(10), e0163978. https://doi.org/10.1371/journal.pone.0163978

[87] Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee food calls are directed at specific individuals. *Animal behaviour*, *86*(5), 955-965. https://doi.org/10.1016/j.anbehav.2013.08.013

[88] Neumann, C., Agil, M., Widdig, A., & Engelhardt, A. (2013). Personality of wild male crested macaques (Macaca nigra). *PloS one*, *8*(8), e69383. https://doi.org/10.1371/journal.pone.0069383

[89] Berthier, J. M. (2024). *The role of emotions in the communication system and behaviour of female crested macaques (Macaca nigra): A new approach using infrared thermography.* Doctoral dissertation, University of Kent. https://doi.org/10.22024/UniKent/01.02.107477



925

926 **Figure 1.** Positions of the five stimuli used in field experiments on a two-dimensional core affect space
927 (adapted from [64]). Valence indicates whether the stimulus is perceived as positive or negative, while
928 arousal reflects emotional excitement. Colored dots represent stimuli, with labels and expected
929 macaque perceptions in italics. Color gradient shows anticipated facial temperature changes (blue for
930 decrease, red for increase), with darker shades indicating greater deviation from baseline.

931

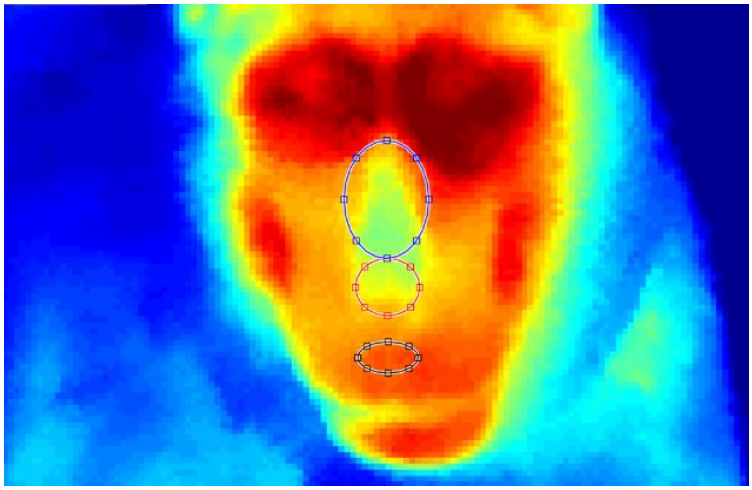
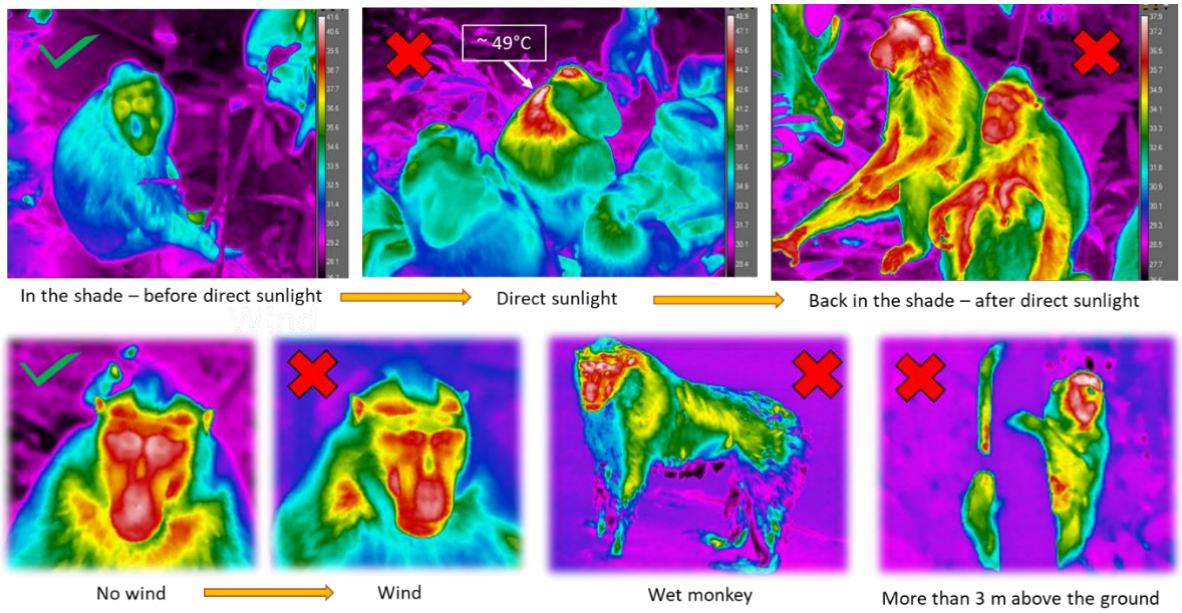


Figure 2. ROIs - nose tip, nose bridge and upper lip - delimited by ellipses using a customized MATLAB script. The ellipse for the nose tip excludes nostrils to prevent measurement bias, as nostrils are typically cooler than nose skin due to airflow.



938

939 **Figure 3.** Illustrations of appropriate (green tick) and inappropriate (red cross) environmental

940 conditions to record thermal data.

941

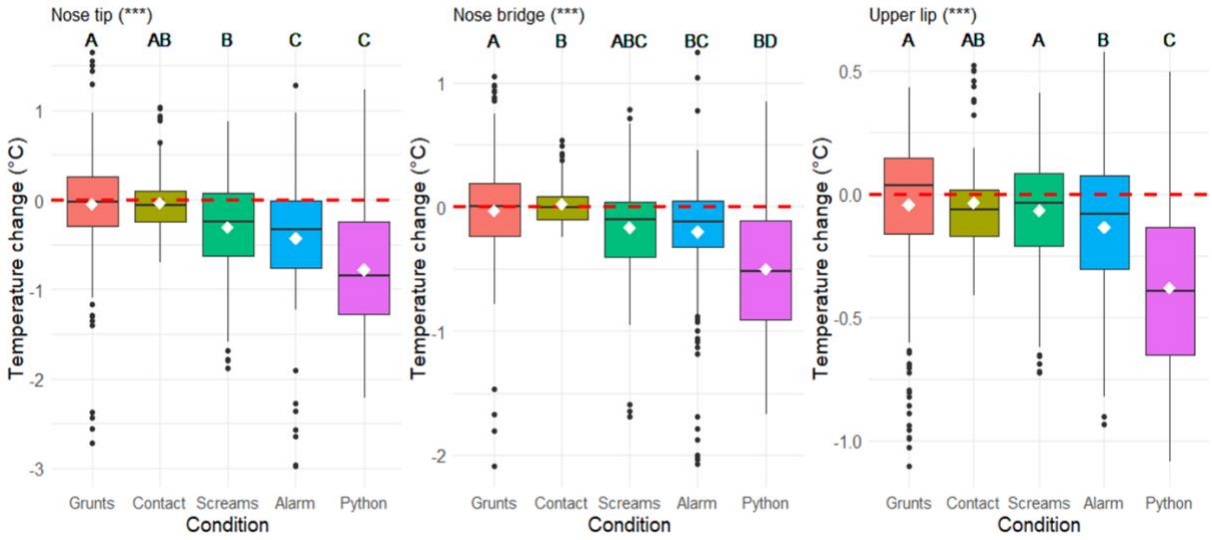
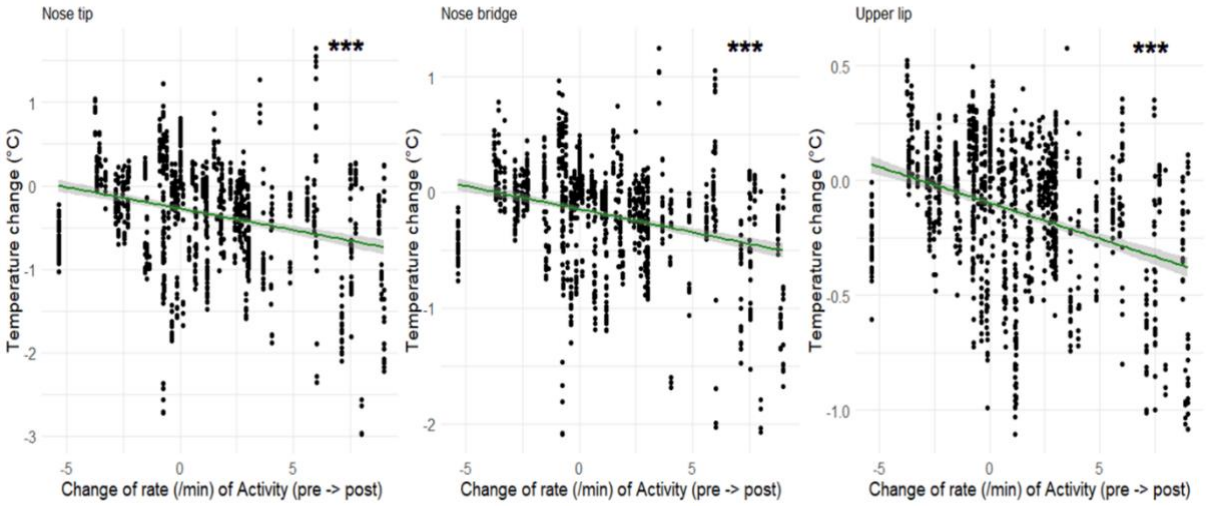


Figure 4. Temperature variations of nose tip (left), nose bridge (middle) and upper lip (right) depending on the condition ordered in their assumed arousal intensity. The mean is indicated by white rhombus. Red dashed line represents the centered baseline. Boxplots sharing the same letters are not statistically different (**** P < 0.001*).

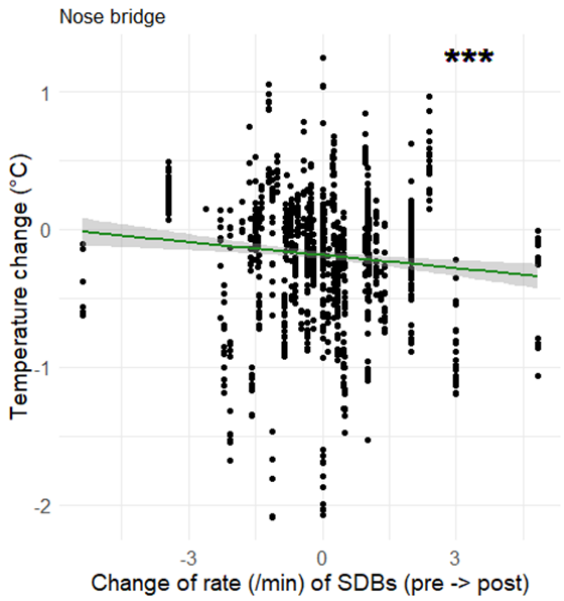


948

949 **Figure 5.** Relationship between the change of activity rates post-stimuli and the temperature changes

950 from the baseline of the nose tip (left), nose bridge (middle), the upper lip (right). (*** *P < 0.001*).

951

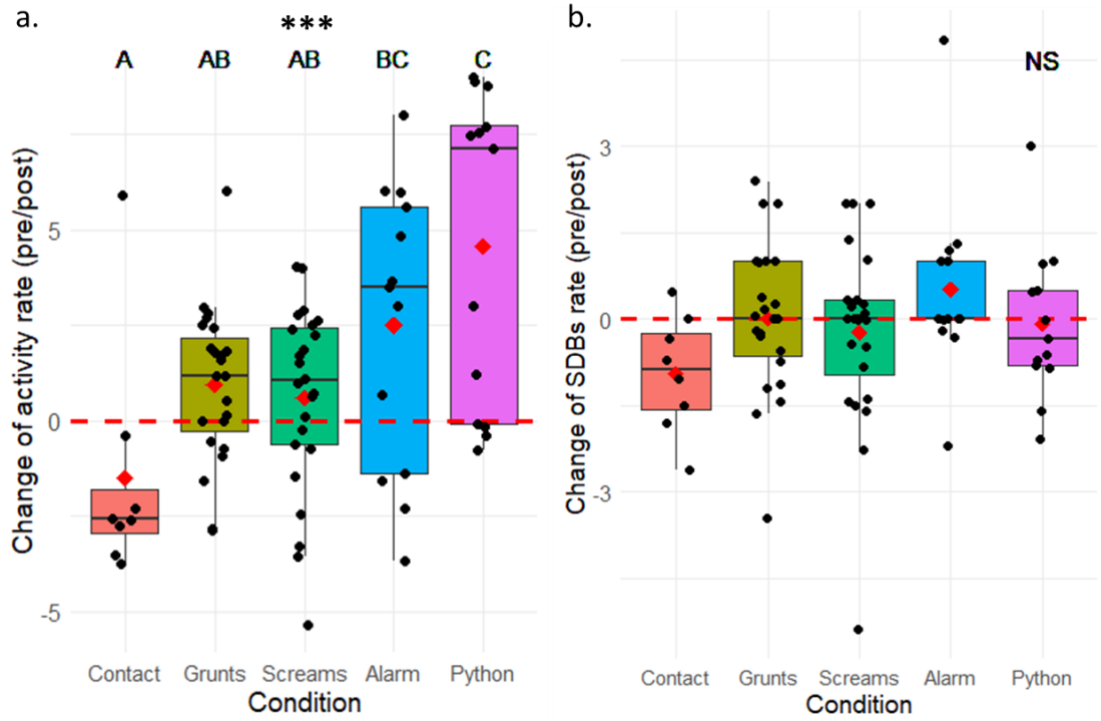


952

953 **Figure 6.** Relationships between the temperature changes from the baseline of the nose bridge and the

954 change of SDB rate post-stimuli (*** *P < 0.001*).

955



956

957 **Figure 7.** Changes of (a) physical activity and (b) SDB rate depending on the condition. The mean is

958 indicated by red rhombus, and black dots show individual data points. Red dashed line represents the

959 centred baseline. Boxplots sharing the same letters are not statistically different. (**** P* < 0.001; *NS:*

960 *Non-Significant*).

961

962 **Supplementary material**

963

964

965 **Figure S1.** Picture of the two python models used during field experiments – model 1 on the left and model 2 on
966 the right. During the experiments, only the head of the python model was made visible to the focal animal by a
967 field assistant.

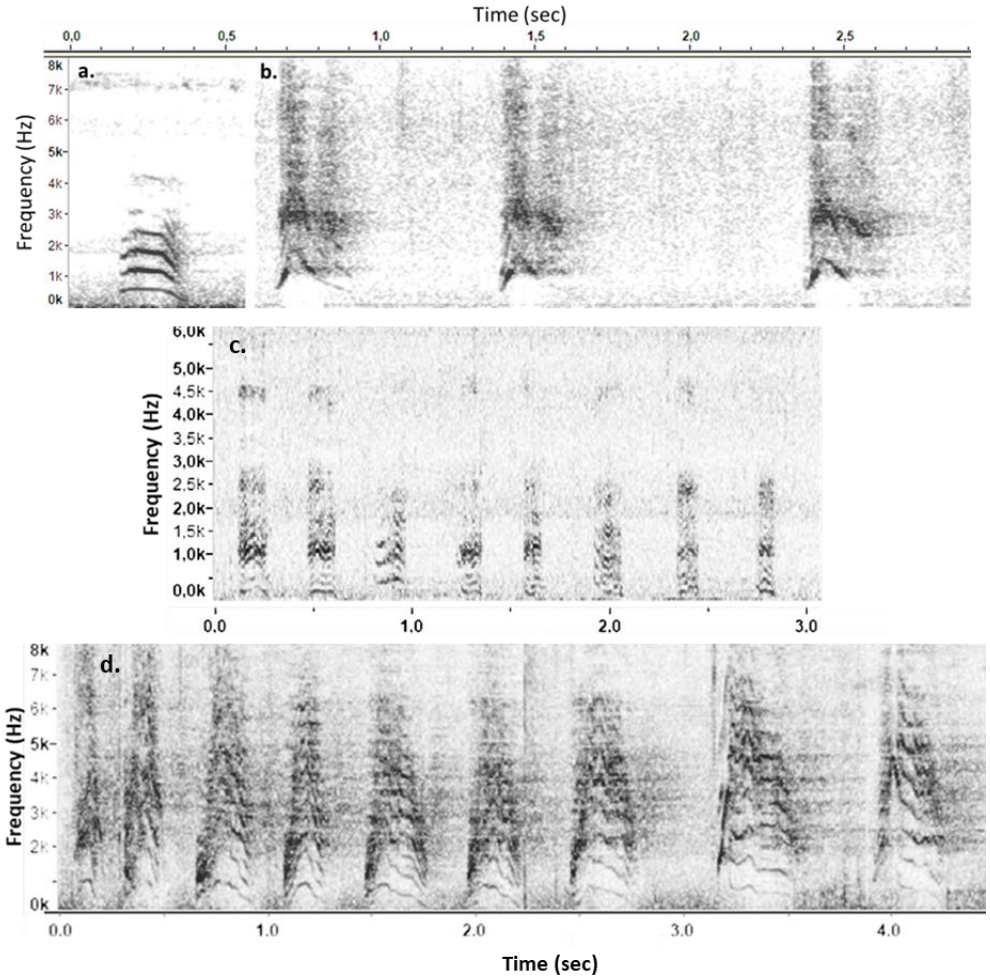


Figure S2. Spectrogram of (a.) a contact call, (b.) series of three alarm calls, (c.) affiliative grunts, and (d.) aggressive screams of female crested macaques. Spectrograms parameters: FFT length = 1024 points; window = Hann; frame size = 100% (Generated with the software Audacity® 2.3.2).

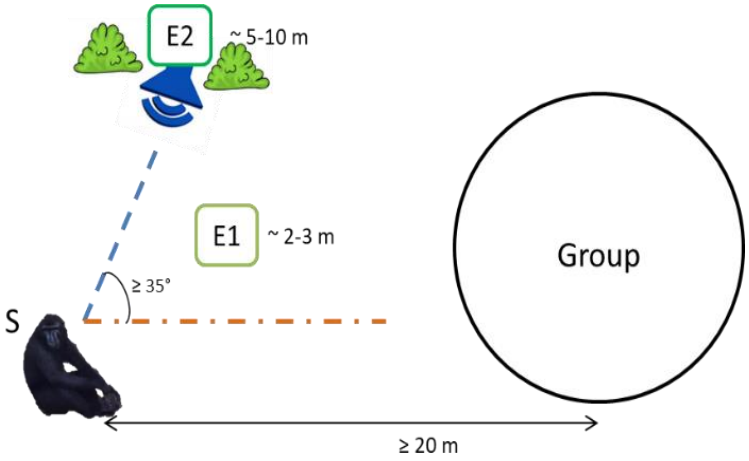
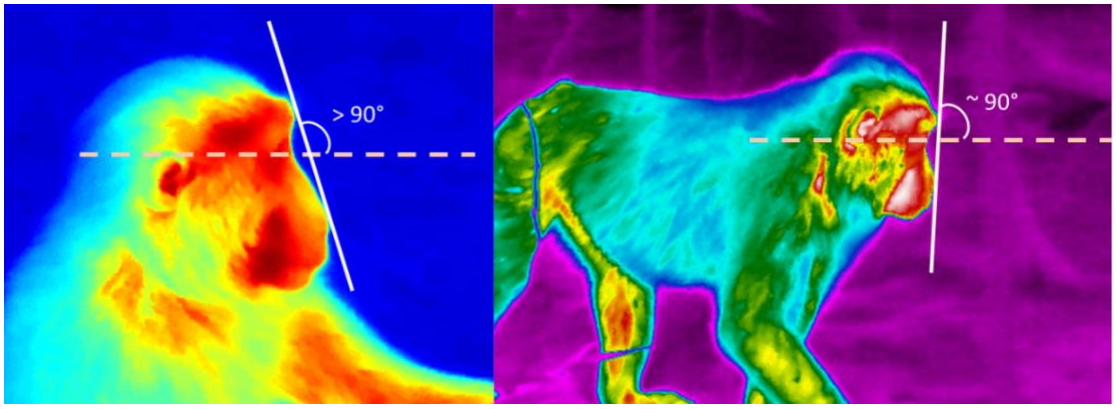
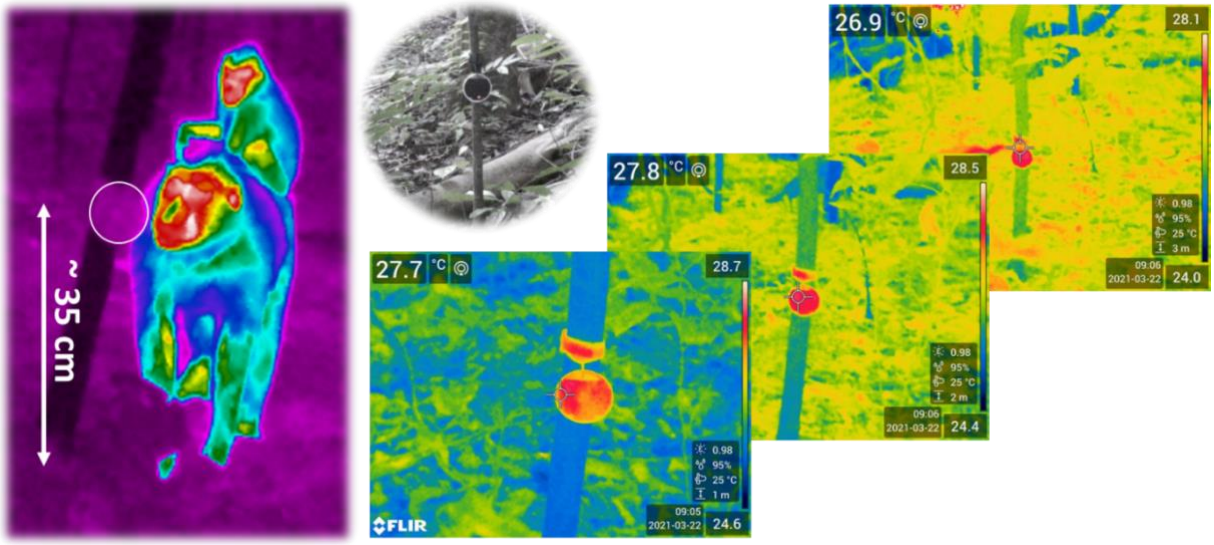


Figure S3. Drawing and picture illustrating the set-up of the playback experiments. 'S': Subject; 'E1': Experimenter 1 (the camera holder); 'E2': Experimenter 2 (the loudspeaker holder).

977 

978 **Figure S4.** Natural angles of the head when the monkey is looking straight and when it is sitting (left) or standing
979 (right).



980

981 **Figure S5.** Illustration of the use of the black target of known emissivity to control for the potential effect of the
982 distance on the thermal measurements in the same environmental condition as the experimental trials (procedure
983 followed just after each successful experimental trial).

Table S1. Descriptions of focal subjects among the 19 adult females present in the group PB1B at the beginning of data collection, and field experiments information.

Adult females			Playback experiments			Python		Caller
ID-Code	Age	Rank	Grunts	Screams	Alarm	Contact Call	Models	
AA	11	1	QP,HA *(o,p)*	HA,QP,BA *(o,o,p)*	.	HA,QP *(p,i)*	2,2	1
NA	6	2	KP,CP *(o,p)*	.	.	.	.	3
KP	15	3	EP,NA *(o,p)*	NA,EP *(o,o)*	DP,CA *(p,p)*	EP *(o)*	1	5
HA	7	4	.	.	.	.	.	7
FP	16	6	.	AA,BA *(p,i)*	.	UP *(o)*	1	1
CP	15	8	UP,KA *(s,o)*	KA,UP *(o,s)*	UP,CA *(o,o)*	UP,JA *(o,o)*	2,2	2
XP	13	9	DP,NA *(o,o)*	.	DP *(p)*	.	.	5
DP	16	10	HA,NP *(o,o)*	NP,HA *(p,p)*	.	NP *(p)*	1	5
BA	11	11	NP,CA *(o,o)*	CA,NP *(o,o)*	NP,CA *(o,o)*	CA,NP *(o,o)*	2,2	4
EP	16	13	.	.	.	.	.	3
QP	12	15	.	.	.	.	.	3
DA	9	16	LA *(p)*	FP,LA *(p,p)*	.	.	.	1
GA	7	17	.	.	.	.	.	.
KA	7	20	.	.	.	.	.	8
NP	16	21	XP,KA *(o,o)*	XP,KA *(o,o)*	CA *(p)*	KA,XP *(s,o)*	1,2	9
UP	12	22	KA,KP *(s,o)*	KP,KA *(s,o)*	KP,DP *(o,p)*	KP,KA *(o,p)*	2,1	5
CA	11	23	JA,BA *(o,s)*	BA,JA *(o,o)*	DP (o)	JA *(o)*	2	7
LA	7	24	HA,XP *(p,i)*	HA,XP *(p,p)*	NP,CP *(i,i)*	.	.	2
JA	7	25	.	DA,NP (o,s)	.	.	.	4
TOTAL	***19 ind.***		***23***	***25***	***13***	***14***		***75***

*Legend: **Age** estimate based on the year of first birth recorded by the MNP + 4 years (in crested macaques, females reach sexual maturity and have their first offspring at 4 years old on average); **Rank** calculated through Elo ratings at the end of data collection and including the 25 adult females present in the group at the end of data collection. **Playback experiments** performed with pre-recorded calls (affiliative grunts, aggressive screams, and alarm calls): the table shows the ID-codes of the callers in the chronological order of passage when several repetitions have been conducted. The column '**Contact Call**' gives the ID-codes of the caller in the chronological order of passage when several repetitions have been conducted; the column '**Model**' gives the python model(s) used in the chronological order of presentation if two repetitions have been performed – (1) for model 1 and (2) for model 2. In brackets is the reproductive status of the subject at the time of the experiment: 's' for anogenital swelling, 'p' for pregnant, 'i' for infant of less than 6 months old, and 'o' for other. **Audience/Caller**: the number of times the vocalisations (among contact calls, screams, grunts and alarm calls) of the individual has been used for an experiment.*

997 **Table S2.** Environmental and physical factors that could impact thermal measurements considered in this study.

FIXED EFFECT		DEFINITIONS
F (3)	**Posture**	
	Sitting	The bottom is touching the substrate (ground/branch, etc.)
	Standing still	The four limbs are touching the substrate, not the bottom
	Moving	Including walking, jumping, climbing up and down a tree
C (m)	**Distance**	Distance between the face of the subject animal and the camera sensor
C (°C)	**Air temperature**	Recorded with an automatic temperature and humidity logger (OMEGA OM-HL-SP-TH)
C (%)	**Relative humidity**	
F (2)	**Weather**	Sunny or cloudy
F (3)	**Environment**	
	Sea front	Open area near the sea but still under the cover of high vegetation; often a bit windy; dark substrate (rock or sand)
	Open forest	Primary forest with large and high trees; only sparse vegetation at ground level
	Dense forest	Secondary or damaged forest; dense low vegetation at ground level (scrub)
F (2)	**Vegetation dryness**	Whether the vegetation was wet following rain or not (i.e., dry)

998 *Legend: The first column indicates whether the variable is a factor (F) or continuous (C) and number of levels or the*

999 *unit of measurement in brackets*

1000

1001 **Table S3.** Definitions of the self-directed behaviors considered in this study.

Behavior	Definitions
SELF-DIRECTED BEHAVIOURS	
Self-grooming	to clean its own fur.
Scratch	to rake the skin repeatedly using fingers of hands or feet.
Yawn	a gaping movement of the mouth.
Body shake	to shake the body (often the upper part) vigorously from side to side.

1002 *These definitions have been adapted from the ethogram established by the Macaca Nigra Project.*