NeuroImage 247 (2022) 118800

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Classification of emotion categories based on functional connectivity patterns of the human brain ☆

Heini Saarimäki [a,b,*], Enrico Glerean [b,c,d,e,f], Dmitry Smirnov [b], Henri Mynttinen [b], Iiro P. Jääskeläinen [b,f], Mikko Sams [b,e], Lauri Nummenmaa [d]

[a] Faculty of Social Sciences, Tampere University, FI-33014 Tampere University, Tampere, Finland
[b] Department of Neuroscience and Biomedical Engineering, School of Science, Aalto University, Espoo, Finland
[c] Advanced Magnetic Imaging (AMI) Centre, Aalto NeuroImaging, School of Science, Aalto University, Espoo, Finland
[d] Turku PET Centre and Department of Psychology, University of Turku, Turku, Finland
[e] Department of Computer Science, School of Science, Aalto University, Espoo, Finland
[f] International Laboratory of Social Neurobiology, Institute for Cognitive Neuroscience, HSE University, Moscow, Russian Federation

ARTICLE INFO

Keywords:
fMRI
Functional connectivity
Emotion
Pattern classification
MVPA

ABSTRACT

Neurophysiological and psychological models posit that emotions depend on connections across wide-spread corticolimbic circuits. While previous studies using pattern recognition on neuroimaging data have shown differences between various discrete emotions in brain activity patterns, less is known about the differences in functional connectivity. Thus, we employed multivariate pattern analysis on functional magnetic resonance imaging data (i) to develop a pipeline for applying pattern recognition in functional connectivity data, and (ii) to test whether connectivity patterns differ across emotion categories. Six emotions (anger, fear, disgust, happiness, sadness, and surprise) and a neutral state were induced in 16 participants using one-minute-long emotional narratives with natural prosody while brain activity was measured with functional magnetic resonance imaging (fMRI). We computed emotion-wise connectivity matrices both for whole-brain connections and for 10 previously defined functionally connected brain subnetworks and trained an across-participant classifier to categorize the emotional states based on whole-brain data and for each subnetwork separately. The whole-brain classifier performed above chance level with all emotions except sadness, suggesting that different emotions are characterized by differences in large-scale connectivity patterns. When focusing on the connectivity within the 10 subnetworks, classification was successful within the default mode system and for all emotions. We thus show preliminary evidence for consistently different sustained functional connectivity patterns for instances of emotion categories particularly within the default mode system.

1. Introduction

Our emotional experiences cluster around emotion categories. Emotional experiences, or feelings, refer to the subjectively felt part of individual's emotional state (Adolphs, 2017). Each emotion category describes some unified, shared features of the current emotional state and experience, even though the boundaries between the exemplars within a category (and also between categories) are fuzzy (e.g., Cowen and Keltner, 2017; Horikawa et al., 2020). Emotion concepts such as *fear, joy*, or *pride* are often used to label the categories as a summary of the shared features within a category. However, the exact features that define category boundaries at behavioral, physiological, and neural level remain elusive.

Current neurophysiological and psychological models of emotions largely agree that emotions are characterized by large-scale changes in brain activity spanning both cortical and subcortical areas (see, e.g., Hamann, 2012; Lindquist et al., 2012; Pessoa, 2012; Scarantino, 2012; Adolphs, 2017; Barrett, 2017). Accumulating evidence from both patient and neuroimaging studies in humans shows that different brain regions support different aspects of emotional processing, corresponding to functionally distinct components rather than to specific emotions *per se* (for reviews, see, e.g., Hamann, 2012; Feinstein, 2013; Kragel and LaBar, 2016; Sander et al., 2018; Nummenmaa and Saarimäki, 2019; Saarimäki, 2021). Accordingly, theoretical frameworks also emphasize the importance of networks of cortical and subcortical regions underlying the representation of emotion categories in the brain (Pessoa, 2017; Barrett and Satpute, 2013; Barrett, 2017).

☆ The authors declare no competing conflict of interest.

* Corresponding author at: Faculty of Social Sciences, Tampere University, FI-33014 Tampere University, Tampere, Finland.
E-mail address: heini.saarimaki@tuni.fi (H. Saarimäki).

https://doi.org/10.1016/j.neuroimage.2021.118800.
Received 20 January 2021; Received in revised form 5 December 2021; Accepted 8 December 2021
Available online 9 December 2021.

Multiple lines of evidence suggest that different emotion categories involve distributed patterns of behavior, physiology, and neural activation. Instances of different emotion categories have been successfully classified from functional components including peripheral physiological responses and bodily activation patterns (Kragel and Labar, 2015; Kreibig et al., 2007; Nummenmaa et al., 2014a; Hietanen et al., 2016), and subjective feelings (Saarimäki et al., 2016, 2018), suggesting that different emotion categories have distinguishable characteristics in behavior, bodily sensations, and subjective experience (see also review in Nummenmaa and Saarimäki, 2019). At the neural level, multivariate pattern classification studies have revealed that discrete, distributed activity patterns in the cortical midline and somatomotor regions, and subcortical regions such as amygdala and thalamus, underlie different emotion categories (Chikazoe et al., 2014; Wager et al., 2015; Kragel and LaBar, 2015; Peelen et al., 2010; Saarimäki et al., 2016; 2018; Horikawa et al., 2020). Thus, it seems that different instances of the same emotion category share at least some underlying characteristics of neural activity that distinguish the emotions in one category from those in the other. Importantly, previous studies show that classification accuracy is consistently higher for the whole-brain activity patterns than for any single region of interest alone (e.g. Saarimäki et al., 2016; 2018). This suggests that while emotions cannot always be distinguished from each other based on changes in a single functional component, such as changes in a single autonomic nervous system (Siegel et al., 2018), integration of information from multiple component processes might underlie the experienced differences between emotion categories (Satpute and Lindquist, 2019).

Yet, as the brain functions as a hierarchy of networks (Bullmore and Sporns, 2009, 2012; Bassett and Sporns, 2017), it is possible that also the functional connectivity between different regions, rather than local activity patterns alone, vary between different emotions (Pessoa, 2017; Barrett and Satpute, 2013). In BOLD-fMRI studies, functional connectivity is defined as stimulus-dependent co-activation of different brain regions, usually measured as the correlation of the BOLD time series between two regions. Thus, functional connectivity does not necessarily reflect a true physical connection between two brain regions – two regions might not be structurally connected but still share the same temporal dynamics. For instance, for an emotion category such as *anger*, functional connectivity between two regions might reflect either true interaction between them or that the two regions follow the stimulation independently but with similar temporal dynamics. Given the promising results obtained with multivariate pattern analysis of brain activity patterns underlying different emotions, it has been suggested that the study of the brain basis of emotion would benefit from studying the functional relationship between brain regions by employing, for instance, machine learning on functional connectivity patterns (Pessoa, 2018). However, the applications of multivariate pattern analysis in functional networks are sparse. The few studies reporting classification of connectivity measured with fMRI have focused on mental or cognitive states (Richiardi et al., 2011; Shirer et al., 2012; Gonzalez-Castillo et al., 2015) or inter-individual differences in functional connectivity during rest and various tasks (Finn et al., 2015; Shen et al., 2017), demonstrating the plausibility of this approach.

Preliminary evidence highlighting the emotion-related changes in functional connectivity has shown decreased connectivity in specific brain networks after emotional stimulation containing negative affect (Bochardt et al., 2015). Also, dynamic valence- and arousal-related variation in the stimulus modulates functional connectivity (Nummenmaa et al., 2014a; Young et al., 2017). So far, only a handful of studies have compared how specific emotions modulate functional brain connectivity, and most of these studies have focused on a limited set of a-priori-defined regions-of-interest, including salience and amygdala networks (Eryilmaz et al., 2011; Tettamanti et al., 2012; Raz et al., 2016; Huang et al., 2018; Sachs et al., 2020) or investigated emotion-specific intrinsic connectivity (Touroutoglou et al., 2015). Sustained affective states such as sad mood and anxiety have been found to modulate functional connectivity between the core emotion-related areas such as midline regions including anterior and posterior cingulate cortices, orbitofrontal cortex, insula, and subcortical regions including thalamus and amygdala (Harrison et al., 2008; Seeley et al., 2007; Hermans et al., 2011, 2014; McMenamin et al., 2014). Connectivity especially in salience, frontoparietal control, attention and subcortical networks is altered in depression both during generation and regulation of negative emotions as well as during positive emotions (Siegle et al., 2007; Kaiser et al., 2015; Hasler and Northoff, 2011; Price et al., 2017). However, the differences in connectivity between emotion categories have not been directly compared before using large-scale brain networks. To our knowledge, only one study has investigated whole-brain-level connectivity changes during emotional stimuli, but instead of comparing different emotion categories, this study modeled valence- and arousal-dependent connectivity changes (Nummenmaa et al., 2014b). Accordingly, it remains unresolved whether emotion-specific connectivity patterns distinguish between emotions, and which networks might code for sustained changes in the overall emotional content of the stimulus.

The goals of the current study were two-fold. First, we aimed to demonstrate a pattern classification pipeline as a *proof-of-concept* for studying the whole-brain, fMRI-derived functional connectivity patterns underlying different subjective mental states. Second, given the previous studies that have successfully applied machine learning to brain activity patterns underlying instances of different emotions, we tested whether also the functional connectivity patterns during different emotional states could be separated from each other. To reach these goals, we induced emotions targeting six emotion categories (anger, fear, disgust, happiness, sadness, surprise) and a neutral state using auditory narratives during fMRI scanning. We focused on the sustained connectivity changes across each 1-minute-long narrative which reflects the overall emotional content of the narrative rather than the rapid, dynamic changes during the story. Brain activity related to each emotion was modeled as a functional network. We hypothesized that if instances of same emotion category share similar connectivity patterns, the classifier separates between the emotion categories reliably. The classification was performed with 264 nodes derived from a functional parcellation (Power et al., 2011) and for either whole-brain connectivity (i.e., all nodes together) or for within and between subnetwork connectivity (i.e., for links within and between each subnetwork separately).

2. Materials and methods

2.1. Participants

Sixteen female volunteers (ages 20–30, mean age 24.3 years) participated in the fMRI experiment. All participants were right-handed, healthy with normal or corrected-to-normal vision and gave written informed consent. The studies were run in accordance with the guidelines of the Declaration of Helsinki, and Research Ethics Committee of Aalto University approved the study protocol.

2.2. Stimuli

The stimuli were 35 one-minute-long narratives representing six emotional states (anger, fear, disgust, happiness, sadness, surprise) and a neutral state (five narratives per category in Finnish, transcribed English translations of the stimuli are available in Supplementary Table S4). The narratives described personal life events spoken by a female speaker with natural emotional prosody and included emotional expressions, such as weeping and laughing, and have been shown to elicit strong affect in listeners (Smirnov et al., 2019). Prior to the audio recording of the spoken narratives, the speakers were given the target emotion categories as well as five narrative topics per category and were instructed to develop and rehearse 1-minute-long narratives that described a corresponding autobiographical experience that elicited an instance of the

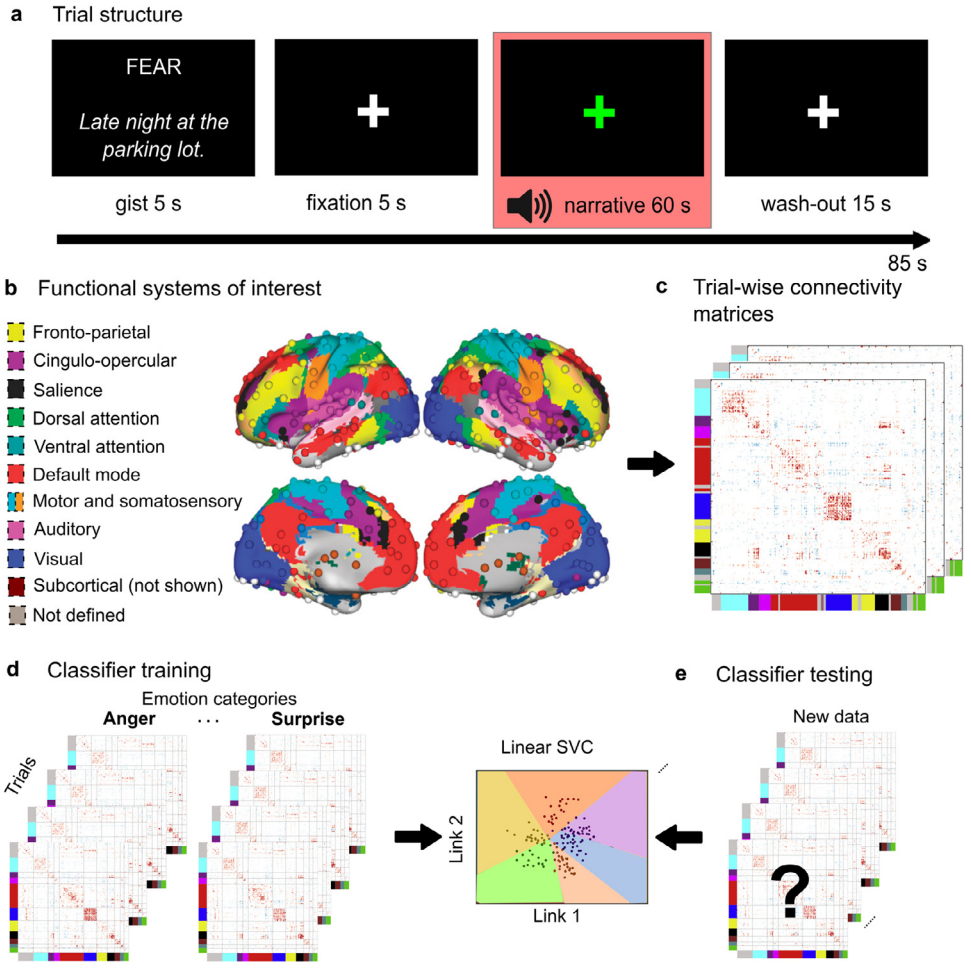

Fig. 1. **(a)** Trial structure. The highlighted time period (HRF-corrected) was used for calculating the connectivity matrices. **(b)** Functional brain systems analyzed in the present study, based on Power et al. (2011). Dots denote network nodes and colors denote subnetworks. **(c)** Connectivity matrices were calculated using Pearson correlation between each pair of 264 node time series for each subject and for each 60-s narrative. **(d)** The connectivity matrices were fed as input for a linear support vector classifier. **(e)** The classifier performance was evaluated by calculating the accuracy (percentage of correct classifier guesses per target category) and the confusion matrix (classifier guesses per category).

target emotion category. In a pilot study, a separate sample of twenty-four females (ages 20–37, mean age = 24.4 years) rated the stories for the experienced, categorical emotional content (Supplementary Fig. S1; Supplementary Table S1). Behavioral ratings confirmed that the stories successfully elicited the *a priori* defined target emotion; however, to strengthen the effect in the fMRI experiment, the participants also saw a word corresponding the target emotion prior to each story and a short description of the narrative gist (Fig. 1a). To ensure that the stimuli successfully elicited emotions during the scanning, the fMRI participants also rated the valence and arousal elicited by the stories (see below). However, as they were presented with the target emotion category during the fMRI scanning, we did not collect ratings for the emotion categories.

2.3. Experimental design

During fMRI, the stimuli were presented in five runs, with one narrative from each emotion category per run. Each run lasted for approximately 10 min (365 volumes) and consisted of seven trials (Fig. 1a). The order of trials within a run was the same for all participants and the order of runs was counterbalanced across participants (see Supplementary Fig. S2). Each narrative was thus presented only once during the experiment. A trial started with a fixation cross presented for 5 s, followed by a 5-s presentation of the target emotion (e.g. 'happiness') and a short description of the narrative gist (e.g. 'lovers under a tree'). Next, a fixation cross appeared on the screen and the narrative was played through earphones. The trial ended with a 15-s wash-out period allowing emotional state to recover to baseline level following each trial. Subjects were instructed to listen to the narratives similarly as if they would listen to their friend describing a personal life event. The auditory stimuli were delivered through Sensimetrics S14 insert earphones (Sensimetrics Corporation, Malden, Massachusetts, USA). Sound was adjusted for each subject to be loud enough to be heard over the scanner noise. The visual stimuli were back-projected on a semitransparent screen using a 3-micromirror data projector (Christie X3, Christie Digital Systems Ltd., Mönchengladbach, Germany) and from there via a mirror to the participant. Stimulus presentation was controlled with Presentation software (Neurobehavioral Systems Inc., Albany, CA, USA). After the scanning, participants listened to the narratives twice again using an on-line rating tool to continuously rate their subjectively experienced valence (ranging from unpleasant to pleasant) and arousal (ranging from calm to excited) during the narrative. Ratings were acquired post-experiment rather than during fMRI, as a reporting task is known to influence neural response to emotional stimulation (Hutcherson et al., 2005; Lieberman et al., 2007) and as repeating a specific emotional stimulus has only a negligible effect on self-reported emotional feelings (Hutcherson et al., 2005).

2.4. Psychophysiological recordings

To remove effects of heart rate and respiration from the BOLD signal during preprocessing, we successfully recorded heart rate and respiration data from 14 subjects with BIOPAC MP150 Data Acquisition System (BIOPAC System, Inc.). Heart rate was measured using BIOPAC TSD200 pulse plethysmogram transducer, which records the blood volume pulse waveform optically. The pulse transducer was placed on the palmar surface of the participant's left index finger. Respiratory movements were measured using BIOPAC TSD201 respiratory-effort transducer attached to an elastic respiratory belt, which was placed around

each participant's chest to measure changes in thoracic expansion and contraction during breathing. Both signals were sampled simultaneously at 1 kHz using RSP100C and PPG100C amplifiers for respiration and heart rate, respectively, and BIOPAC AcqKnowledge software (version 4.1.1). Respiration and heart rate signals were then used to extract and clean the time-varying heart and respiration rates out of the data with the DRIFTER toolbox (Särkkä et al., 2012).

2.5. MRI data acquisition and preprocessing

MRI data were collected on a 3T Siemens Magnetom Skyra scanner at the Advanced Magnetic Imaging centre, Aalto NeuroImaging, Aalto University, using a 20-channel Siemens volume coil. Whole-brain functional scans were collected using a whole brain T2*-weighted EPI sequence with the following parameters: 33 axial slices, interleaved order (odd slices first), TR = 1.7 s, TE = 24 ms, flip angle = 70°, voxel size = 3.1 × 3.1 × 4.0 mm^3, matrix size = 64 × 64 × 33, FOV 198.4 × 198.4 mm^2. A custom-modified bipolar water excitation radio frequency (RF) pulse was used to avoid signal from fat. High-resolution anatomical images with isotropic 1 × 1 × 1 mm^3 voxel size were collected using a T1-weighted MP-RAGE sequence. fMRI data were preprocessed using FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl) and inhouse MATLAB (The MathWorks, Inc., Natick, Massachusetts, USA, http://www.mathworks.com) tools (code available at: https://version.aalto.fi/gitlab/BML/bramila). Non-brain matter was removed from functional and anatomical images with FSL BET. After slice timing correction, the functional images were realigned to the middle scan by rigid-body transformations with MCFLIRT to correct subject's head motion. Next, DRIFTER was used to clean respiratory and heart rate signal from the data (Särkkä et al., 2012). Functional images were registered to the MNI152 standard space template with 2-mm resolution using FSL FLIRT two-step co-registration method with 9 degrees of freedom registration from subject's space EPI to subject's space anatomical volume, and 12 degrees of freedom from anatomical to MNI152 standard space. Removal of scanner trend was performed with a 240-seconds long cubic Savitzky-Golay filter (Çukur et al., 2013). To control for head motion artefacts, we followed the procedure as described in JD Power et al. (2014). The 6 motion parameters were expanded into 24 confound regressors and regressed out. Furthermore, signal at deep white matter, ventricles and cerebrospinal fluid were also regressed out. Finally, temporal bandpass filtering (0.01–0.08 Hz) was applied with the second-order butterworth filter. Spatial smoothing was done with a Gaussian kernel of FWHM 6 mm. All subsequent analyzes were performed with these preprocessed data.

2.6. Extracting the networks

Because working at voxel-by voxel time series would yield a very high dimensionality and thus be computationally prohibitive, the emotion-specific functional networks were estimated for 264 nodes based on the functional parcellation by Power et al. (2011). We extracted the BOLD time course for each node by averaging the activity of voxels within a 1-cm diameter sphere centered at each node's coordinates (list of coordinates and module assignments available at https://web.archive.org/web/20160127134525/http://www.nil.wustl.edu/labs/petersen/Resources_files/Consensus264.xls). For each of the 35 narratives, we calculated the Pearson correlation coefficient between the BOLD time course of each of the nodes during the 60-s-long story, which resulted in a connectivity matrix of 264 × 264 nodes for each narrative (Fig. 1b). Next, we removed the baseline connectivity pattern from emotion-wise connectivity matrices by taking the average of the five connectivity matrices for neutral narratives, and regressing it from each of the remaining 30 emotion-specific connectivity matrices separately.

Finally, in addition to the full network of 264 × 264 nodes, we extracted also subnetworks based on the 10 functional systems of interest as proposed by Power et al. (2011, 2014). The subnetworks included were motor and somatosensory (35 nodes), cingulo-opercular (14 nodes), auditory (13 nodes), default mode (58 nodes), visual (31 nodes), fronto-parietal (25 nodes), salience (18 nodes), subcortical (13 nodes), ventral attention (9 nodes), and dorsal attention (11 nodes) networks.

2.7. Correlation between connectivity matrices

To test the similarity between connectivity matrices of different emotions, we quantified the correlation between averaged emotion-wise connectivity matrices with Mantel test employing Spearman correlation as implemented by Glerean et al. (2016)(function bramila_mantel in https://github.com/eglerean/hfASDmodules/blob/master/ABIDE/bramila_mantel.m). P values were obtained with 5000 permutations (FDR corrected at $p<0.05$ for multiple comparisons).

2.8. Classification of connectivity matrices

The classification of emotion categories was performed in Python 2.7.11 (Python Software Foundation, http://www.python.org) using the Scikit learn package (Pedregosa et al., 2011). A between-subjects support vector machine classification algorithm with linear kernel was trained to recognize the correct emotion category out of 6 possible ones (anger, disgust, fear, happiness, sadness, surprise; Fig. 1c). Naïve chance level, derived as a ratio of 1 over the number of categories, was 16.6%. The samples for the classifier consisted of the 30 connectivity matrices (5 matrices for each emotion category) from each subject, resulting in altogether 480 samples (80 per category). A leave-one-subject-out cross-validation was performed and the classification accuracy was calculated as an average percentage of correct guesses across all the cross-validation runs (Fig. 1d).

For full network classification, the classifier was trained and tested with the full connectivity matrix of each trial. For subnetwork classification, the classifier was trained and tested with the connectivity matrix of each sample either *within* one subnetwork (e.g. connectivity of the nodes within default node network) or *between* two subnetworks (e.g. connectivity between the nodes of default mode network and visual networks, omitting connections of nodes within each network). A separate classifier was trained for each within/between subnetwork division. Based on the subnetwork classifier results (see below), we also investigated the default mode system's subnetworks in more detail. Therefore, we further split the default mode system into four subnetworks (right temporal, left temporal, midline frontal, and midline posterior) based on clustering of the spatial distances between pairs of nodes within the default mode system and trained a separate classifier for each within/between default subnetworks.

For all classification approaches, we used a permutation test to assess the significance of the results (see, e.g., (Combrisson and Jerbi, 2015)). To obtain a null distribution, we generated 5000 surrogate accuracy values for the full network and for each subnetwork separately by shuffling the rows of the upper triangle of the connectivity matrix. The null cumulative distribution function was obtained using kernel smoothing, and the average classification accuracies were compared to the permuted null distribution to obtain their p values. Multiple comparisons were corrected for by using the Benjamini-Hochberg FDR correction.

To visualize the differences between connectivity patterns of different emotions, we ran permutation-based t-tests (Glerean et al., 2016) to compare the connectivity matrix of each emotion to that of the rest of the emotions (FDR-corrected for multiple comparisons at $p < 0.05$). This resulted in a contrast connectivity matrix showing the connections that were unique to each emotion.

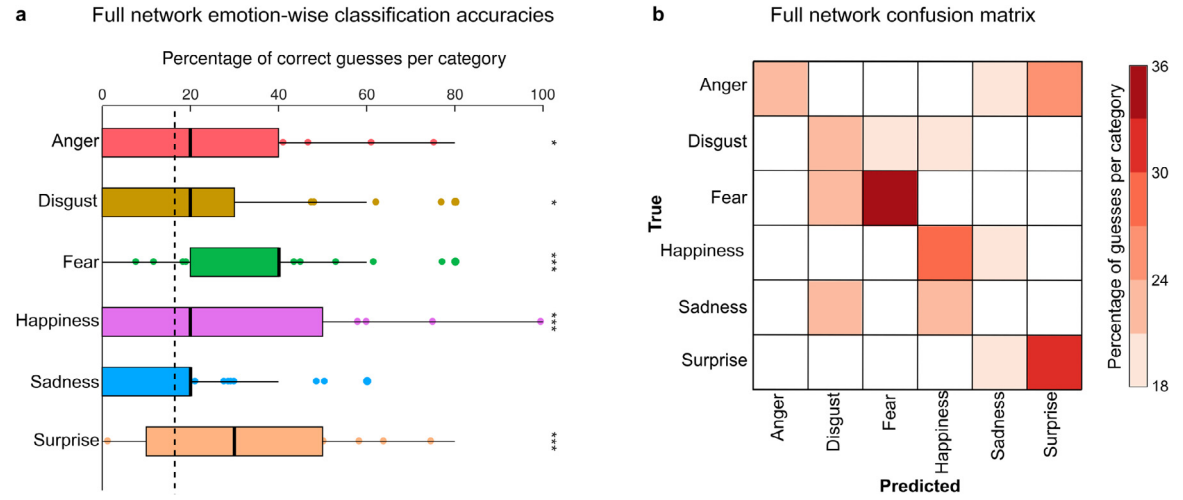

Fig. 2. **(a)** Emotion-wise classification accuracies for the full-network classification. Dashed line represents naïve chance level (16.6%). Asterisks denote significance relative to chance level ($^{*}p < 0.01$, $^{***}p < 0.0001$). Thick black line represents median of classification accuracies. Boxes show the 25th to 75th percentiles of classification accuracies and values outside this range are plotted as circles. Whiskers extend from box to the largest value no further than 1.5 * inter-quartile range from the edge of the box. **(b)** Classifier confusions from full network classification. Color code denotes average classifier accuracy over the cross-validation runs, cells shown in white have guesses below naïve chance level.

3. Results

3.1. Classifying emotions from full-brain connectivity patterns

To test whether different emotions are characterized by distinct connectivity patterns, we trained a between-subject classifier with the full network data to recognize the corresponding emotion category out of the six possible ones. Mean classification accuracy was 26% (naïve chance level of 16.6%; permuted $p < 0.00001$). The mean classification performance was above the permutation-based significance level for all emotions except sadness (Fig. 2A; see Fig. 2B for a confusion matrix): anger 23% (FDR corrected $p = 0.003$), disgust 23% ($p = 0.003$), fear 35% ($p < 0.00001$), happiness 28% ($p < 0.00001$), sadness 18% ($p = 0.291$), and surprise 31% ($p < 0.00001$). Finally, we compared the connectivity-based classification to voxel-based classification (see Supplementary Methods and Results). The connectivity-based classification accuracies were comparable to the voxel-based classification accuracies (26% and 25%, respectively; see Supplementary Fig. S3).

3.2. Classifying emotions from within and between subnetwork connectivity patterns

Our previous studies employing MVPA on brain activity patterns investigated both whole-brain and different regions-of-interest (Saarimäki et al., 2016; 2018). Here we followed a comparable pipeline by implementing the region-of-interest analysis as a *subnetwork-of-interest* analysis. To this end, we separated the connectivity matrices for each functional subnetwork (Power et al., 2014), and trained and tested the across-subject classifiers separately for the connections *within* each subnetwork as well as *between* all possible pairs of subnetworks. Mean classification accuracies and confusion matrices for each within and between subnetwork classifier are shown in Fig. 3. Classification accuracy was highest for connections within the default mode system (30%) which after correcting for multiple comparisons remained the only subnetwork showing significant classification accuracy ($p<0.0001$; see Supplementary Table S3 for all subnetwork accuracies and p values).

To visualize the emotion-specific functional connectivity, we plotted the connectivity matrices (Supplementary Fig. S5; see Supplementary Fig. S4 for connectivity matrix of the neutral state). Correlations between each pair of connectivity matrices for all emotions and the neutral state were significant (Supplementary Fig. S5; correlations ranging from rho = 0.79–0.84, all Bonferroni-corrected Ps < 0.01; see Supplementary Table S2). The average connectivity patterns for each emotion with neutral baseline removed are shown in Supplementary Fig. S6 and pairwise t-tests between the connectivity matrices in Supplementary Fig. S8.

3.3. Classifying emotions from default mode network connectivity

Because default mode system was the only subnetwork where connectivity-based classification accuracies were above chance-level after post-hoc correction, we next investigated its connectivity patterns in more detail. Within the default mode system, all emotions could be classified with above-chance-level accuracy (see Fig. 4a): anger 24% ($p = 0.001$), disgust 33% ($p < 0.00001$), fear 30% ($p < 0.00001$), happiness 31% ($p < 0.00001$), sadness 26% ($p = 0.0002$), and surprise 35% ($p < 0.00001$). See Supplementary Fig. S7 for emotion-specific connectivity patterns for default mode network.

We next investigated whether there are node-wise differences in classification accuracy within the default mode network by sub-grouping the DMN nodes from Power et al. (2011) to four subnetworks based on their spatial proximity in the brain. The spatial clustering of DMN resulted in four DMN subnetworks: left temporal, right temporal, frontal, and posterior midline. We trained classifiers to recognize emotions from the connections within and between these DMN subnetworks separately. Classification was successful for connections within midline posterior subnetwork (22.5%, permuted and FDR-corrected $p = 0.008$), between left temporal and midline frontal subnetworks (23.5%, $p < 0.001$), and between right temporal and midline posterior subnetworks (21.5%, $p = 0.032$). The classification accuracies and confusion matrices are shown in Fig. 4.

4. Discussion

The goals of the current study were to (i) provide a *proof-of-concept* for the application of machine learning to functional connectivity patterns, and to (ii) test whether instances of different emotional states can be classified from functional connectivity. We show differences in sustained whole-brain functional connectivity patterns during emotional narratives (targeting anger, fear, disgust, happiness, and surprise), as evidenced by significantly above chance-level classification accuracy. The classifier was trained across participants, demonstrating that the connectivity patterns were consistent across subjects. Especially, default

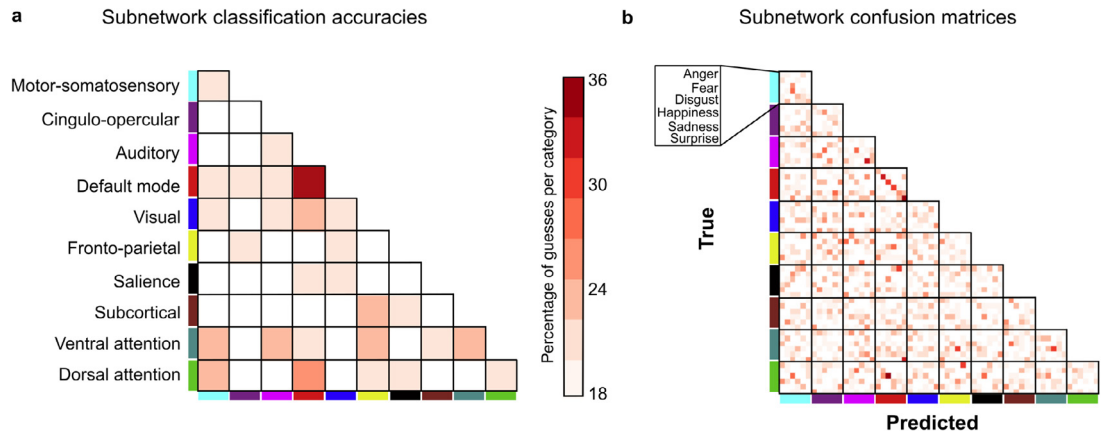

Fig. 3. **(a)** Classification accuracies for connectivity within and between each ROI. Color code denotes classifier accuracy; cells shown in white have guesses below naïve chance level (16.6%). After correcting for multiple comparisons, only the accuracy for within default mode network connections remained significant. **(b)** Classifier confusions for subnetwork classification.

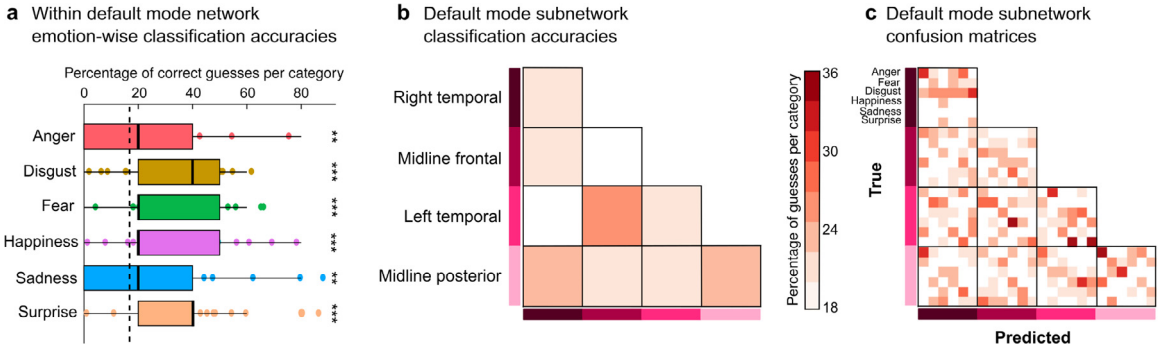

Fig. 4. **(a)** Emotion-wise classification accuracies for connections within the default mode system. Dashed line represents naïve chance level (16.6%). Asterisks denote significance relative to chance level ($^{**}p < 0.001$, $^{***}p < 0.0001$). Thick line represents median of classification accuracies. Boxes show the 25th to 75th percentiles of classification accuracies and values outside this range are plotted as dots. Whiskers extend from box to the largest value no further than 1.5 * inter-quartile range from the edge of the box. **b)** Classification accuracies and **c)** subnetwork confusion matrices for DMN subnetwork classification. Color code denotes classifier accuracy; cells shown in white have guesses below naïve chance level.

mode system contained emotion-specific connectivity patterns sustained during the whole one-minute-long narrative. Thus, our results show preliminary evidence for distinct connectivity patterns underlying instances of emotion categories and suggest that the methodological approach is plausible for future studies with larger samples.

4.1. Emotion-specific changes in default mode network connectivity

Connectivity-based classification accuracy was highest when we considered only the connections within the default mode system. The role of DMN regions in integrating information about one's internal state and accessing representations for self-relevant mentalization is well-known (Klasen et al., 2011; Mar 2011; Kleckner et al., 2017; Summerfield et al., 2009; D'Argembeau et al. 2010; Andrews-Hanna et al. 2014; Barrett, 2012). Furthermore, DMN areas have been linked to manipulating abstract representations and conscious thought, and emotional experiences are one type of consciousness with specific somatomotor and interoceptive inputs that distinguishes them from other conscious experiences (LeDoux and Brown, 2017; Nummenmaa et al., 2018). It has been suggested that DMN is anatomically well-positioned to process transmodal information that is unrelated to immediate sensory and motor processing (Marguelis et al., 2016). In line with this, our results show that sustained connectivity patterns within the DMN differ across instances of emotion categories. In the present study, the functional connectivity patterns were calculated over a period of one minute; thus, rapid changes in stimulus or connectivity were not picked up by the analysis. Previous studies have also shown that emotions lead to sustained changes in DMN connectivity that persist even minutes after the emotion-evoking episode ends (Eryilmaz et al., 2011). Sad mood and depression also modulate DMN connectivity (Harrison et al., 2008; Whitfield-Gabrieli and Ford, 2012), further supporting the role of DMN holding sustained emotion-evoked representations.

Default mode system regions are also important in distinguishing between experiences of discrete emotions. These areas contain the most robust distinct voxel activity patterns that allow machine learning based classification of different emotions (Saarimäki et al., 2016; Wager et al., 2015; Kragel and LaBar 2015; Horikawa et al., 2020). While these studies show robust regional differences between emotion categories in DMN, its exact role in emotional processing remains unclear. It has been suggested that instances of emotion categories arise from the dynamic interaction of wide-scale brain networks (see, e.g., Barrett and Satpute, 2013; Kleckner et al., 2017) and, indeed, the aforementioned role of DMN in sustained transmodal associations makes it a good candidate for supporting this interaction (Barrett, 2012). Accordingly, it has recently been suggested that DMN plays a constitutive role in holding representational content to create a discrete experience of emotion (Satpute and Lindquist, 2019). The DMN activity patterns resulting from the integration of different representations, such as those of salience, potential motor actions, bodily sensations, or sensory features, may constitute a core feature of the emotional experience, demonstrated also as discrete neural patterns for discrete emotions (Saarimäki et al., 2016).

Default mode network can be divided into subsystems that serve different cognitive functions (Buckner et al., 2008; Andrews-Hanna et al., 2010, 2014), and the subregions of DMN vary in their involvement in the representation of different emotion categories (Horikawa et al., 2020). Therefore, we examined whether emotions can be classified from connectivity patterns within and between different DMN subsystems. In our spatial parcellation, the default mode subnetworks comprised medial PFC, posterior midline regions, and left and right lateral temporal areas (Power et al., 2011). Emotions could be classified most accurately from connectivity patterns within the posterior midline DMN connections including posterior cingulate cortex and precuneus. These regions have been linked to integration of information from other DMN subsystems and other brain regions (Buckner et al., 2008; Andrews-Hanna et al. 2014). Successful classification of emotions from the connectivity patterns in this region suggest that emotions might vary in how and which information is integrated during the emotional state. For instance, interpretation of emotions with a strong action component (e.g., fear) might rely more on motor inputs, while the role of such inputs is smaller with emotional states that do not require immediate motor actions (e.g., sadness). This accords with the view that emotions arise from integrated activity across multiple physiological, behavioral and neural systems (Sander et al., 2018; Nummenmaa and Saarimäki, 2019; Satpute and Lindquist, 2019).

4.2. Functional connectivity is modulated differently by different emotions

We showed that different whole-brain connectivity patterns underlie instances of specific emotional categories similarly as has previously been found for regional activity patterns (for reviews, see Kragel and LaBar 2016; Nummenmaa and Saarimäki, 2019). We have previously shown that instances of anger, fear, disgust, happiness, sadness, and surprise can be classified from brain activity in the cortical midline regions, subcortical regions, somatomotor regions, but also regions related to cognitive functions such as memory and language (Saarimäki et al., 2016), thus suggesting that different emotions modulate each of these regions differently. These areas are also consistently activated in studies using univariate analysis of emotional brain responses (Kober et al., 2008). Together, the results from the voxel-based and connectivity-based pattern classification of emotions support a model where emotions are brought about by distributed net activation of different regions together with their connectivity patterns.

We found no significant classification performance for the connectivity within or between areas other than default mode network despite the evidence showing that activation patterns within a large number of other areas, including subcortical, somatomotor, and frontal areas, differ between emotions (Saarimäki et al., 2016). This is probably partly due to how connectivity was calculated in the current study: connectivity was calculated over a time period of one minute, which is less sensitive to rapid temporal changes in connectivity that potentially underlie different emotional states (see, e.g., Pessoa, 2018). For instance, Nummenmaa et al. (2014b) have shown that such dynamic changes reveal large connectivity differences in positive versus negative valence and high versus low arousal. Furthermore, it is unlikely that emotional state remains the same over the 1-minute-long stimulation, probably restricting the stability of connectivity patterns and affecting their distinctness. However, connectivity measures with shorter time windows usually contain more noise. Thus, classification of connectivity patterns in general is a more difficult task than that of activation patterns; yet, the current proof-of-concept work shows that it is possible. Finally, in future it would be interesting to compare the stability of regional activity and network connectivity, and their role in decoding emotion states from the brain. However, this would necessitate a dedicated study designed for this purpose.

Labeling the emotion category is a descriptive summary of the current emotional which consists of several different aspects (see, e.g., Sander et al., 2018). When classifying instances of emotion categories, we might thus miss other important, emotion-related distinctions between the stimuli, such as appraisal or semantic properties. Indeed, the differences between categories probably stem from differences in these kinds of functional components (Sander et al., 2018; Satpute and Lindquist, 2019). For instance, Horikawa et al. (2020) showed that DMN encoded emotion above and beyond appraisal and semantic categories, suggesting that these components alone are not enough to explain variation in DMN activity between emotion categories. However, it is possible that appraisal and semantic features do not grasp the whole variation in emotion categories. For instance, emotion-related bodily changes are ignored in these models. As we have suggested earlier (Nummenmaa and Saarimäki, 2019), brain activity patterns underlying different emotion categories probably reflect activity in different components and the net activity across all these components is integrated and labeled as a specific emotion category. This is further supported by our current finding that DMN connectivity – probably highlighting differences in integration between emotions – varies between emotions.

All emotions except for sadness could be classified significantly above chance level from each other. Confusion matrix shows that sadness was most often confused with disgust and happiness in whole-network classification and with surprise in within DMN classification, suggesting no consistent confusions with any specific emotion category. Furthermore, sadness was still successfully classified for some participants while the classification average was at chance. The reasons for this can only be speculated. When looking at the content of the narrative stimuli, sadness stories describe death and separation from loved ones, sharing the social closeness aspect with most happiness stories. When comparing sadness and disgust stories, relative to other categories, these stories describe visual details of the appearance of other persons in more detail than other stories. While similarities regarding the content might contribute to the experience of emotions and lead to similarities in sustained connectivity patterns due to the similarities in underlying input systems, a more detailed analysis regarding the input differences between emotion categories is an interesting empirical question for future experiments.

4.3. Across-participant classification and individual differences

The overall, whole-brain accuracy (26% against chance level of 16.6%) in the current study corresponds to the 23–34% (against chance levels of 16.7% and 20%, correspondingly) across-participant classification accuracies reached by regions classifiers trained to classify an equal number of emotional states (Saarimäki et al., 2016). The within-participant classification accuracies of 50–60% by regional classifiers are consistently higher than the across-participant classification.

As highlighted by the Figs. 2a and 4b, classification accuracy varied across cross-validation runs. With our leave-one-participant-out classifier, this means that there is variation in how well the connectivity patterns of different individuals can be classified. Given the number of participants (N = 16), the across-participant classification is affected by individual variation and, thus, individual variation in emotional experiences has a potentially larger effect on the results than when using a larger sample size. Therefore, the results are tentative and exploratory and should be investigated further in larger samples.

Potential sources for individual variation include the idiosyncratic emotional experiences evoked by the stories. We focused on six emotion categories that were defined *a priori* by the researchers and the storyteller that provided the narratives. Thus, the ground truth categories reflect the emotion concepts used by the storyteller that she deliberately tried to convey. This model naturally assumes that consistent emotions are also evoked in listeners, and earlier analysis of the same data has shown that the emotional experiences and brain activations match well across the speakers and listeners (Smirnov et al., 2019). To confirm this alignment between speaker's and listener's emotions, we collected ratings of experienced emotion categories from an independent sample of individuals drawn from the same population. These ratings show that,

on average, instances of emotion categories in the speaker elicit experiences of the same emotion categories in listeners (Supplementary Fig. S1). Also, between-subject variability in emotion category ratings was low, demonstrating that individuals from the same population experienced similar emotions (Supplementary Table S1). Yet, while we further guided the interpretation of the story by presenting the target emotion category for the listeners before each story, different individuals might interpret the emotional content of the story differently, leading to differences in emotion-related brain activity (see Putkinen et al., 2021, for similar arguments on classification of musical emotions). This obviously explains why the classification accuracy in across-participant classification is in general lower than in within-participant classification (see also Saarimäki et al., 2016). However, these six emotion categories – anger, fear, disgust, happiness, sadness, and surprise – capture only ~30% of the systematic variance in reported emotional experience (Cowen et al., 2019). Also, Horikawa et al. (2020) found that emotion-related variations in DMN activity go beyond these six emotion categories. Therefore, the stories we used to elicit each emotion probably did not evoke similar emotional experiences across all participants, decreasing the classification accuracy.

Taken together, the across-participant classification results in this and previous studies show the importance of investigating individual differences in emotional processing. Addressing the individual differences in this classification performance is out of the scope of the current paper but constitutes an important topic for future research.

4.4. Limitations

Our study uses a naturalistic paradigm to elicit emotions. Naturalistic stimuli such as stories and movies have been shown to elicit strong emotions (e.g., Westermann et al., 1996); however, compared to controlled, static stimuli, the continuous nature of naturalistic stimuli also makes controlling the other stimulus features more difficult (e.g., (Huk et al., 2018); Saarimäki, 2021). While our stimuli vary also in speech, prosody, and content, the careful piloting of the stimuli ensured that the common nominator between instances of the same category is emotion and not some other feature (see also Supplementary Table S4 for content of the stories). Furthermore, naturalistic stimuli require simultaneous processing of speaker's emotions as well as one's own emotions; thus, they evoke both perceived and experienced emotions. Thus, we cannot rule out that the emotion classification in the current study would not also contain some elements of perceived emotions. Separating between the perceived and experienced emotions is not s straight-forward task, as perception of emotions likely elicits similar emotional experiences in the listener (e.g.., Nummenmaa et al., 2014). Future studies should parse the neural basis of perceived and experienced emotions more carefully, for instance, by creating separate model timeseries for perceived and experienced emotions.

The current experiment had a 25-second wash-out period between emotional stimuli. It is likely that emotional effects from the previous trial are still present in the following trial, as it has been shown that emotional induction lasts for minutes after the stimulation ends (Eryilmaz et al., 2011)). However, the experimental stimulation of the next trial begins immediately after the wash-out period, allowing new emotional stimulus to modulate the emotion systems and likely overriding the previous input. Moreover, the problem was alleviated by varying the order of emotions between trials.

In our analysis of the haemodynamic data, we removed heart rate and respiration rate variation as noise, as removing heart rate and respiration variation improves strength of connectivity estimates but does not influence the network structure (Yoshikawa et al., 2020). Also, our previous research has shown only modest correlations between experienced emotion ratings and heart rate and respiration rate (Nummenmaa et al., 2014b), and a recent meta-analysis has shown that autonomic nervous system variation alone does not distinguish between emotion categories (Siegel et al., 2018). However, we cannot rule out the possibility that by removing heart rate and respiration rate variation we remove some physiological variation between emotion categories, and potentially also variation regarding their neural basis in brain regions processing interoceptive information (Nguyen et al., 2016). This might remove part of the somatic differences between emotion categories. However, it seems that ANS patterns alone do not reflect the full scope of bodily changes related to emotions (Nummenmaa et al., 2014a). Thus, even though we remove ANS variation, other somatic differences should still be present (see also Eisenbarth et al., 2016).

There are also other factors might affect the classifier performance especially when applying a classification paradigm to functional connectivity. First, classifier accuracy might be limited by the low general power on fMRI connectivity analysis due to slow-frequency signal of interest. Future studies should address this issue using, for example, MEG experiments yielding better temporal resolution for dynamic connectivity measures. Second, the classification accuracy might depend on the number of data points available for the classification. To raise the number of data points, we trained the classifier between subjects. However, this probably leads to variation between the instances of emotions due to individual differences in emotion processing and brain anatomy. A more individualized analysis scheme using, for instance, inter-subject representational similarity analyzes in a larger sample might allow modeling individual differences (see, e.g., Finn et al., 2020). Third, while multivariate pattern analysis results are not straightforward to interpret in a neurophysiological sense, they still suggest that instances of the same emotion category share similar enough properties to allow for successful classification.

5. Conclusion

We conclude that instances of emotion categories including anger, fear, disgust, happiness, sadness, and surprise differ in the connectivity patterns across the brain in a manner that is consistent across individuals. Connectivity-based classification of emotions was most accurate within the default mode system, suggesting that connectivity across this region contains the most accurate, sustained representation of the individual's current emotional state. Together with previous regional voxel-based pattern classification results, the present findings support the view that emotions are represented in a distributed fashion across the brain and speak for the importance of DMN in sustaining the emotional state.

Credit authorship contribution statement

Heini Saarimäki: Methodology, Formal analysis, Investigation, Data curation, Writing – original draft. **Enrico Glerean:** Methodology, Software, Formal analysis, Writing – original draft. **Dmitry Smirnov:** Investigation, Software. **Henri Mynttinen:** Methodology, Software. **Iiro P. Jääskeläinen:** Conceptualization, Funding acquisition, Supervision, Writing – original draft. **Mikko Sams:** Conceptualization, Supervision, Writing – original draft. **Lauri Nummenmaa:** Conceptualization, Funding acquisition, Supervision, Writing – original draft.

Acknowledgments

This work was supported by Academy of Finland (#265917 to L.N. and #138145 to I.P.J.), ERC Starting Grant (#313000 to L.N.); Finnish Cultural Foundation (#00140220 to H.S.), Kordelin Foundation (#160387 to H.S.), and by the International Laboratory of Social Neurobiology ICN HSE RF Government grant ag. No. 075–15- 2019–1930 (to I.P.J and E.G.). We thank Marita Kattelus for herhelp with the data acquisition. We acknowledge the computational resources provided by the Aalto Science-IT project.

Data sharing statement

The ethical pre-evaluation of the current study does not allow sharing of raw data. Connectivity matrices and stimulus ratings are

available from the corresponding author upon request. Code for running the analyzes can be found at https://github.com/hpsaarimaki/Aalto-emotion-networks.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118800.

References

Adolphs, R, 2017. How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences. Social Cognitive and Affective Neuroscience 12, 24–31.

Andrews-Hanna, J.R., Reidler, J.S., Sepulcre, J., Poulin, R., Buckner, R.L., 2010. Functional-anatomic fractionation of the brain's default network. Neuron 65, 550–562.

Andrews-Hanna, J.R., Smallwood, J., Spreng, R.N., 2014. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann. N. Y. Acad. Sci. 1316, 29–52.

Barrett, L.F., 2012. Emotions are real. Emotion 12, 413–429.

Barrett, L.F., 2017. The theory of constructed emotion: an active inference account of interoception and categorization. Social Cognitive and Affective Neuroscience 12, 1–23.

Barrett, L.F., Satpute, A.B., 2013. Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. Curr. Opin. Neurobiol. 23, 361–372.

Bassett, D.S., Sporns, O., 2017. Network neuroscience. Nat. Neurosci. 20, 353–364.

Borchardt, V., Krause, A.L., Li, M., van Tol, M.J., Demenescu, L.R., Buchheim, A., Metzger, C.D., Sweeney-Reed, C.M., Nolte, T., Lord, A.R., Walter, M., 2015. Dynamic disconnection of the supplementary motor area after processing of dismissive biographic narratives. Brain Behav. 5, e00377.

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network. Ann. N. Y. Acad. Sci. 1124, 1–38.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.

Bullmore, E., Sporns, O., 2012. The economy of brain network organization. Nat. Rev. Neurosci. 13, 336–349.

Chikazoe, J., Lee, D.H., Kriegeskorte, N., Anderson, A.K., 2014. Population coding of affect across stimuli, modalities and individuals. Nat. Neurosci. 17, 1114–1122.

Combrisson, E., Jerbi, K., 2015. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. Journal of Neuroscience Methods 250, 126–136.

Cowen, A.S., Keltner, D., 2017. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. Proc. Natl. Acad. Sci. 114, E7900–E7909.

Cowen, A., Sauter, D., Tracy, J.L., Keltner, D., 2019. Mapping the passions: toward a high-dimensional taxonomy of emotional experience and expression. Psychol. Sci. Public Interest 20, 69–90.

Cukur, T., Nishimoto, S., Huth, A.G., Gallant, J.L., 2013. Attention during natural vision warps semantic representation across the human brain. Nature Neuroscience 16, 763–770.

D'Argembeau, A., Stawarczyk, D., Majerus, S., Collette, F., Van der Linden, M., Feyers, D., Maquet, P., Salmon, E., 2010. The neural basis of personal goal processing when envisioning future events. J. Cogn. Neurosci. 22, 1701–1713.

Eisenbarth, H., Chang, L.J., Wager, T.D., 2016. Multivariate brain prediction of heart rate and skin conductance responses to social threat. J. Neurosci. 36, 11987–11998.

Eryilmaz, H., Van De Ville, D., Schwartz, S., Vuilleumier, P., 2011. Impact of transient emotions on functional connectivity during subsequent resting state: a wavelet correlation approach. Neuroimage 54, 2481–2491.

Feinstein, J.S., 2013. Lesion studies of human emotion and feeling. Curr. Opin. Neurobiol. 23, 304–309.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., … Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671.

Finn, E.S., Glerean, E., Khojandi, A.Y., Nielson, D., Molfese, P.J., Handwerker, D.A., Bandettini, P.A., 2020. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. Neuroimage 215, 116828.

Glerean, E., Pan, R.K., Salmi, J., Kujala, R., Lahnakoski, J.M., Roine, U., Nummenmaa, L., Leppämäki, S., Nieminen-von Wendt, T., Tani, P., Saramäki, J., Sams, M., Jääskeläinen, I.P., 2016. Reorganization of functionally connected brain subnetworks in high-functioning autism. Hum. Brain Mapp. 37, 1066–1079.

Gonzalez-Castillo, J., Hoy, C.W., Handwerker, D.A., Robinson, M.E., Buchanan, L.C., Saad, Z.S., Bandettini, P.A., 2015. Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns. Proc. Natl. Acad. Sci. 112, 8762–8767.

Hamann, S., 2012. Mapping discrete and dimensional emotions onto the brain: controversies and consensus. Trends Cogn. Sci. 16, 458–466.

Harrison, B.J., Pujol, J., Ortiz, H., Fornito, A., Pantelis, C., Yücel, M., 2008. Modulation of brain resting-state networks by sad mood induction. PLoS One 3, e1794.

Hasler, G., Northoff, G., 2011. Discovering imaging endophenotypes for major depression. Mol. Psychiatry 16, 604–619.

Hermans, E.J., van Marle, H.J., Ossewaarde, L., Henckens, M.J., Qin, S., van Kesteren, M.T., Schoots, V.C., Cousijn, H., Rijpkema, M., Oostenveld, R., Fernández, G., 2011. Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. Science 334, 1151–1153.

Hermans, E.J., Henckens, M.J., Joëls, M., Fernánzez, G., 2014. Dynamic adaptation of large-scale brain networks in response to acute stressors. Trends Neurosci. 37, 304–314.

Hietanen, J.K., Glerean, E., Hari, R., Nummenmaa, L., 2016. Bodily maps of emotions across child development. Dev. Sci. 19, 1111–1118.

Horikawa, T., Cowen, A.S., Keltner, D., Kamitani, Y., 2020. The neural representation of visually evoked emotion is high-dimensional, categorical, and distributed across transmodal brain regions. iScience 23, 101060.

Huang, Y.A., Jastorff, J., Van den Stock, J., Van de Vliet, L., Dupont, P., Vandenbulcke, M., 2018. Studying emotion theories through connectivity analysis: evidence from generalized psychophysiological interactions and graph theory. Neuroimage 172, 250–262.

Huk, A., Bonnen, K., He, B.J., 2018. Beyond trial-based paradigms: Continuous behavior, ongoing neural activity, and natural stimuli. Journal of Neuroscience 38, 7551–7558.

Hutcherson, C.A., Goldin, P.R., Ochsner, K.N., Gabrieli, J.D., Barrett, L.F., Gross, J.J., 2005. Attention and emotion: does rating emotion alter neural responses to amusing and sad films? Neuroimage 27, 656–668.

Kaiser, R.H., Andrews-Hanna, J.R., Wager, T.D., Pizzagalli, D.A., 2015. Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. JAMA Psychiatry 72, 603–611.

Klasen, M., Kenworthy, C.A., Mathiak, K.A., Kircher, T.T., Mathiak, K., 2011. Supramodal representation of emotions. J. Neurosci. 31, 13635–13643.

Kleckner, I.R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W.K., Barrett, L.F., 2017. Evidence for a large-scale brain system supporting allostasis and interoception in humans. Nat. Hum. Behav. 1, 0069.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D., 2008. Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies. Neuroimage 42, 998–1031.

Kragel, P.A., LaBar, K.S., 2015. Multivariate neural biomarkers of emotional states are categorically distinct. Soc. Cogn. Affect. Neurosci. 10, 1437–1448.

Kragel, P.A., LaBar, K.S., 2016. Decoding the nature of emotion in the brain. Trends Cogn. Sci. 20, 444–455.

Kreibig, S.D., Wilhelm, F.H., Roth, W.T., Gross, J.J., 2007. Cardiovascular, electrodermal, and respiratory response patterns to fear-and sadness-inducing films. Psychophysiology 44, 787–806.

LeDoux, J.E., Brown, R., 2017. A higher-order theory of emotional consciousness. Proc. Natl Acad. Sci. 114, E2016–E2025..

Lieberman, M.D., Eisenberger, N.I., Crockett, M.J., Tom, S.M., Pfeifer, J.H., Way, B.M., 2007. Putting feelings into words. Psychol. Sci. 18, 421–428.

Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F., 2012. The brain basis of emotion: a meta-analytic review. Behav. Brain Sci. 35, 121–143.

Mar, R.A., 2011. The neural bases of social cognition and story comprehension. Annu. Rev. Psychol. 62, 103–134.

Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., Bezgin, G., Eickhoff, S.B., Castellanos, F.X., Petrides, M., Jefferies, E., 2016. Situating the default-mode network along a principal gradient of macroscale cortical organization. Proc. Natl Acad. Sci. 113, 12574–12579.

McMenamin, B.W., Langeslag, S.J., Sirbu, M., Padmala, S., Pessoa, L., 2014. Network organization unfolds over time during periods of anxious anticipation. J. Neurosci. 34, 11261–11273.

Nguyen, V.T., Breakspear, M., Hu, X., Guo, C.C., 2016. The integration of the internal and external milieu in the insula during dynamic emotional experiences. Neuroimage 124, 455–463.

Nummenmaa, L., Saarimäki, H., 2019. Emotions as discrete patterns of systemic activity. Neurosci. Lett. 693, 3–8.

Nummenmaa, L., Glerean, E., Hari, R., Hietanen, J.K., 2014a. Bodily maps of emotions. Proc. Natl. Acad. Sci. U. S. A. 111, 646–651.

Nummenmaa, L., Saarimäki, H., Glerean, E., Gotsopoulos, A., Hari, R., Sams, M., 2014b. Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. Neuroimage 102, 498–509.

Nummenmaa, L., Hari, R., Hietanen, J.K., Glerean, E., 2018. Maps of subjective feelings. Proc. Natl. Acad. Sci. U. S. A. 115, 9198–9203.

Pedregosa, F., Varoquaux, G., Gramfort, A., Vincent, M., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Peelen, M.V., Atkinson, A.P., Vuilleumier, P., 2010. Supramodal representations of perceived emotions in the human brain. J. Neurosci. 30, 10127–10134.

Pessoa, L., 2012. Beyond brain regions: network perspective of cognition–emotion interactions. Behav. Brain Sci. 35, 158–159.

Pessoa, L., 2017. A network model of the emotional brain. Trends Cogn. Sci. 21, 357–371.

Pessoa, L., 2018. Understanding emotion with brain networks. Curr. Opin. Behav. Sci. 19, 19–25.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional network organization of the human brain. Neuron 72, 665–678.

Power, J.D., Schlaggar, B.L., Petersen, S.E., 2014. Studying brain organization via spontaneous fMRI signal. Neuron 84, 681–696.

Price, R.B., Lane, S., Gates, K., Kraynak, T.E., Horner, M.S., Thase, M.D., Siegle, G.J., 2017. Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. Biol. Psychiatry 81, 347–357.

Putkinen, V., Nazari-Farsani, S., Seppälä, K., Karjalainen, T., Sun, L., Karlsson, H.K., Hudson, M., Heikkilä, T.T., Hirvonen, J., Nummenmaa, L., 2021. Decoding music-evoked emotions in the auditory and motor cortex. Cereb. Cortex doi:10.1093/cercor/bhaa373.

Raz, G., Touroutoglou, A., Wilson-Mendenhall, C., Gilam, G., Lin, T., Gonen, T., Jacob, Y., Atzil, S., Admon, R., Bleich-Cohen, M., Maron-Katz, A., Hendler, T., Barrett, L.F., 2016.

Functional connectivity dynamics during film viewing reveal common networks for different emotional experiences. Cogn. Affect. Behav. Neurosci. 16, 709–723.
Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P., Van De Ville, D., 2011. Decoding brain states from fMRI connectivity graphs. Neuroimage 56, 616–626.
Saarimäki, H., 2021. Naturalistic stimuli in affective neuroimaging: a review. Front. Hum. Neurosci. 15, 318.
Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I.P., Lampinen, J., Vuilleumier, P., Hari, R., Lampinen, J., Nummenmaa, L., 2016. Discrete neural signatures of basic emotions. Cereb. Cortex 26, 2563–2573.
Saarimäki, H., Ejtehadian, L.F., Glerean, E., Jääskeläinen, I.P., Vuilleumier, P., Sams, M., Nummenmaa, L., 2018. Distributed affective space represents multiple emotion categories across the human brain. Soc. Cogn. Affect. Neurosci. 13 (5), 471–482.
Sachs, M.E., Habibi, A., Damasio, A., Kaplan, J.T., 2020. Dynamic intersubject neural synchronization reflects affective responses to sad music. Neuroimage 218, 116512. doi:10.1016/j.neuroimage.2019.116512.
Sander, D., Grandjean, D., Scherer, K.R., 2018. An appraisal-driven componential approach to the emotional brain. Emot. Rev. 10, 219–231.
Särkkä, S., Solin, A., Nummenmaa, A., Vehtari, A., Auranen, T., Vanni, S., Lin, F.H., 2012. Dynamic retrospective filtering of physiological noise in BOLD fMRI: DRIFTER. Neuroimage 60, 1517–1527.
Satpute, A.B., Lindquist, K.A., 2019. The default mode network's role in discrete emotion. Trends Cogn. Sci. 23, 851–864.
Scarantino, A., 2012. Functional specialization does not require a one-to-one mapping between brain regions and emotions. Behav. Brain Sci. 35, 161–162.
Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J. Neurosci. 27, 2349–2356.
Shen, X., Finn, E.S., Scheinost, D., Rosenberg, M.D., Chun, M.M., Papademetris, X., Constable, R.T., 2017. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat. Protoc. 12, 506–518.
Shirer, W.R., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D., 2012. Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb. Cortex 22, 158–165.
Siegel, E.H., Sands, M.K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K.S., Barrett, L.F., 2018. Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. Psychol. Bull. 144, 343.
Siegle, G.J., Thompson, W., Carter, C.S., Steinhauer, S.R., Thase, M.E., 2007. Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. Biol. Psychiatry 61, 198–209.
Smirnov, D., Saarimäki, H., Glerean, E., Hari, R., Sams, M., Nummenmaa, L., 2019. Emotions amplify speaker–listener neural alignment. Hum. Brain Mapp. 40 (4777), 4788.
Summerfield, J.J., Hassabis, D., Maguire, E.A., 2009. Cortical midline involvement in autobiographical memory. Neuroimage 44, 1188–1200.
Tettamanti, M., Rognoni, E., Cafiero, R., Costa, T., Galati, D., Perani, D., 2012. Distinct pathways of neural coupling for different basic emotions. Neuroimage 59, 1804–1817.
Touroutoglou, A., Lindquist, K.A., Dickerson, B.C., Barrett, L.F., 2015. Intrinsic connectivity in the human brain does not reveal networks for "basic" emotions. Soc. Cogn. Affect. Neurosci. 10, 1257–1265.
Wager, T.D., Kang, J., Johnson, T.D., Nichols, T.E., Satpute, A.B., Barrett, L.F., 2015. A Bayesian model of category-specific emotional brain responses. PLoS Comput. Biol. 11, e1004066.
Westermann, R., Spies, K., Stahl, G., Hesse, F.W., 1996. Relative effectiveness and validity of mood induction procedures: a meta-analysis. Eur. J. Soc. Psychol. 26, 557–580. doi:10.1002/(sici)1099-0992(199607)26:4<557::aid-ejsp769>3.0.co;2-4.
Whitfield-Gabrieli, S., Ford, J.M., 2012. Default mode network activity and connectivity in psychopathology. Annu. Rev. Clin. Psychol. 8, 49–76.
Yoshikawa, A., Masaoka, Y., Yoshida, M., Koiwa, N., Honma, M., Watanabe, K., Kubota, S., Natsuko, I., Ida, M., Izumizaki, M., 2020. Heart rate and respiration affect the functional connectivity of default mode network in resting-state functional magnetic resonance imaging. Front. Neurosci. 14, 631.
Young, C.B., Raz, G., Everaerd, D., Beckmann, C.F., Tendolkar, I., Hendler, T., et al., 2017. Dynamic shifts in large-scale brain network balance as a function of arousal. J. Neurosci. 37, 281–290. doi:10.1523/jneurosci.1759-16.2017.