



OPEN

Data-driven analysis of facial thermal responses and multimodal physiological consistency among subjects

Saurabh Sonkusare^{1,2✉}, Michael Breakspear^{1,4}, Tianji Pang^{1,3}, Vinh Thai Nguyen¹, Sascha Frydman¹, Christine Cong Guo¹ & Matthew J. Aburn^{1✉}

Facial infra-red imaging (IRI) is a contact-free technique complimenting the traditional psychophysiological measures to characterize physiological profile. However, its full potential in affective research is arguably unmet due to the analytical challenges it poses. Here we acquired facial IRI data, facial expressions and traditional physiological recordings (heart rate and skin conductance) from healthy human subjects whilst they viewed a 20-min-long unedited emotional movie. We present a novel application of motion correction and the results of spatial independent component analysis of the thermal data. Three distinct spatial components are recovered associated with the nose, the cheeks and respiration. We first benchmark this methodology against a traditional nose-tip region-of-interest based technique showing an expected similarity of signals extracted by these methods. We then show significant correlation of all the physiological responses across subjects, including the thermal signals, suggesting common dynamic shifts in emotional state induced by the movie. In sum, this study introduces an innovative approach to analyse facial IRI data and highlights the potential of thermal imaging to robustly capture emotion-related changes induced by ecological stimuli.

Galvanic skin response (GSR) and heart rate (HR) metrics are commonly employed in psychophysiological studies^{1–3}. However, it has been argued that measures targeting a variety of somatic effects may be best able to capture emotion-related brain-body states². An innovative technique which compliments the traditional measures is facial infra-red imaging (IRI). IRI quantifies the temperature fluctuations of the face, predominantly resulting from blood flow changes⁴ although it has also been shown to capture the temperature effect of perspiration⁵. As a contact-free technique, it allows ecologically valid experimental conditions which recent studies have exploited⁶. Although it has not yet been as widely used as GSR and HR measures, thermal imaging has shown promising results in affective research^{7–9}.

The current use of facial thermal imaging analysis is based on ad hoc choices of facial regions, usually chosen a priori and used with a region of interest (ROI) analysis i.e. extracting the thermal signals from specific facial regions like the nose-tip, the cheeks and the forehead^{7–12}. ROI-based thermal signal extraction can be strongly influenced by motion artefacts¹³ with only a few studies implementing motion tracking or motion correction^{14,15}. Furthermore, ROI based approaches have their own unique challenges and limitations. First, they only exploit a small portion of the information encoded in the whole face. Second, manual steps are time consuming and only feasible for the analysis of small data sets with low sampling rate. While several useful improvements in methods have been proposed such as in tracking¹⁶, displacement correction of frames by matching anatomical landmarks¹⁷ and analysis based on machine learning¹⁸, these have yet to be adopted in general practice. Lastly, different *a priori* choices for the size of ROIs, their shapes and placement can further introduce biases⁴.

Although thermal signal extraction based on ROIs has provided many insights into thermal effects in affective research, these methods can be complemented by data-driven approaches which uncover underlying data features with less *a priori* assumptions. Independent component analysis (ICA), specifically spatial ICA (sICA), is a subset of such methods for blind signal separation employed under assumptions of statistical independence of the source signals¹⁹. Due to the effectiveness in capturing the essential structure of diverse kinds of data, ICA

¹QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia. ²School of Medicine, The University of Queensland, Brisbane, Australia. ³Northwestern Polytechnical University, Xi'an, Shaanxi, China. ⁴The University of Newcastle, University Drive, Callaghan, NSW 2308, Australia. ✉email: sonkusaresaurabh@gmail.com; Matthew.Aburn@qimrberghofer.edu.au

has been widely used in many applications, such as functional magnetic resonance imaging (fMRI) analysis^{20,21} which measures the haemodynamic changes as a proxy for neuronal activity. Similar to the brain, facial skin has an extensive distribution of blood vessels. Furthermore, temperature changes associated with the face can have various contributions, including stimulus induced changes but also respiratory confounds, perspiration, and noise. The diverse nature of these signals suggests that blind signal separation techniques could be advantageous for isolating the different sources of signal¹⁹ and distinct spatial components.

With these considerations in mind, we sought to employ an sICA approach to characterize the dynamic facial thermal responses during naturalistic emotional experience. For this we acquired facial IRI data alongside other traditional physiological measures while subjects watched a 20-min emotional movie. Naturalistic stimuli have emerged as an alternative to strictly controlled paradigms and stimuli, such as pictures and sounds, as they offer better ecological validity^{22,23} and induce robust physiological responses²⁴. Furthermore, movies have been previously used in a thermal imaging study to induce emotions²⁵. To analyse thermal signals in a maximally naturalistic setting where participants were free to move, we combined state-of-the-art motion correction method based on optical flow with sICA to extract dynamic temperature changes that manifest in different facial regions. Since ROI-based facial IRI studies have recognized the nose-tip as the most sensitive region for temperature fluctuations, we hypothesised that a nose-tip spatial component would be reliably captured with the sICA method as well.

The secondary aim of this study was to test the shared nature of physiological responses between subjects. For the three main modalities GSR, HR and thermal imaging we aimed to test whether these measures show consistency of the dynamic responses across subjects. In neuroimaging studies using naturalistic stimuli such as movies, inter-subject correlation (ISC) analyses have demonstrated common covarying patterns of brain activity^{22,26,27}. Here we leverage this analytical technique to investigate whether subjects share a common temporal response in these physiological signals (heart rate metrics, skin conductance and thermal responses).

Materials and methods

Thermal imaging and physiological data acquisition. Twenty healthy human participants (11 females, aged 22–30 years, mean = 25.7 years) were recruited for the study. Informed consent was obtained from all participants. All participants had normal or corrected-to-normal vision (using contact lenses if applicable). Exclusion criteria included habitual smoking, the presence of a chronic illness (e.g., cardiovascular or thyroid conditions), psychological disorders (e.g. depression or anxiety) or any other illness requiring regular medication. The study was approved by the Research Ethics Board of QIMR Berghofer and performed in agreement with the Declaration of Helsinki. The participants had the choice to withdraw from the study at any time. Each participant was compensated with a \$50 voucher for their time. At the end of the data acquisition session a questionnaire was completed by the participants regarding subjective ratings of emotional response to the movie (Table S1). Physiological data (thermal, GSR and HR) from three participants were excluded from analysis due to the failure of accurate trigger information and incomplete data acquisition. Questionnaire ratings could not be recorded from one subject. Informed consent to publish identifying images (RGB and thermal) was obtained from two subjects.

Prior to the data acquisition, participants acclimatized for about 10 min within the experimental room. We adopted this practice as per prior studies^{7,11} and recommendations suggested⁴. The temperature and the humidity of the room were kept within a steady range (22 ± 2 °C; 55–65% relative humidity). Where necessary, participants' hair was secured away from their forehead with an unobtrusive hat. Participants were also asked to avoid alcohol and caffeine for at least 2 h prior to the experimental session to minimize the vasoactive effects that these substances have on the skin temperature. Testing was performed exclusively between 2 and 5 p.m. to avoid any potential confounding effects of the circadian rhythm. After the participants had assumed a comfortable posture in a fixed chair, GSR and electrocardiograph (ECG) electrodes were attached to fingers and arms, respectively. A thermal imaging camera and an RGB video camera were then manually focused on the subject's face. The researcher checked the recording quality and left the experimental room but retained audio-visual contact via CCTV. The paradigm consisted of a ~20-min short movie *The Butterfly Circus*²⁸. This stimulus has been used in our prior study and was selected by a film critic²⁶ because of its critical acclaim and its ability to evoke an emotional experience. Ten seconds of countdown was prefixed to the movie. Stimuli were presented on a 24-inch computer screen 40 cm in front of the subject. The sound was kept to the same levels for all subjects and was presented via two loudspeakers placed beside the stimulus screen.

An in-house built integrated hardware and software experimentation platform LabNeuro was used to integrate the multi-modal data acquisition. Thermal imaging of the face was performed by a FLIR A615 camera with a 15 mm lens, an uncooled Vanadium Oxide (VoX) detector producing images of 640×480 pixels in size placed approximately 100 cm away from the participants' faces. The FLIR A615 provides a temperature detection range from -20 to 2000 °C with the NEDT (noise equivalent differential temperature) smaller than 0.05 °C at 30 °C. The thermal camera response was blackbody-calibrated to nullify noise-effects related to the sensor drift/shift dynamics and optical artifacts. For this camera emissivity correction was configurable from 0.01 to 1.0 and was set at 0.98. The sampling rate for thermal imaging was set at 5 Hz. This was performed in order to generate sufficient frames to balance out any potential movement artifacts by the participant. The cameras were incorporated into LabNeuro using the LabVIEW Image Acquisition library.

GSR recordings were obtained by two Ag/AgCl electrodes (0.8-cm diameter) filled with a conductive paste and attached to the distal phalanges of the index and ring fingers of the participant's left hand. Skin conductance was recorded using a standard constant voltage system of 0.5 V and recordings were continuously digitized by an A/D converter with a sampling rate of 2 kHz. The recording of GSR data was acquired using National Instruments (NI) CompactDAQ modular IO hardware and software for the system was written using NI LabVIEW and NI

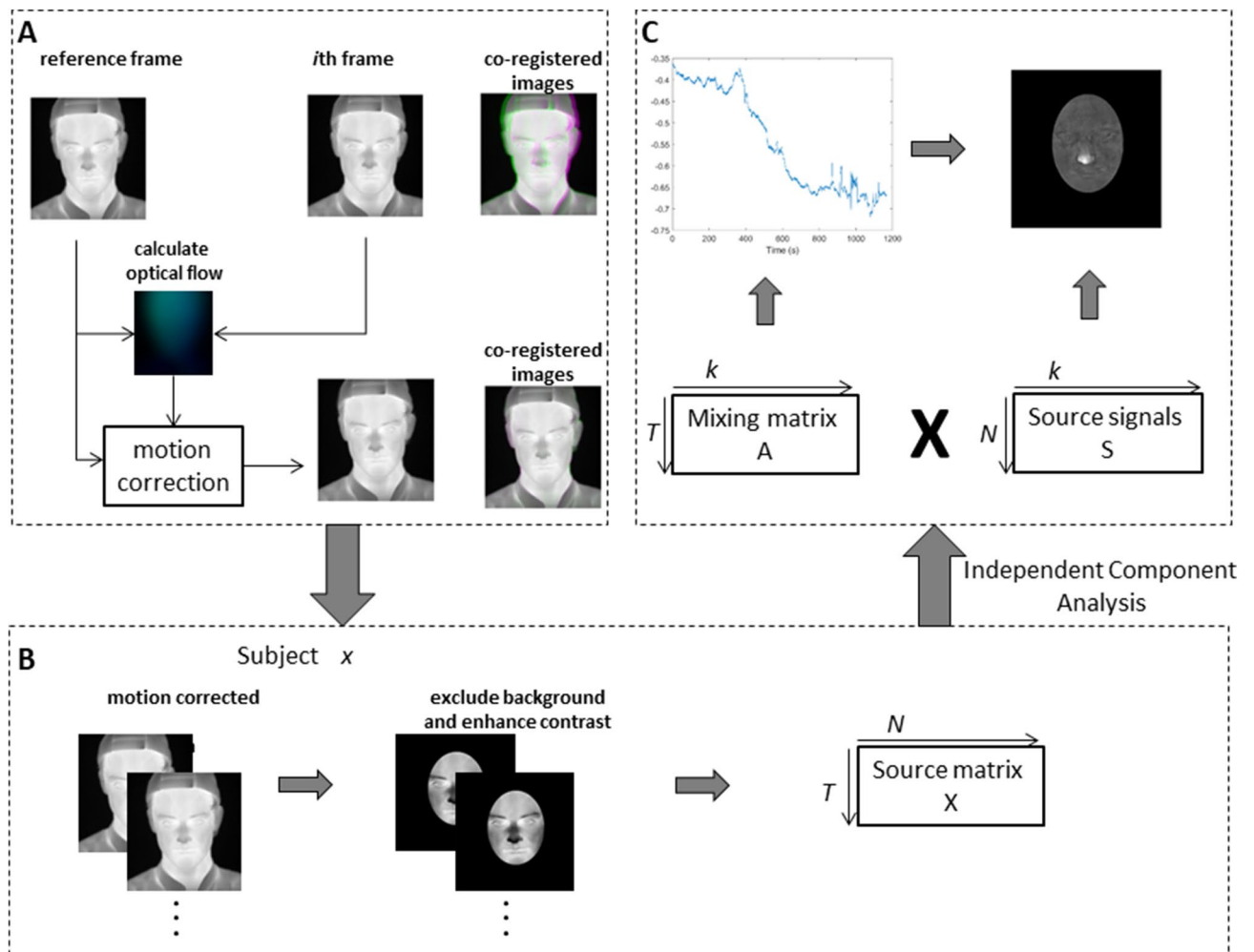


Figure 1. The computational pipeline for employing spatial independent component analysis (sICA) on thermal imaging data (A) Motion correction framework showing co-registered images with and without application of optical flow. For each subject, the whole thermal imaging data are aggregated into a single matrix, in which each row represents the thermal imaging data in one time point and each column stands for the time series of one single pixel (B) A mask to exclude background, i.e. neck and clothes, and retain only the face was applied to each frame. The data from these images were used as the source matrix (C) sICA—illustration of the mixing matrix, each column of which represents the time course of the corresponding source signal. An exemplar time series of nose components is shown. Each row of the source signal matrix represents one independent spatial map. Thermal image of subject L06 is used for illustrative purposes.

Biomedical Toolkit. For ECG data collection, electrodes were attached to the mid-upper left arm, left wrist, and mid-upper right arm. We asked participants to put their hands on chair hand-rests to minimize motion artifacts. ECG data was recorded with a sampling rate of 2 kHz. GSR signals and ECG data were collected simultaneously with the thermal imaging data.

Extraction of facial thermal responses. The pipeline for this workflow is illustrated in Fig. 1. First the thermal frames were pre-processed to eliminate motion by applying a dense optical flow algorithm²⁹. Specifically, the nonlinear motion vector field between each frame pair was first estimated by this method. Discontinuities in the transformation, which could distort the boundary of the face, were effectively removed by iterating motion correction and a Gaussian smoothing step applied to the motion vector field until convergence (where the displacement of all pixels is less than 1).

Similar to its use in fMRI data, spatial ICA was then applied to remove artifacts and to unmix statistically independent (true) sources of facial thermal signal change. To reduce noise and the number of parameters to be estimated, pre-whitening and dimension reduction were employed before the application of ICA. sICA was applied using a publicly available software package FastICA (<http://research.ics.aalto.fi/ica/fastica/code/dlcode.shtml>)^{30,31}. The specific parameters inputted in the code are provided in supplementary Table S3. The complete pipeline of the implementation of ICA together with motion correction is shown in Fig. 1.

Motion correction is a separate pre-processing step before sICA is employed. It is the first step (Fig. 1A) and does not depend on any ROI or facial components. It eliminates head and face motion by non-linearly morphing each frame back to coincide with preceding frames, leaving a video where the head does not move

in space or rotate. This motion correction is important for sICA because the ICA model assumes that changes in a pixel's intensity represent changes in temperature and not underlying movement of the face. The novel pre-processing method to remove motion is also valuable in combination with a traditional ROI signal extraction, as it serves to eliminate the issue of motion artifacts. In the research field of computer vision, this general problem of optical flow motion correction is a well-researched problem, so we are here bringing its novel application to psychophysics.

Though the motion correction and the ICA are both automated processes, the selection of physiologically meaningful independent components still needs manual inspection. To identify independent components with the most anatomico-physiological meaning we restricted attention to spatially localized components corresponding to broad anatomical boundaries such as nose, cheeks and forehead. The time series corresponding to each component was then normalized to have values between 0 and 1.

Comparison of nose component signal with region-of-interest based nose-tip signal. For validation and comparison of sICA against an ROI approach, a circle of 9-pixel radius (area ~ 250 pixels) was used to define an ROI to extract thermal signals from the nose-tip of the thermal data (Fig. 4A left). An identical size of ROI was chosen to have uniformity across participants irrespective of face morphology. The same motion correction was used with the ROI and sICA to allow comparison. The time series extracted by ROI were normalized to have values between 0 and 1 for each subject to mitigate against spurious inter-subject differences. The correlation of ROI based nose-tip signal to that obtained from nose IC was computed, and significance of this correlation assessed using a non-parametric permutation test. Specifically, we calculated the group mean of the Pearson correlation between each subject's nose component signal and their ROI based nose-tip signal. To generate the null distribution, each permutation used the ROI nose-tip signal with a random circular shift, to preserve temporal autocorrelation, and the group mean of the correlation between these shifted data and the thermal signals was computed. The null distribution was generated from 5000 permutation realizations.

Heart rate variability. The ECG signal was pre-processed using QRSTool software³² to detect the R peaks with the ability to manually correct for missed peaks. Inter-beat interval (IBI) time series was then computed from this and normalised to have values between 0 and 1. R peak data were further analysed using HRVAS toolbox³³ to obtain heart rate variability (HRV) frequency domain measures. These were calculated via the autoregressive method using a window size of 16 s, with 15 samples overlap, nfft of 1024 and cubic spline interpolation rate of 2 Hz. These parameters were chosen as they had previously been demonstrated to capture robust responses to simple stimuli even with a short duration of data²⁵. HRV data metrics were computed for the whole stimulus but edge effects for frequency estimation led to loss of approximately 8 s of data at the beginning and the end. Time-frequency decompositions of inter-beat intervals are typically linked to autonomic influences in distinct frequency bands. Lower frequency (LF) HRV (0.04–0.14 Hz) mainly reflects changes in sympathetic and parasympathetic outflows, while high frequency (HF) variability (0.15 to 0.4 Hz) is primarily due to modulation of parasympathetic outflow^{34,35}. However, the interpretation of low frequency HRV has been questioned^{36,37}. Respiratory sinus arrhythmia (RSA) is a major contributor to HF HRV and is thought to be due to respiration modulating the cardiac vagal activity³⁸.

Comparison of nose component signal with skin conductance (GSR). GSR is a standard measure of arousal. Event based studies have found an inverse relation between GSR and nasal temperature^{15,39}. The GSR time courses were first detrended and low pass filtered at 5 Hz and subsequently normalized to have values between 0 and 1 for each subject to mitigate against spurious inter-subject differences in baseline. For statistically comparing GSR with nose component signal, identical permutation testing procedure as used for comparison of thermal signals to that ROI based signal was used.

Inter-subject correlation of physiological signals. Pearson correlation coefficients between each pair of participants were computed separately for both thermal signals, GSR data and heart rate data. Inter-subject correlation (ISC) was computed as the mean pairwise correlation between participants. With 17 participants, each analysis comprised a total of 136 correlation pairs. Non-parametric permutation tests were used to identify statistically significant inter-subject correlation ($p=0.05$)^{40,41}. Specifically, the null distribution was generated from 5000 permutation realizations. These permutations comprised circularly shifted data in a way to preserve temporal autocorrelation in the physiological signals but disrupt correlations between subjects.

Testing for similarity between two different recordings at group level. To assess statistically whether two different physiological responses were significantly correlated (such as the nose IC and nose-tip ROI thermal signals, the nose IC thermal signals with GSR, and the nose IC thermal signals with movie emotions), we employed non-parametric permutation testing. We first calculated for each subject the Pearson correlation coefficient between the two signals of interest. The correlation values were then averaged for all the subjects to obtain a group mean correlation value. To statistically test the significance of the obtained group mean correlation value, we applied non-parametric permutation tests with a null model where data from one modality was randomly circularly shifted relative to the other (thus preserving autocorrelation), before computing the group mean correlation of the null. Five thousand realizations were used to generate the null distribution.

Facial expression analysis. Facial emotions expression scores, both for the movie actors and for the subject's responses were obtained by processing the movie frames, and the RGB video recordings of the each subject

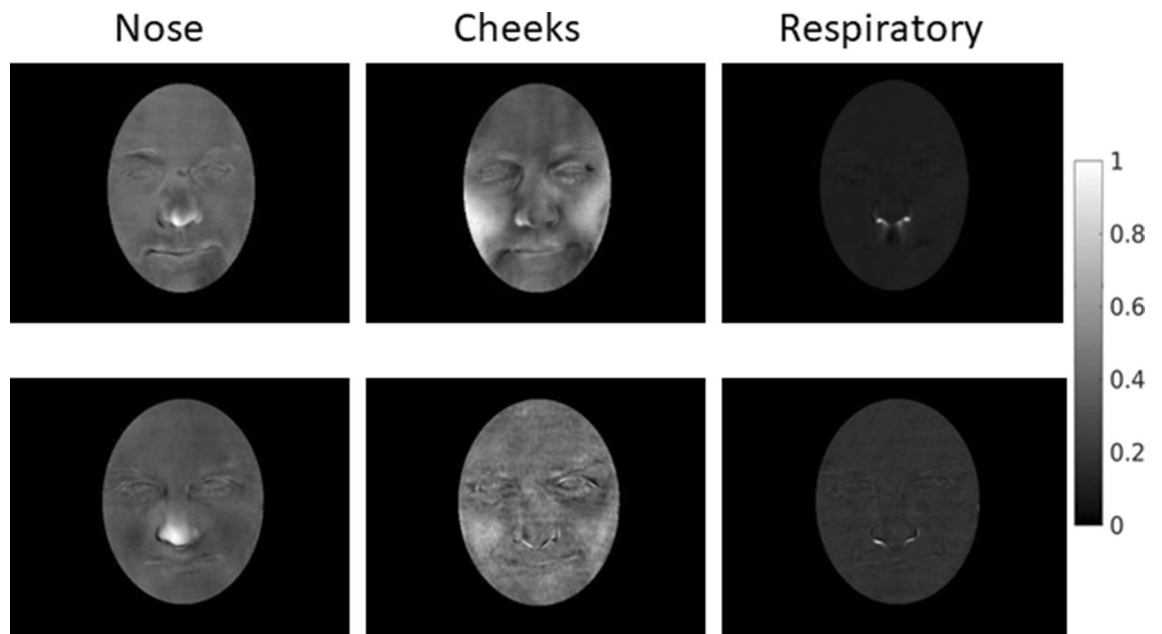


Figure 2. Distinct spatial components. Representative components from one subject (L08) are shown. Three components were consistently identified in all subjects (except respiratory component absent in one subject). Color scale normalized between 0 and 1 for each component for display. Results from two subjects are shown here.

through the Microsoft CNTK Face API (Cognitive Toolkit)⁴². This software has been rated as one of the best for emotion categorisation⁴³. The software takes a facial image as an input and returns the likelihood of each emotion class, across a pre-specified set of emotions, for each face in the image. The emotion classes permissible are anger, contempt, disgust, fear, happiness, neutral, sadness and surprise. They are ranked in descending order, with the overall total score summing to 1. The score was averaged for all the frames and averaged for each second. Briefly, emotion recognition is usually based on three key steps: (1) face detection, (2) feature detection—such as eyes and eye corners, brows, mouth corners, the nose-tip etc. and (3) feature classification—translation of the features into Action Unit codes, emotional states, and other affective metrics⁴⁴. We summed the emotion scores for each emotion category for each subject before undertaking inter-subject correlation analysis to investigate the consistency of facial expressions among subjects.

Results

Subjective ratings (SI Table S1, Fig. S1) showed that the participants perceived the movie to be positively valenced (Table S1 Question 1: mean = 5.95, SD = 1.43) as well as emotionally intense (Table S1 Question 2: mean = 5.53, SD = 1.26).

Consistent facial components identified by sICA. We extracted fifty independent components (ICs) from each subject. Of these, three components could be consistently identified across all subjects based on spatial and temporal characteristics. These were a nose component which mainly included pixels on the nose-tip, a cheek component which mainly contained pixels distributed on bilateral cheeks, and a respiratory component. The spatial patterns of these components clearly follow anatomical boundaries of facial features, supporting that they are driven by underlying physiological processes (Fig. 2).

The temporal courses of these three ICs showed gradual evolution over the course of the movie (Fig. 3 left). To investigate whether the nose and the cheek ICs identified were distinct from the respiratory component, we compared the power spectra of these signals (Fig. 3 right). A frequency range of 0.16–0.35 Hz is indicative of normal respiration frequency⁴⁵. The nose IC was minimally affected by respiration whereas the power spectra of both the cheek and the respiration IC exhibited a peak at frequency range of 0.16–0.35 Hz. The similar time courses and power spectra of the bilateral cheeks to that of the respiration IC is consistent with the cheek IC's diffuse spatial spread which likely includes respiration affected signals.

Validation of sICA nose component by ROI tracking method. To validate the ICA-based method, we compared the time courses of the nose IC with those of thermal signals obtained using a ROI located at the nose tip (Fig. 4A). Note that for thermal signal extraction with ROI method, we used the same motion correction algorithm based on optical flow. To our knowledge no other facial thermal imaging study has used an optical flow-based motion correction algorithm. The similarity between signals from the two methods was assessed by evaluating the group average Pearson correlation coefficient of nose IC and the thermal signal extracted from the corresponding ROI of the subject and statistically tested with non-parametric permutation testing (Fig. 4B) (see Methods) ($r = 0.94$, $p < 10^{-7}$, SI Fig. S2 left).

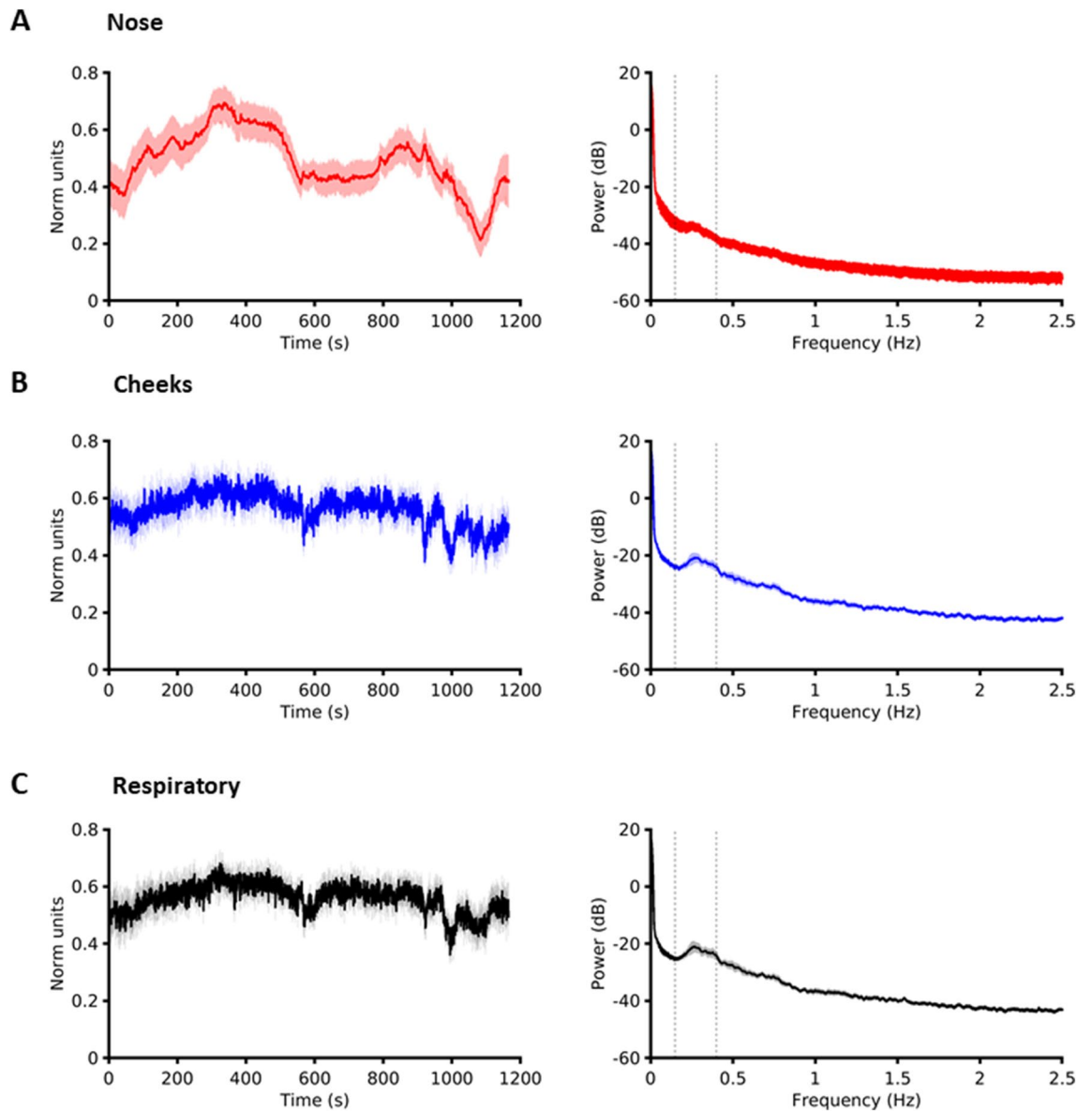


Figure 3. Group averaged component signals and their spectral signatures. (A) nose component signal and its power spectra (right). (B) bilateral cheek component signal and its power spectra (right). (C) respiratory component signal and its power spectra (right). Vertical dashed lines on the spectral plots indicates the normal respiratory frequency range of 0.16–0.35 Hz. Respiratory and cheek components both seem to be affected by respiration whereas the nose component seems minimally affected by it. Shading indicates SEM.

Inverse relation between nose component signals and GSR. GSR has been extensively validated as an arousal measure in psychophysiological studies². GSR thus provides a valuable benchmark to compare other psychophysiological measures and for facial thermal imaging this has previously been done, albeit for short event-based stimuli^{11,39}. We thus compared the time courses of the nose component with that of corresponding GSR signals to further investigate their response relationship during naturalistic emotional experiences. Permutation testing showed a significant inverse relationship ($r = -0.18$, $p = 0.03$) (Fig. 4C,D).

Similarity of physiological changes across subjects. We then examined the consistency of all the corresponding physiological variables across subjects using inter-subject correlation (ISC) analysis. We found significant consistency between subjects in the dynamic thermal response patterns for the nose IC (mean $r = 0.12$, $p = 0.007$, Fig. 5A), and still greater consistency for dynamic GSR (mean $r = 0.50$, $p < 0.0001$, Fig. 5B). The high ISC of GSR signals seemed to be mostly attributable to the gradual decrease of GSR at the beginning of the movie viewing—by contrast the ISC of GSR signals during the second half of the movie was 0.08 ($p = 0.03$), far lower than the ISC coefficient value for the thermal signal in the second half of the movie of 0.14 ($p = 0.003$). ISC of frequency components of heart rate variability (HRV) also showed significant inter-subject correlation (low

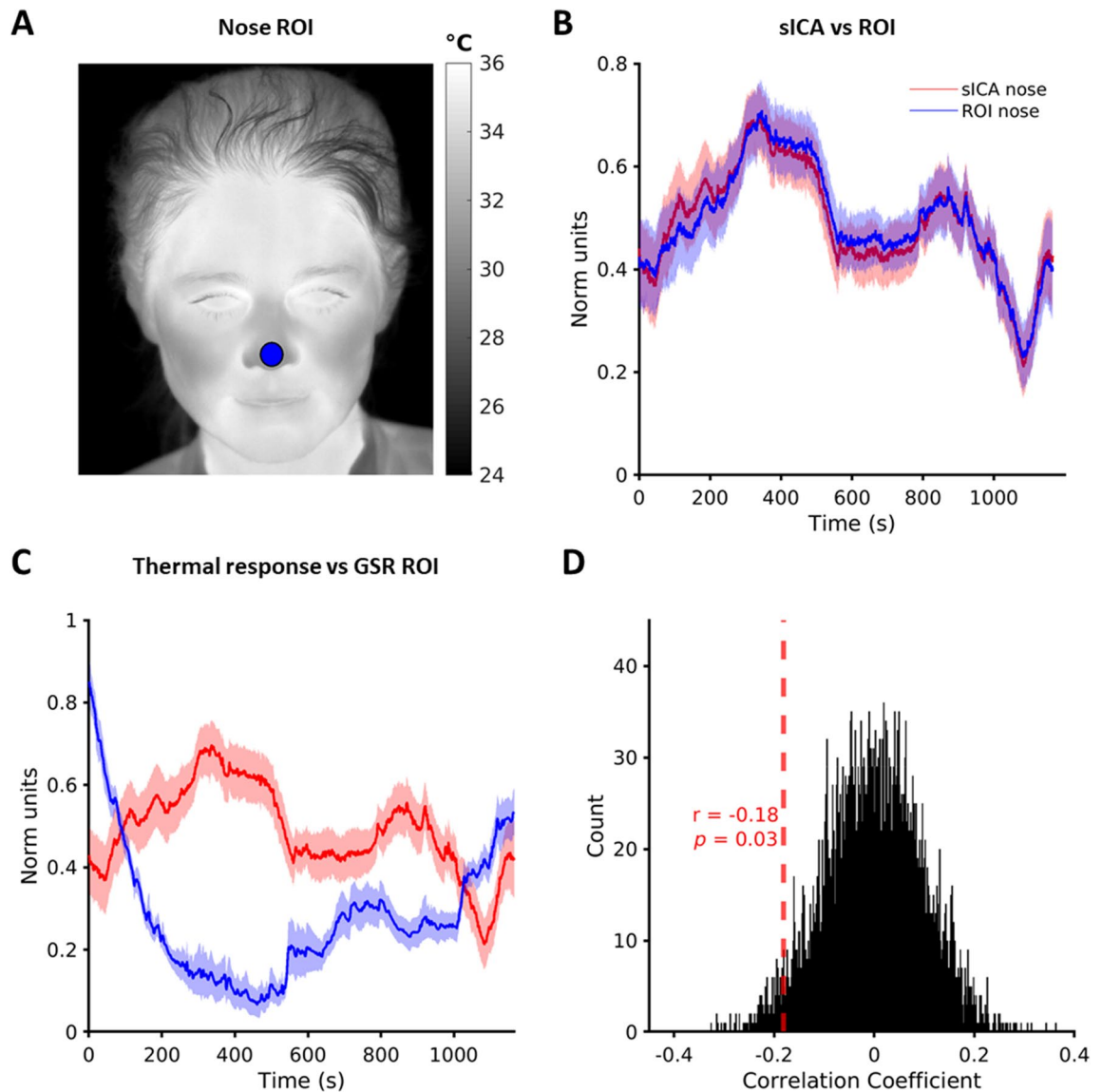


Figure 4. Nose component signal validation by comparison to ROI (region of interest) method and to GSR. (A) ROI location on motion-corrected thermal image of a subject (L08) with nine-pixel radius. (B) Comparison of group average nose sICA thermal component (red) with that of group average thermal signal obtained by ROI method (blue). (C) Group average comparison of thermal response obtained from nose sICA component (red) to GSR (blue). Shading indicates SEM. (D) Null distribution obtained with 5000 permutations showing statistical significance of negative correlation between thermal response and GSR.

frequency HRV: mean $r = 0.03$, $p = 0.0002$, Fig. 5C left, high frequency HRV: mean $r = 0.03$, $p = 0.00004$, Fig. 5C right, $p_{FDR} = 0.05$).

Thermal correlates of emotions depicted in the movie. With an exploratory approach, we evaluated whether thermal responses were correlated with the facial emotional scores of the actors in the movie frames. Previous studies, using an event-related design, have found the latency of thermal responses to be around 3.8 s⁴⁶ and 4–5 s¹⁵ after the stimulus onset. In this study we employed a naturalistic paradigm with continuous narrative instead of discrete events, hence, to compare the emotions of the movie and thermal signals we correlated a 3 s lagged thermal signal with the movie emotion time series. The emotional content of the movie was quantified by computing the scores of facial emotional expressions of the actors in each frame of the movie averaged for each second (see Methods). Thus, each second of the movie was assigned a facial emotional score according to the mean emotion detected from those frames (happiness, sadness, surprise, anger, fear, disgust and contempt), excluding the neutral category (SI Fig. S3A). A non-parametric permutation test revealed a significant negative correlation between the thermal signal and happiness scores ($r = -0.06$, $p_{FDR} = 0.0004$, SI Fig. S3B) and a significant positive correlation between thermal signal and anger scores ($r = 0.05$, $p_{FDR} = 0.004$, SI Fig. S3B). These results are also replicable when using a 4 s lagged thermal signals (happy: $r = -0.06$, $p_{FDR} = 0.0001$, anger $r = 0.05$,

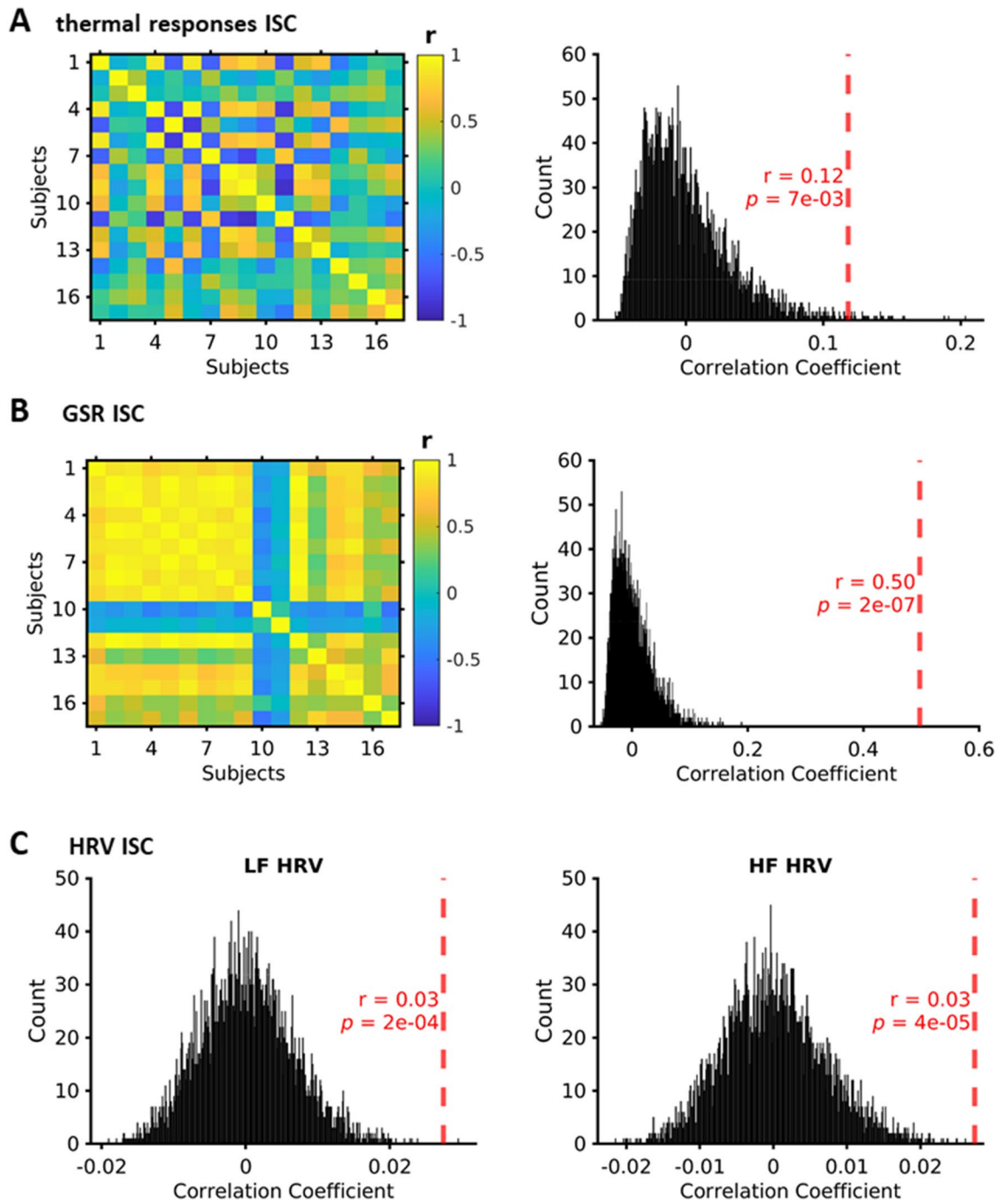


Figure 5. Inter-subject correlation (ISC) analysis. (A) correlation matrix of nose component thermal response time series (left) and null distribution obtained with 5000 permutations (right) (see Methods) showing statistical significance of positive mean correlation shown in red. (B) correlation matrix of GSR responses curves (left) and right—null distribution obtained with 5000 permutations (right) (see Methods) showing statistical significance of positive mean correlation shown in red (C) ISC analysis of low frequency heart rate variability (LF HRV) and its statistical significance (shown in red) (left) across participants when tested with 5000 permutations. ISC analysis for high frequency (HF) HRV (right). FDR corrected for statistical comparisons for LF and HF HRV. Colour bar r denotes Pearson's correlation coefficient. Histogram r denotes mean correlation coefficient at group level.

$p_{\text{FDR}} = 0.004$). Furthermore, the results are replicable without introducing lag with (happy: $r = 0.06$, $p_{\text{FDR}} = 0.0004$; anger: $r = 0.05$, $p_{\text{FDR}} = 0.004$).

Discussion

We sought to develop a new data-driven method for facial thermal imaging analysis and introduced novel application of optic flow for motion correction and that of sICA. We validated our method with traditional ROI analysis showing the robustness of our approach. We further demonstrated common physiological responses across subjects as measured by thermal signals, heart rate metrics and GSR, underlining common changes in emotional states induced by the naturalistic stimulus. Crucially these results were obtained in an ecologically valid context.

We restricted our analysis to the nose tip avoiding analysing the cheek and respiratory components since those signals shared what appeared to be noisy respiratory related signatures. The spatial components corresponding to the cheeks were quite broad, and often encroached regions close to lateral aspects of the philtrum. Thus, it is likely that larger regions may have captured respiratory signals in some participants. By contrast, the nose was a distinct and spatially constrained component in all subjects without any overlap of philtrum or cheeks. As a proof-of-concept study, with limited sample size, we wanted to prevent potential multiple comparisons by focusing subsequent analysis on the nose component only. We validated our methodology by comparison to the traditionally employed ROI method and demonstrated remarkably similar results for nose responses with both these methods. The robust correspondence of results with the two methods can be due to state-of-the-art motion correction technique employed. As far as we are aware no other thermal imaging study has utilized optical flow for motion correction of facial thermal imaging. Existing thermal imaging studies employing motion correction have used spatial cross-correlation^{12,14} but these are of limited effectiveness when used with a 3D surface that is not flat, such as the human face. The successful applications of optical flow and spatial ICA in this work thus provide two separate novel methodological contributions, as the former can be used to provide motion correction for ROI based methods as well.

The spatial components of nose and cheeks identified in this study correspond to facial regions widely studied in previous studies^{7,11,39,46,47}. The forehead has also shown stimulus induced temperature changes in prior studies^{46,48,49}. While some subjects exhibited a distinct forehead spatial component, this was not consistently present in all subjects. In addition to nose and cheeks, a respiration related component was found consistently in all subjects. Respiration monitoring via thermal imaging has already been demonstrated^{50,51}. The different facial components were associated with distinct thermal responses. Specifically, the cheek component and respiratory component showed similar time courses and with both being affected by respiration noise. Moreover, the spatial maps of the cheek component showed a diffuse distribution and hence was likely noisier. The successful application of motion correction and spatial ICA in this work demonstrates the utility of this approach for extracting robust thermal signals in longer-time naturalistic paradigms. A promising avenue of future research will be to use spatio-temporal ICA in place of spatial ICA^{52,53} to attempt to separate sympathetic and parasympathetic influences in the same facial regions. Specifically, spatial ICA, as used in the present paper, seeks spatial maps that are maximally independent, so it cannot separate signals that have different causes but manifest in the same facial region, e.g., separate effects of sympathetic and parasympathetic system on blood flow. Temporal ICA separates signals that have independent time courses but has no constraint on spatial maps. Spatiotemporal ICA is a trade-off between these two extremes, so may be useful in separating underlying source signals that are not fully independent in either time or space.

Facial blood flow changes detected by thermal imaging are caused not only by sympathetic (predominantly) but also parasympathetic influences^{54,55}. Thus, it is a more complex signal in comparison to GSR which is a uni-dimensional sympathetic response representing arousal^{56,57}. Thermal signals thus seem to convey more nuanced representation of sympathetic-parasympathetic balance. The more complex nature of thermal signals may also contribute to greater subject to subject variability and may explain the unexpected anti-correlations between a few subjects (Fig. 5B). GSR signals, due to their reliability, are most widely used in peripheral neurophysiological and psycho-physiological studies, thus providing an optimal benchmark for validation in IRI research⁵⁸. Overall, these results further highlight the biological underpinnings that might make thermal responses useful in differentiating emotional valence.

All the physiological variables (namely thermal imaging, heart rate metrics and GSR) showed significant similarity among subjects, thus further advocating for its use to induce common emotional states. Naturalistic stimuli offer a trade-off between completely uncontrolled stimuli and unconstrained conditions on the one hand (for instance resting state paradigms in neuroimaging), and the strict control of simplistic stimuli (for instance auditory odd-ball paradigm), placing relevant ecological constraints on physiological processes²³. As emotion is built over a longer narrative, these stimuli can simulate common, everyday emotional experiences and evoke robust and consistent responses in subjects as evident from our inter-subject correlation analysis. These analytical methods of inter-subject correlation analysis, when applied to neuroimaging, have also shown shared neuronal processes underlying emotional states^{22,27,41}. Thus, the consistent responses across subjects reported in our study for diverse physiological variables provide convergence of findings on a body level, which at the brain level have already been shown^{22,26,40}. This reinforces the wider use of naturalistic paradigms to comprehensively investigate human emotions. In similar vein however, we did not find a temporal consistency of facial expressions among subjects (SI Fig. S4). Facial emotion displays encountered in everyday life situations show high variability^{59,60} and spontaneous expressive behaviour is more complex and ambiguous (Pantic, 2009). This was true in our subjects as well, as they had low emotion expression scores.

Limitations of the study

Some caveats and limitations of this study should be noted. The three consistent ICA components were identified manually based on visual inspection. An objective automated component selection method can help overcome this limitation in the future⁶¹. Furthermore, we restricted the number of ICs to fifty, however the spatial components were consistent even when this was increased to eighty (SI Fig. S5). In that case, the bilateral cheek component was often decomposed in separate individual components. Future work could more deeply investigate the spatial patterns of the components, such as their variation across participants, and whether they change in spatial extent (and not just expression) according to different affective responses. Motion correction of facial data has inherent challenges due to the 2D nature of the data while the actual face is a 3D object. We obtained motion correction using 2D dense optical flow, while future work may explicitly map 2D thermal data to a 3D facial model to achieve more precise motion correction. A common standard space for registration of all subjects' faces will also improve the spatial decomposition. We did not analyse the cheek component signal as the power spectral signature was similar to that of the respiration component. However, we cannot be sure that cheek component signal is devoid of autonomic activation and given these components covered broad areas, we omitted this from analyses. Future studies can delineate these differences. Cross-validation of neuroimaging applications of ISC suggests that to ensure reliability for fMRI data, sample sizes should be $N \geq 20$ ⁶². Two things are worth noting here—first, that (outside of the very strong responses in V1), ISC in fMRI is typically lower than the ISC we observe in thermal responses (compare Fig. 5A with Table 2⁶²). The dice coefficient used in Pajula & Tohka encompasses the reliability of all suprathreshold voxels and hence reflects the weaker ISC in e.g. the insula (see their Fig. 1). Second, in fMRI, ISC amongst ICA components is stronger than in voxel space—mirroring the results here. Therefore, while $N = 20$ is likely a useful rule-of-thumb in fMRI, it is unlikely to generalize exactly to thermal responses when using sICA. We also note that the average sample size of thermal imaging is $17.9 \pm \text{SEM of } 3.1$ ⁴. Our sample size is consistent with these studies. Nonetheless, and despite conservative statistical correction, although our results are significant, the effect sizes are relatively small. Therefore, as with fMRI, future studies using ISC and thermal responses should use larger samples, particularly if aiming to covary against phenotypic or other independent measures. Future studies could compliment naturalistic stimulus with separately acquired continuous subjective ratings of valence and arousal to further probe temperature perturbations dependence on emotions. Lastly, challenges regarding experimental conditions remain such as the tight experimental control over ambient temperature and humidity.

In sum, we developed a novel data-driven analytical technique for facial thermal imaging analysis and showed robust facial thermal changes evoked by an emotional movie which were correlated with the emotional content of facial expressions in the movie. We also uncovered a physiological consistency among subjects thus signifying common responses elicited by the movie. The effect sizes observed here although modest provide a helpful guide for power analyses in future studies. Future studies could also use simultaneous thermal and brain recordings such as combined facial thermal imaging with functional near infrared spectroscopy¹² or electroencephalography (EEG), to investigate the brain-body interaction.

Data and code availability

Code related to this paper is available from the authors. Data used in the study includes personal identifying facial images of participants, and local ethics approval mandated strict privacy restrictions around their availability outside of the named investigator team. Researchers wishing to access these data will require local ethics approval and a data sharing agreement with QIMR Berghofer.

Received: 6 August 2020; Accepted: 24 May 2021

Published online: 08 June 2021

References

- Cacioppo, J. T. & Tassinary, L. G. *Principles of Psychophysiology: Physical, Social, and Inferential Elements* (Cambridge University Press, Cambridge, 1990).
- Kreibig, S. D. Autonomic nervous system activity in emotion: A review. *Biol. Psychol.* **84**(3), 394–421 (2010).
- Rainville, P. *et al.* Basic emotions are associated with distinct patterns of cardiorespiratory activity. *Int. J. Psychophysiol.* **61**(1), 5–18 (2006).
- Ioannou, S., Gallese, V. & Merla, A. Thermal infrared imaging in psychophysiology: Potentialities and limits. *Psychophysiology* **51**(10), 951–963 (2014).
- Jarlier, S. *et al.* Thermal analysis of facial muscles contractions. *IEEE Trans. Affect. Comput.* **2**(1), 2–9 (2011).
- Srivastava, N. *et al.* Exploring the usage of thermal imaging for understanding video lecture designs and students' experiences. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 250–259). (2020, March).
- Ebisch, S. J. *et al.* Mother and child in synchrony: Thermal facial imprints of autonomic contagion. *Biol Psychol.* **89**(1), 123–129 (2012).
- Engert, V. *et al.* Exploring the use of thermal infrared imaging in human stress research. *PLoS One* **9**(3), e90782 (2014).
- Ioannou, S. *et al.* The autonomic signature of guilt in children: A thermal infrared imaging study. *PLoS One* **8**(11), e79440 (2013).
- Kuraoka, K. & Nakamura, K. The use of nasal skin temperature measurements in studying emotion in macaque monkeys. *Physiol. Behav.* **102**(3), 347–355 (2011).
- Pavlidis, I., Levine, J., & Baukol, P. Thermal image analysis for anxiety detection. In *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)* (Vol. 2, pp. 315–318). IEEE. (2001, October).
- Pinti, P., Cardone, D. & Merla, A. Simultaneous fNIRS and thermal infrared imaging during cognitive task reveal autonomic correlates of prefrontal cortex activity. *Sci. Rep.* **5**, 17471 (2015).
- Strakowska, M., Strakowski, R., Wiecek, B., & Strzelecki, M. Cross-correlation based movement correction method for biomedical dynamic infrared imaging. In *Proc. of the 11th International Conference on Quantitative InfraRed Thermography, Naples, Italy.* (2012, June).
- Manini, B. *et al.* Mom feels what her child feels: Thermal signatures of vicarious autonomic response while watching children in a stressful situation. *Front. Hum. Neurosci.* **7**, 299 (2013).

15. Sonkusare, S. *et al.* Detecting changes in facial temperature induced by a sudden auditory stimulus based on deep learning-assisted face tracking. *Sci. Rep.* **9**(1), 1–11 (2019).
16. Dowdall, J., Pavlidis, I. T. & Tsiamyrtzis, P. Coalitional tracking. *Comput. Vis. Image Underst.* **106**(2–3), 205–219 (2007).
17. Paolini, D. *et al.* “The face of ostracism”: The impact of the social categorization on the thermal facial responses of the target and the observer. *Acta Physiol. (Oxf.)* **163**, 65–73 (2016).
18. Derakhshan, A. *et al.* Identifying the optimal features in multimodal deception detection. *Multimodal Technol. Interact.* **4**(2), 25 (2020).
19. McKeown, M. J., Hansen, L. K. & Sejnowski, T. J. Independent component analysis of functional MRI: What is signal and what is noise?. *Curr. Opin. Neurobiol.* **13**(5), 620–629 (2003).
20. Beckmann, C. F. *et al.* Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. B Biol. Sci.* **360**(1457), 1001–1013 (2005).
21. Calhoun, V. D., Adali, T., Hansen, L. K., Larsen, J., & Pekar, J. J. ICA of functional MRI data: an overview. In *Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation.* (2003).
22. Hasson, U. *et al.* Intersubject synchronization of cortical activity during natural vision. *Science.* **303**(5664), 1634–1640 (2004).
23. Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends in cognitive sciences*, **23**(8), 699–714.
24. Gross, J. J. & Levenson, R. W. Emotion elicitation using films. *Cogn. Emot.* **9**(1), 87–108 (1995).
25. Ioannou, S. *et al.* Sympathy crying: Insights from infrared thermal imaging on a female sample. *PLoS One* **11**(10), e0162749 (2016).
26. Nguyen, V. T. *et al.* Distinct cerebellar contributions to cognitive-perceptual dynamics during natural viewing. *Cereb. Cortex* **27**(12), 5652–5662 (2016).
27. Pajula, J., Kauppi, J.-P. & Tohka, J. Inter-subject correlation in fMRI: Method validation against stimulus-model based analysis. *PLoS One* **7**(8), e41196 (2012).
28. Weigel, J., Williams, T., & Weigel, R. *The Butterfly circus.* Evolution Entertainment. (2009).
29. Weinzaepfel, P., Revaud, J., Harchaoui, Z., & Schmid, C. DeepFlow: Large displacement optical flow with deep matching. In *Proceedings of the IEEE international conference on computer vision* (pp. 1385–1392). (2013).
30. Ahmad, J. *et al.* Barker-coded thermal wave imaging for non-destructive testing and evaluation of steel material. *IEEE Sens. J.* **19**(2), 735–742 (2018).
31. Hyvarinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* **10**(3), 626–634 (1999).
32. Allen, J. J., Chambers, A. S. & Towers, D. N. The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biol. Psychol.* **74**(2), 243–262 (2007).
33. Ramshur, J. T. *Design, evaluation, and application of heart rate variability analysis software (HRVAS)* (Doctoral dissertation, University of Memphis). (2010).
34. Akselrod, S., *et al.*, Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Science* **220**–222 (1981).
35. Pomeranz, B. *et al.* Assessment of autonomic function in humans by heart rate spectral analysis. *Am. J. Physiol. Heart Circ. Physiol.* **248**(1), H151–H153 (1985).
36. Berntson, G. G., & Cacioppo, J. T. Heart rate variability: Stress and psychiatric conditions. *Dynamic electrocardiography*, **41**(2), 57–64. (2004).
37. Thomas, B. L. *et al.* Validity of commonly used heart rate variability markers of autonomic nervous system function. *Neuropsychobiology* **78**(1), 14–26 (2019).
38. Bernardi, L. *et al.* Modulatory effects of respiration. *Auton. Neurosci.* **90**(1–2), 47–56 (2001).
39. Shastri, D. *et al.* Imaging facial signs of neurophysiological responses. *IEEE Trans. Biomed. Eng.* **56**(2), 477–484 (2009).
40. Kauppi, J.-P. *et al.* Inter-subject correlation of brain hemodynamic responses during watching a movie: Localization in space and frequency. *Front. Neuroinform.* **4**, 5 (2010).
41. Nguyen, V. T. *et al.* The integration of the internal and external milieu in the insula during dynamic emotional experiences. *Neuroimage* **124**, 455–463 (2016).
42. Del Sole, A. Introducing microsoft cognitive services. In *Microsoft Computer Vision APIs Distilled.* 1–4 (Springer, 2018).
43. Khanal, S. R., Barroso, J., Lopes, N., Sampaio, J., & Filipe, V. Performance analysis of microsoft’s and google’s emotion recognition api using pose-invariant faces. In *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion* (pp. 172–178). (2018, June).
44. Cowie, R. *et al.* Emotion recognition in human-computer interaction. *IEEE Signal Process. Mag.* **18**(1), 32–80 (2001).
45. Russo, M. A., Santarelli, D. M. & O’Rourke, D. The physiological effects of slow breathing in the healthy human. *Breathe* **13**(4), 298–309 (2017).
46. Merla, A., & Romani, G. L. Thermal signatures of emotional arousal: a functional infrared imaging study. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 247–249). IEEE. (2007, August).
47. Nakanishi, R. & Imai-Matsumura, K. Facial skin temperature decreases in infants with joyful expression. *Infant Behav. Dev.* **31**(1), 137–144 (2008).
48. Puri, C., Olson, L., Pavlidis, I., Levine, J., & Starren, J. StressCam: non-contact measurement of users’ emotional states through thermal imaging. In *CHI’05 extended abstracts on Human factors in computing systems* (pp. 1725–1728). (2005, April).
49. Zhu, Z., Tsiamyrtzis, P., & Pavlidis, I. Forehead thermal signature extraction in lie detection. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 243–246). IEEE. (2007, August).
50. Cho, Y., Bianchi-Berthouze, N., & Julier, S. J. DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 456–463). IEEE. (2017, October).
51. Cho, Y. *et al.* Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging. *Biomed. Opt. Express* **8**(10), 4480–4503 (2017).
52. Stone, J. V., Porrill, J., Buchel, C., & Friston, K. Spatial, temporal, and spatiotemporal independent component analysis of fMRI data. In *Proc. Leeds Statistical Research Workshop* (pp. 7–9). (1999, July).
53. Stone, J. *et al.* Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. *Neuroimage* **15**(2), 407–421 (2002).
54. Drummond, P. D. Sweating and vascular responses in the face: Normal regulation and dysfunction in migraine, cluster headache and harlequin syndrome. *Clin. Auton. Res.* **4**(5), 273–285 (1994).
55. Segade, L. A. & Sua, J. Distribution of postganglionic parasympathetic fibers originating in the pterygopalatine ganglion in the maxillary and ophthalmic nerve branches of the trigeminal nerve; HRP and WGA-HRP study in the guinea pig. *Brain Res.* **522**(2), 327–332 (1990).
56. Bach, D. R., Friston, K. J. & Dolan, R. J. Analytic measures for quantification of arousal from spontaneous skin conductance fluctuations. *Int. J. Psychophysiol.* **76**(1), 52–55 (2010).
57. Boucsein, W. & Hoffmann, G. A direct comparison of the skin conductance and skin resistance methods. *Psychophysiology* **16**(1), 66–70 (1979).
58. Boucsein, W. *Electrodermal Activity* (Springer, New York, 2012).

59. Nummenmaa, L. & Calvo, M. G. Dissociation between recognition and detection advantage for facial expressions: A meta-analysis. *Emotion* **15**(2), 243 (2015).
60. Scherer, K. R. & Ellgring, H. Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal?. *Emotion* **7**(1), 113 (2007).
61. Salimi-Khorshidi, G. *et al.* Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* **90**, 449–468 (2014).
62. Pajula, J., & Tohka, J. How many is enough? Effect of sample size in inter-subject correlation analysis of fMRI. *Comput. Intell. Neurosci.* (2016).

Author contributions

Conceptualization, C.G., S.S., V.N.; Data Curation, S.S., V.N.; Methodology, M.J.A., T.P., S.S.; Investigation, S.S., M.J.A., T.P.; Writing—Original Draft, S.S., T.P., Writing—Review & Editing, M.J.A., S.S., V.N., T.P., M.B., S.F., C.G.; Funding Acquisition, M.B., C.G.; Resources, S.F., M.B.; Supervision, M.J.A., C.G., M.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-91578-5>.

Correspondence and requests for materials should be addressed to S.S. or M.J.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2021