Title

Unsupervised learning of brain state dynamics during emotion imagination using high-density EEG

Authors

Hsu, Sheng-Hsiou

Lin, Yayu

Onton, Julie

et al.

Publication Date

2022-04-01

DOI

10.1016/j.neuroimage.2022.118873

Peer reviewed

NeuroImage 249 (2022) 118873

Contents lists available at ScienceDirect

NeuroImage

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

Unsupervised learning of brain state dynamics during emotion imagination using high-density EEG

Sheng-Hsiou Hsu*, Yayu Lin, Julie Onton, Tzyy-Ping Jung, Scott Makeig

Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, La Jolla, CA, United States

ARTICLE INFO

Keywords:
Emotion
Affective computing
Electroencephalography (EEG)
Brain states
Non-stationarity
Independent component analysis (ICA)
Adaptive mixture ICA (AMICA)
Unsupervised learning
Source localization

ABSTRACT

This study applies adaptive mixture independent component analysis (AMICA) to learn a set of ICA models, each optimized by fitting a distributional model for each identified component process while maximizing component process independence within some subsets of time points of a multi-channel EEG dataset. Here, we applied 20-model AMICA decomposition to long-duration (1–2 h), high-density (128-channel) EEG data recorded while participants used guided imagination to imagine situations stimulating the experience of 15 specified emotions. These decompositions tended to return models identifying spatiotemporal EEG patterns or states within single emotion imagination periods. Model probability transitions reflected time-courses of EEG dynamics during emotion imagination, which varied across emotions. Transitions between models accounting for imagined "grief" and "happiness" were more abrupt and better aligned with participant reports, while transitions for imagined "contentment" extended into adjoining "relaxation" periods. The spatial distributions of brain-localizable independent component processes (ICs) were more similar within participants (across emotions) than emotions (across participants). Across participants, brain regions with differences in IC spatial distributions (i.e., dipole density) between emotion imagination versus relaxation were identified in or near the left rostrolateral prefrontal, posterior cingulate cortex, right insula, bilateral sensorimotor, premotor, and associative visual cortex. No difference in dipole density was found between positive versus negative emotions. AMICA models of changes in high-density EEG dynamics may allow data-driven insights into brain dynamics during emotional experience, possibly enabling the improved performance of EEG-based emotion decoding and advancing our understanding of emotion.

1. Introduction

Emotional experiences are functional states internally generated by the brain and body that give rise to diverse subjective experiences and behaviors in our daily lives (Scherer, 2005). Advancing our understanding of emotions, while devising technology for affective computing, could enable widespread applications in education, healthcare, gaming, and human-computer interaction (Picard, 2000). Despite enormous progress in cognitive science, psychology, and computer science studies of emotion, much emotion research relies on subjective verbal or written reports of participants (Cowen and Keltner, 2017). However, these subjective reports reflect the conscious awareness of their emotional experiences or the expressed emotion, which might differ from the underlying emotional states that could be unconscious or shaped by an individual's cognitive construct (Barrett et al., 2019; Lindquist et al., 2012; Picard, 2000). As well, often the reporting task may interrupt and distract from the emotional experience itself. To complement results from subjective reports, objective measures may be recorded including behavioral expressions (Ekman, 1993) and physiological (e.g. electrodermal Sequeira et al., 2009) and/or brain signals (e.g., functional magnetic resonance imaging (fMRI) Horikawa et al., 2020; Kober et al., 2008; Lindquist et al., 2012 and electroencephalography (EEG) Coan and Allen, 2004; Kim et al., 2013; Lin et al., 2010). Among these, EEG provides a direct measurement of distributed cortical brain dynamics with high temporal resolution.

Previous EEG-based emotion studies focused mainly on developing supervised learning approaches to optimize classification accuracy using subjective ratings of emotions as labels, e.g., affective dimensions or emotion categories. Such approaches contributed to the building of emotion classifiers applicable in affective computing, and to the identification of EEG biomarkers that maximally separate different labeled emotional conditions. However, these studies often overlooked the temporal dynamics of the emotional states, which normally develop over seconds to minutes of their elicitation by video (Nie et al., 2011) or other presented material (Lin et al., 2010). By contrast, Onton and Makeig (2009) conducted a self-paced emotion imagination experiment during which participants induced 15 distinct emotional states, each lasting 3–5 min, following their own imagination in response to

* Corresponding author.
E-mail address: shhsu@ucsd.edu (S.-H. Hsu).

https://doi.org/10.1016/j.neuroimage.2022.118873.
Received 16 May 2021; Received in revised form 8 November 2021; Accepted 4 January 2022
Available online 5 January 2022.

verbally-guided narrative suggestions. A conventional supervised learning approach assuming that emotion-related brain activity was stationary within each 3–5 min emotion imagination period, applied to the resulting EEG datasets, struggled to achieve high emotion classification accuracy with large differences in classification performance across emotion categories (Kothe et al., 2013).

Furthermore, most EEG-based emotion studies often assumed that subjective ratings of emotional experiences (e.g., valence and arousal used in popular DEAP dataset Koelstra et al., 2011) or group-validated labels of emotional stimuli (e.g., averaged across a pool of subjective ratings Lin et al., 2010) represent the actual emotional state changes in the participants. Few studies have aimed to validate or challenge these assumptions by exploring the relationship between the underlying EEG dynamics and the emotion labels, providing evidence for answers to some key questions in the field: (a) Is the subjective report of emotional experiences consistent with objective measures from neurophysiological recordings?, (b) Should differences between emotions be better described using affective dimensions or emotional categories (Barrett, 1998; Cowen and Keltner, 2017; Mauss and Robinson, 2009)?, (c) What are the most prominent spatiotemporal EEG dynamic changes during emotional experience?, and (d) Are activities in specific brain regions associated with specific emotions (Lindquist et al., 2012)?

These questions motivate the application of unsupervised learning approaches to characterize emotional state changes by clustering emotional states in terms of EEG activity differences rather than using subjective labels. To our knowledge, only a few studies have used an unsupervised-learning approach in EEG-based emotion studies. For example, a Gaussian mixture model (GMM) was used for learning segment-level variability in non-stationary EEG (Zhuang et al., 2014), a hypergraph representation was introduced to capture hidden structures in EEG signals across emotion trials (Liang et al., 2019), and a deep belief network (DBN) and hidden Markov model (HMM) were employed for unsupervised feature learning and tracking of emotional stage switching (Zheng et al., 2014). However, these studies did not explore the relationship between the spatiotemporal structure of the EEG and the associated emotion state labels.

Alternatively, adaptive mixture independent component analysis (adaptive mixture ICA, AMICA) (Hsu et al., 2018a; Palmer et al., 2008) assumes multidimensional data can be modeled by an independent component analysis mixture model (ICAMM) (Lee et al., 2000) in which each model represents a decomposition of the data into (possibly contiguous) segments associated with a distinct set of statistically independent component processes. Previous studies have shown that ICAMM, or its more advanced instantiation in the AMICA algorithm (Palmer et al., 2008), can separate EEG activities during different sleep stages (Hsu et al., 2018a; Salazar et al., 2010), fluctuations of drowsiness (Hsu et al., 2018a; Jung et al., 2000), mental state changes during memory test (Safont et al., 2017), guided meditation (Hsu et al., 2018b), and emotional video watching (Ran et al., 2020).

This study aims to employ multi-model AMICA as an unsupervised-learning approach for exploring the brain-state dynamics during different emotional experiences. We applied AMICA to a dataset from Onton and Makeig (2009), containing high-density (250-channel EEG data) collected from 31 participants during a self-paced emotion imagination experiment in which the participants induced a series of 15 emotional experiences by following recorded voice narratives with eyes closed, using their own imagination. Marrying a promising unsupervised method for modeling EEG nonstationarity with a unique high-density EEG dataset containing fifteen 3–5 min periods of emotional experience targeting fifteen different emotions enabled us to investigate, (1) whether such data-driven segmentation might self-organize and then be clustered according to affective dimension (e.g., valence), or separate into distinct models for each emotion, and (2) how does EEG spatiotemporal dynamics vary across self-paced experiences of imagined emotion and across participants. We study the putative cortical brain-based sources active during emotion imagination periods and differences in AMICA models across emotions and participants. These results provide evidence that emotional processes vary across emotions, but not necessarily consistent across participants. Finally, we discuss the studys limitations and make suggestions for the future research, such as comparing the results with those obtained from supervised methods to validate the consistency of findings and investigate emotion-specific activities.

2. Materials and methods

2.1. Dataset and preprocessing

2.1.1. Dataset

The dataset used in the study contains 31 recordings collected as described in Onton and Makeig (2009). The data are available at HeadIT website (http://headit.ucsd.edu, *Imagined Emotion with Continuous Data* under *Studies*). EEG data were collected from 250 scalp channels using a Biosemi ActiveTwo system (Amsterdam, Netherlands) at a sampling rate of 256 Hz per channel with 24-bit resolution. Caps with a custom whole-head montage were used to position the electrodes on most of the scalp, forehead, and lateral face surface, omitting chin and fleshy cheek areas. Relative 3D locations of the electrodes for each participant were recorded (Polhemus, Inc.). For more details, see Onton and Makeig (2009) and the description on the HeadIT website.

2.1.2. Experiment and participants

Thirty-one volunteer participants participated in the emotion imagination experiment (19 females; age 25.5 ± 5 years) and gave informed consent in accordance with UCSD institutional review board requirements. All participants reported being able to induce realistic emotional experiences for most of the suggested emotions by following a recorded verbal narrative and using their own imagination. For example, for “anger the narrative suggested that the participant either recall a situation in which they had been angry, or else imagine a scene in which they would become angry, giving two examples (“You find someone smashing your car for no reason with a baseball bat.). Throughout the experiment, participants were seated comfortably with eyes closed in a quiet, dimly-lit room, and listened to the narratives through ear-bud earphones.

Fig. 1 a shows the paradigm of the self-paced emotion imagination experiment. Each session began with 2 min of silent eyes-closed rest, followed by pre-recorded general instructions and then a 5 min guided relaxation induction to promote a relaxed, inwardly-focused state of mind. In 15 following task periods, the participant was presented with a sequence of recorded voice-guided inductions instructing them to recall or imagine scenarios for stimulating a vivid, embodied experience of the suggested emotion. Participants were encouraged to pay attention to their somatic state, as such attention tended to increase the vividness and duration of their emotional experience (Damasio, 1999). Participants were told to take as much time as they needed to recall or imagine a scenario that would induce a realistic experience of the suggested emotion. No external time limits or indicators were provided. Participants pressed a right-hand held thumb button once to signal they had begun to have a somatic experience of (i.e., to “feel) the targeted emotion, then a left-hand held button when the imagined scene and feeling began to wane, triggering (40 s) instructions to again relax (through a following silent 10 s period), letting go of the previously experienced scene and feeling. Then a 10 s audio clip prepared them for the next emotion imagination period.

The 15 emotional imagination periods were presented in a pseudo-random sequence that alternated between one of eight positive-valence emotions (love, joy, happiness, relief, compassion, contentedness, excitement, awe) followed one of seven negative-valence emotions (anger, jealousy, disgust, frustration, fear, sadness, grief). The experimental session was about 80 min in length. Button-press delimited durations of the active emotion imagination periods varied between 43 s and 12 min;

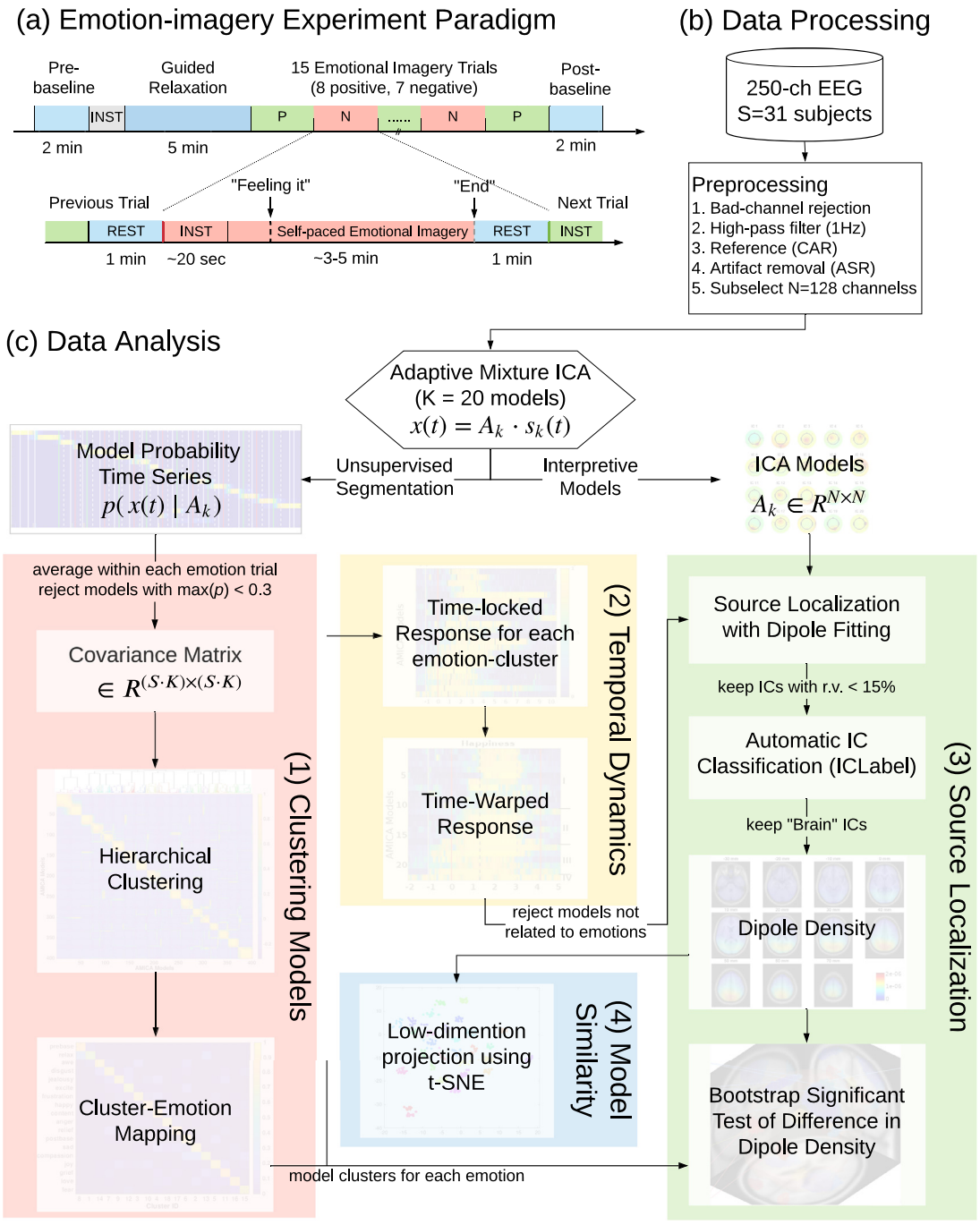

Fig. 1. (a) Experimental paradigm for the self-paced emotion imagination experiment. (b) Flowchart of the EEG data processing prior to data decomposition by adaptive mixture ICA (AMICA). (c) The AMICA model: EEG data, $x(t)$, are modeled as mixing matrices, A_k, times independent components (ICs) activities, $s_k(t)$, for model k. Schematically presented post-AMICA data analysis consists of *(1) Clustering Models*: hierarchical clustering of AMICA models across participants to examine the relationship between emotion imagination and changes in EEG dynamics detected by the multiple AMICA models, *(2) Temporal Dynamics*: exploring temporal dynamics of emotional responses through time-locked and time-warped analysis, *(3) Source Localization*: mapping IC equivalent-dipole models of source location for AMICA model clusters active during the same emotions to dipole density to examine source density differences during different emotional imagination periods, and *(4) Model Similarity*: projecting dipole density of individual AMICA models to low-dimension representations through t-distributed stochastic neighbor embedding (t-SNE) to explore differences across emotions and participants.

most lasted between 3 and 5 min. Transcripts of the verbal narratives are available in the Supplement Information to Onton and Makeig (2009).

2.1.3. Data preprocessing

The data obtained from the HeadIT portal were already pre-processed as described in Onton and Makeig (2009) including removal of bad channels (e.g., involving electrodes with poor skin contact, judged by their grossly abnormal activity patterns), leaving 134–235 channels per subject (214 ± 18, mean ± SD), re-referencing to digitally linked mastoids, and digital filtering above 1 Hz. Data periods containing broadly distributed, high-amplitude muscle noise and other irregular artifacts identified by tests for high-kurtosis or low-probability activity were removed from analysis using EEGLAB functions (Delorme and Makeig, 2004; Delorme et al., 2007). The occurrence of eye blinks,

other eye movements, or tonic muscle tension artifacts were not criteria for data rejection, as these were typically well-separated by ICA decomposition into separate component processes that were withheld from subsequent analyses.

Further preprocessing steps were applied to the data prior to the application of AMICA decomposition as schematized in Fig. 1b. Data before the first eye-closed rest session and after the last eye-closed rest session were removed. Data were then re-referenced to a common average reference. Artifact subspace reconstruction (ASR) (Kothe and Jung, 2016; Mullen et al., 2015) with a cutoff parameter 20 was applied for automatic removal of large-amplitude artifacts like electrode pops and motion artifacts, which has been shown to improve subsequent ICA decomposition (Chang et al., 2019).

Data were reduced to 128 channels by sub-selecting channels involving scalp electrodes that were maximally evenly spaced using the *loc-subsets()* function in EEGLAB. The reason for this decimation was to address the following three practical issues: insufficient samples for learning a larger number of model parameters (the size of the unmixing matrix is proportional to the square of the number of channels); computation time (as constrained by resources available); variations across participants in the number of retained channels. Effects of channel number on AMICA decomposition performance are further discussed in Section 4.5.

2.2. Adaptive mixture independent component analysis (AMICA)

AMICA assumes that the multi-channel time-series data can be described by a mixture of independent component analysis (ICA) models. Each model decomposes some portions of the data (possibly non-contiguous) into independent component (IC) effective source activities. Each IC has an individual probability density function learned from the data and parameterized as a mixture of generalized Gaussians. For mathematical description of the AMICA model, its learning algorithm, and effects and selection of learning parameters, see Section in the Supplemental Materials. For more technical details, see Palmer et al. (2008) for the algorithm and Hsu et al. (2018a) for validation.

In this study, 20 AMICA models, 128 ICs (same as the number of channels), and (only) one generalized Gaussian were used. The choice of 20 models for AMICA was based on an implicit assumption of the number of possible EEG states during the experiment, e.g., at most 15 emotion states, three baseline periods, plus inter-trial rest periods. The effect of the number of models on the learning performance of AMICA has been investigated in Hsu et al. (2018a) and is here further addressed in Section 4.1.

The AMICA sphering transformation option was engaged ($do\_pca = 1$), and data samples with low probabilities of model fit were rejected ($numreg = 5, rejstart = 2, rejint = 5$) from further use for learning AMICA parameters to alleviate the effects of transient artifacts such as electrode pops and discontinuities. AMICA uses expectation-maximization (EM) to estimate the parameters that maximize data likelihood under the learned model(s), using an efficient implementation with a parallel computing capability (Palmer et al., 2008). AMICA code is available at https://github.com/japalmer29/amica and as an open-source plug-in for EEGLAB (Delorme and Makeig, 2004). The computations were run on the Comet high-performance computing resource at the San Diego Supercomputer Center with support from the Neuroscience Gateway (NSG) project (Sivagnanam et al., 2013). Maximum learning steps were set to 2000; this required about 48 h of compute time on one computing node with 24 threads. Retrospective analysis showed that data likelihood reached a plateau near 1000 steps.

It is worth noting that AMICA failed, returning a data likelihood of zero when applied to 2 of the 31 recordings (Participant IDs 26 and 32) regardless of the numbers of models and channels used, possibly due to numerical round-off errors when computing data likelihoods close to zero. These two recordings were thus removed from further analysis.

As an unsupervised learning approach assuming the ICA mixture model, AMICA learns the underlying data source distributions (temporal pdfs) and quantitatively assesses continuous changes in EEG pattern likelihood under each learned model, thereby identifying model transitions. Specifically, the likelihood of each ICA model being active can be represented as the normalized data likelihood given the estimated parameters of each model; this may be referred to as “ICA model probability” that indicates the goodness-of-fit of the ICA model to each data sample. Subsequent learning steps focus learning for each model on data points relatively likely to follow that model, thus segregating the data into (possibly non-contiguous) model domains. A strength of AMICA is that it provides interpretable models allowing the characterization of the typically focal spatial distribution and detailed time series of (here) 128 active brain (as well as non-brain “artifact) source processes for each model for time points within the respective model time domain.

2.3. Post-AMICA data analysis

2.3.1. Clustering AMICA models across participants

To characterize the relationship between AMICA models and emotion states, we clustered across participants the AMICA models active in the same experiment sections. The procedure is summarized in Fig. 1c *(1) Clustering Models* and detailed below. An 18-by-1 feature vector was created for each model consisting of mean model probability between the felt emotion-surge button presses to the end of the emotion trial for each of 18 periods (15 emotion imagination periods, identified by emotion, plus guided relaxation, pre-session baseline, and post-session baseline). The models with low mean probability (i.e., below 0.3) in all 18 periods were rejected from further analysis, resulting in retention of 401 of the 580 (on average, 13.8 of the 20 models for the 29 participants) models. After reordering the emotion periods into an (arbitrary) common order, correlations (ρ) between each pair of these period model probability feature vectors were then computed, giving a correlation matrix of the size 401×401. Agglomerative hierarchical clustering into 18 model clusters was then applied to the correlation distance matrix ($1 - \rho$) to cluster models across participants having highly-correlated feature vector patterns, using the *linkage()* (with group average) and *dendrogram()* functions in MATLAB. The feature vectors within each model cluster were then averaged to obtain a summary cluster-to-emotion mapping to examine the relationship between emotion imagination trials and EEG dynamics segmented into AMICA models with similar probability profiles.

2.3.2. Time-locked and time-warped temporal dynamics of model transition

Fig. 1 c *(2) Temporal Dynamics* illustrates the procedure we used to explore spatiotemporal changes in EEG dynamics during emotion imagination. Model probability time-series of the AMICA models in the cluster identified according to Section 2.3.1 were smoothed using mean probabilities in successive (non-overlapping) 5 s windows time-locked to the guided narrative beginning each emotion-imagination trial, from 2 min before until 8 min after this event. To address the variability in trial lengths across participants, the model-probability time courses were linearly time-warped using the *timewarp()* function in EEGLAB (Delorme and Makeig, 2004) to equate median participant-response delay (i.e., the median time elapsed from imagination period onset to first button press) and median trial length (i.e., the time elapsed from the first button press to end of the trial) so as to be able to compare model probability courses differences across participants.

2.3.3. Criteria for categorizing model types based on activation time courses

To help identify and interpret models associated with particular emotions based on their probability time courses, we categorized model probability time courses into four types, as illustrated in Fig. 3b. For each model, we calculated the average model probabilities in four time periods during the emotion imagination trial: P_{pre}: from 2 min before to

Table 1
Dominance time series Types for models in each of the 15 model clusters, labeled by their dominant period emotion. The criteria for categorizing models into Types are described in Section 2.3.3.

Emotion \ Types	I	II	III	IV	Total	I + II (%)	III (%)	IV (%)
grief	10	8	0	2	20	90	0	10
excitement	9	8	3	2	22	77	14	9
happiness	10	6	5	1	22	73	23	5
love	4	9	5	0	18	72	28	0
fear	10	7	5	2	24	71	21	8
relief	6	7	4	2	19	68	21	11
anger	8	9	3	5	25	68	12	20
sadness	7	5	4	2	18	67	22	11
disgust	8	4	4	2	18	67	22	11
awe	9	6	6	2	23	65	26	9
joy	7	3	5	2	17	59	29	12
compassion	2	5	2	3	12	58	17	25
jealousy	4	6	5	3	18	56	28	17
frustration	3	7	7	4	21	48	33	19
contentment	2	4	8	8	22	27	36	36
average	6.6	6.3	4.4	2.7	19.9	64	22	13

the end of the previous trial; P_{rest}: from the end of the previous trial to the start of the current trial instruction; P_{instr}: from the start of the current trial instruction to the first button press; P_{feel}: from the first button press to the end of the current trial.

Selection criteria for the four model types were: Type IV: $P_{pre} > 0.5$, "activated" (increased in model probability) from the previous emotion imagination trial period; Type III: excluding Type IV, remaining models with $P_{rest} > 0.5$, meaning the model was activated during the rest period and before the current trial started; Type II, remaining models with $P_{instr} > 0.5$, indicating the model became activated after the guided emotion narrative began; and Type I, remaining models with $P_{feel} > 0.5$ - those only activated following the first button press through the end of the trial.

These model types provide a semi-quantitative way to compare temporal dynamics across emotions. The distribution of model types for each emotion-associated model cluster is summarized in Table 1. For the subsequent source localization and dipole density analyses, type III and type IV models were withheld since they became active before the trial started, suggesting they were not associated with the specific trial emotion.

2.3.4. Source localization with dipole fitting

To characterize and compare the spatial distribution of brain effective source ICs for AMICA models in each model cluster, we applied equivalent dipole source localization to the model learned IC scalp maps using DIPFIT2 (Acar and Makeig, 2010), available in EEGLAB. Here, electrode locations were manually co-registered to the template head model (a 4-layer Boundary Element Method model). A single equivalent current dipole in the head model was fit to each IC scalp map (projection pattern) learned for each IC by the AMICA decomposition. For justification for use of this source localization method, see (Delorme et al., 2012). Initial coarse-grid and subsequent fine-grid search for the best-fitting dipole location were applied using DIPTFIT2. Non-dipolar ICs (those with a residual variance of the model equivalent dipole scalp projection from the learned IC scalp map larger than 15%, as well as those ICs whose equivalent dipole was more than 5 mm located outside the template brain compartment were removed from further analysis. On average, 60% of the ICs (approximately 76 of 128) were retained across participants.

2.3.5. Automatic IC classification

To investigate and interpret the AMICA models, we categorized the independent components (ICs) of all the models into seven types (Brain, Eye, Muscle, Heart, Channel Noise, Line Noise, and Other) using an automatic IC classifier (ICLabel) (Pion-Tonachini et al., 2019), a pre-trained neural network based on IC power spectra and spatial projection patterns (scalp map). To improve the performance of ICLabel for multi-model AMICA, we weighted the IC power spectra by the normalized log-likelihood of the model to which the ICs belonged. We used the Lite version of ICLabel as this was faster to compute and gave results comparable to those of the default version. ICLabel code and a detailed tutorial can be found in its github repository https://github.com/sccn/ICLabel.

2.3.6. Equivalent dipole density comparison and bootstrap significance testing

Overall 3D spatial equivalent dipole densities for the selected ICs (i.e. ICs with r.v.<15%, located inside the head model and classified as "Brain" by ICLabel) from all the models active in each experiment period cluster were obtained by summing over each IC's contribution spatially blurred in the template head (to reflect possible dipole location error, participant head differences, etc.) using a spatial Gaussian-kernel with full width at half maximum (FWHM) of 8.5 mm. The equivalent dipole densities were normalized such that the dipole-density values summed to 1 across all voxels, suggested for comparison between models to reduce the effect of the number of dipolar ICs in the models. Here, we used the *dipoleDensity()* function in EEGLAB. Equivalent dipole density was computed for each AMICA model for the t-SNE visualization described in Section 2.3.7.

To test for significant differences between normalized dipole density in pairs of emotion-model clusters, we used bootstrapping. Two surrogate datasets were generated by drawing two sets of ICs randomly with replacement from the pool of ICs from all models within the two clusters. The number of ICs in each surrogate dataset equaled the number in each original emotion dataset. Effective source equivalent dipole densities for the two surrogate datasets were computed (as described above in Methods 2.3.4), and their 3D source density difference was obtained. After repeating this process 500 times, the distribution of surrogate differences for each voxel provided a difference value beyond which only 5% of the surrogate values fell, indicating a probability $p < .05$ (two-sided) than the actual dataset difference lies beyond that value.

To characterize brain model sources activated during emotion imagination, we also combined all models (likely) active during an emotion imagination period into one group (Emotion group) and compared these to models activated during relaxation, pre-session baseline, and post-session baseline (Baseline group) periods. The same bootstrapping approach was applied to obtain significance thresholds for differences in equivalent dipole density between the Emotion and Baseline groups, again using 500 bootstrap repetitions.

For improved visualization, regions with statistically significant dipole density differences were plotted in a reconstructed 3D image us-

ing code snippets from the Measure Projection Toolbox (MPT) (Bigdely-Shamlo et al., 2013). Each 2 × 2×2-mm template brain voxel was color-coded by differences in normalized dipole density values. The figure also shows 2D projections on sagittal, coronal, and axial MR slice images. Areas of significant differences are labeled with their Broadmann areas (BA) defined in the MNI coordinates using the MNI to Talairach mapping application in the Yale BioImage Suite (BIS) package https://bioimagesuiteweb.github.io/webapp/.

2.3.7. Visualization of model similarities using t-distributed stochastic neighbor embedding (t-SNE)

The t-distributed stochastic neighborhood embedding (t-SNE) method (van der Maaten and Hinton, 2008) was used to visualize the similarity between equivalent dipole densities for each AMICA model by projecting the 91 × 109×91-dimensional dipole density down to a best-representative two-dimensional space. t-SNE is a nonlinear dimensionality-reduction method that finds a map in a low-dimensional space best preserving pairwise similarities between neighboring points in the high-dimensional space by minimizing the Kullback-Leibler divergence between the two spatial distributions. Here we used the *tsne()* function in MATLAB with the Euclidean distance metric.

3. Results

3.1. Unsupervised segmentation of state changes during emotion imagination

Applied to the high-density EEG data, 20-model AMICA decomposition was able to characterize and separately distinguish differences in EEG dynamics across the emotion imagination experiment. Fig. 2a shows the model likelihood time courses (normalized data log-likelihoods) across the experiment for the 20 models of a sample subject. Model M1 was the only model with high model likelihood during the initial instruction and pre-session baseline periods. Near the beginning of the "relaxation" narrative, M1 likelihood faded and M2 emerged and became the most likely model. During the first emotion imagination period, following the "happiness" emotion narrative (solid green line), model M3 emerged as most likely until the end of the emotion period (solid white line). Model M4 became dominant in the next "fear" emotion imagination period. Transitions between dominant models sometimes occurred near onsets of the emotion induction narratives (solid green lines for positive emotions, red lines for negative emotions), but in other emotion periods (e.g., for "excitement", "contentment", "anger", and "grief"), transitions occurred at or near first button presses indicating the target emotion was now felt (dashed white lines). For this participant, AMICA separately modeled the EEG in nearly all the emotion imagination periods (all but "disgust") and the two resting baselines, giving a nearly one-to-one mapping between AMICA models and imagined emotions.

A similar correspondence between AMICA models and emotion imagination period markers was observed across most of the participants. The likelihood time series of the twenty AMICA models from all participants are shown in Figs. to in the Supplemental Materials. Across participants, twenty-model AMICA decomposition segmented the EEG activity into 10 to 18 models (mean, 14) that, typically, were dominant within only one task period, e.g., that were dominant (mean likelihood above 0.5) in 12 successive 5 s smoothing windows during the task period. The remaining models were most likely to reach dominance only sporadically or during brief rest periods between emotion imagination periods.

3.2. EEG activity in different emotion imagination periods fit by distinct AMICA models

Hierarchical Clustering (HC) was applied to identify models with similar emotion-related likelihood profiles across participants, e.g., models with high likelihoods in the same emotion imagination period(s). Fig. 2b shows pairwise correlations between all models across participants, sorted by the HC result shown at the top of the panel. Each of the 18 clusters contained 14–30 models. The resulting block-diagonal pattern shows clear separation between model clusters and high similarity within each model cluster, suggesting that models with similar activation patterns were returned consistently across participants.

Furthermore, the mean likelihood dominance patterns for each model cluster (columns in Fig. 2c) were focused on only one of the 18 experiment periods. That is, cluster models were dominant in one emotion imagination period, the two baseline periods, or during relaxation periods. The mean mapping between model clusters and experiment periods is nearly one-to-one. A few exceptions were mean model dominance in both pre-session baseline ("prebase") and "relax" periods (Clusters 1 and 2), and dominance in both "sadness" and "compassion" (Clusters 13 and 14). Note that these cluster pairs are closer in the HC dendrogram at the top of the figure. Hence, multi-model AMICA decomposition was able to blindly separate the EEG dynamics of each of the emotion imagination periods, allowing HC to identify model clusters of models across participants that were dominant during each of the 15 imagined emotions.

3.3. Differences in model transitions during emotion imagination

Once the best-fitting models for each emotion were identified, we examined their model likelihood time series to study the timing of state transitions to and from emotion imagination. Fig. 3a shows the (5 s smoothed) model likelihood time series for AMICA models in Cluster #8 (Fig. 2 c) that were dominant during "happiness" emotion periods, time-locked to period onsets (onset of the "happiness" narrative; red line). The participants emotion induction delay, the duration from narrative onset to first button press (black line), varied across participants from 0.6 to 7.2 min; whole emotion imagination period lengths were 1.5–9.4 min.

To better visualize transitions between dominant AMICA models during the emotion imagination trials, we time-warped the likelihood time series of the AMICA models between first (dashed black line) and final (dash gray line) button presses to their median values across participants, as shown in Fig. 3c for "happiness", "contentment", "anger", and "grief". Results for all 15 emotions are shown in Fig. and in the Supplemental Materials. Models became dominant (moving from low (blue) to high (yellow) model likelihood) at different points during the emotion imagination periods. Based on this observation, models were further categorized into four types based on their activation patterns, as shown in Fig. 3b. Some model dominance onsets were time-locked to first button presses (Type I) or to narrative onsets (Type II). Other models that were dominant before the emotion narrative were presented (Type III), while other models remained the dominance they gained during the previous emotion period (Type IV). Exact criteria for categorizing the four types of models are described in Section 2.3.3.

Fig. 3 c shows that the dominance patterns, as described by model type, varied significantly across emotions. For instance, 45% and 50% of the models in the "happiness" and "grief" clusters, respectively, were of Type I, the highest such percentages among the emotion clusters. This indicates EEG activity during these emotion periods was distinct in some way from that in the previous emotion and rest periods and emerged only as the participant began to feel the targeted emotion. The "anger" cluster included more Type II than Type I models, suggesting the modeled EEG changes tended to occur immediately following the emotion narrative (perhaps consistent with the oft-quoted phrase, "quick to anger."). Types III and IV models were the most common for "contentment" dominant models (of which 36% were of Type III and 36% of Type IV), suggesting the EEG activities during imagined "contentment" were not distinguishable from those during the preceding rest period or near the end of the previous emotion trial.

Numbers and percentages of model types for the 15 emotions are summarized in Table 1. Across all emotion models, 64% of model likelihood time series were Types I and II, which were more likely related to

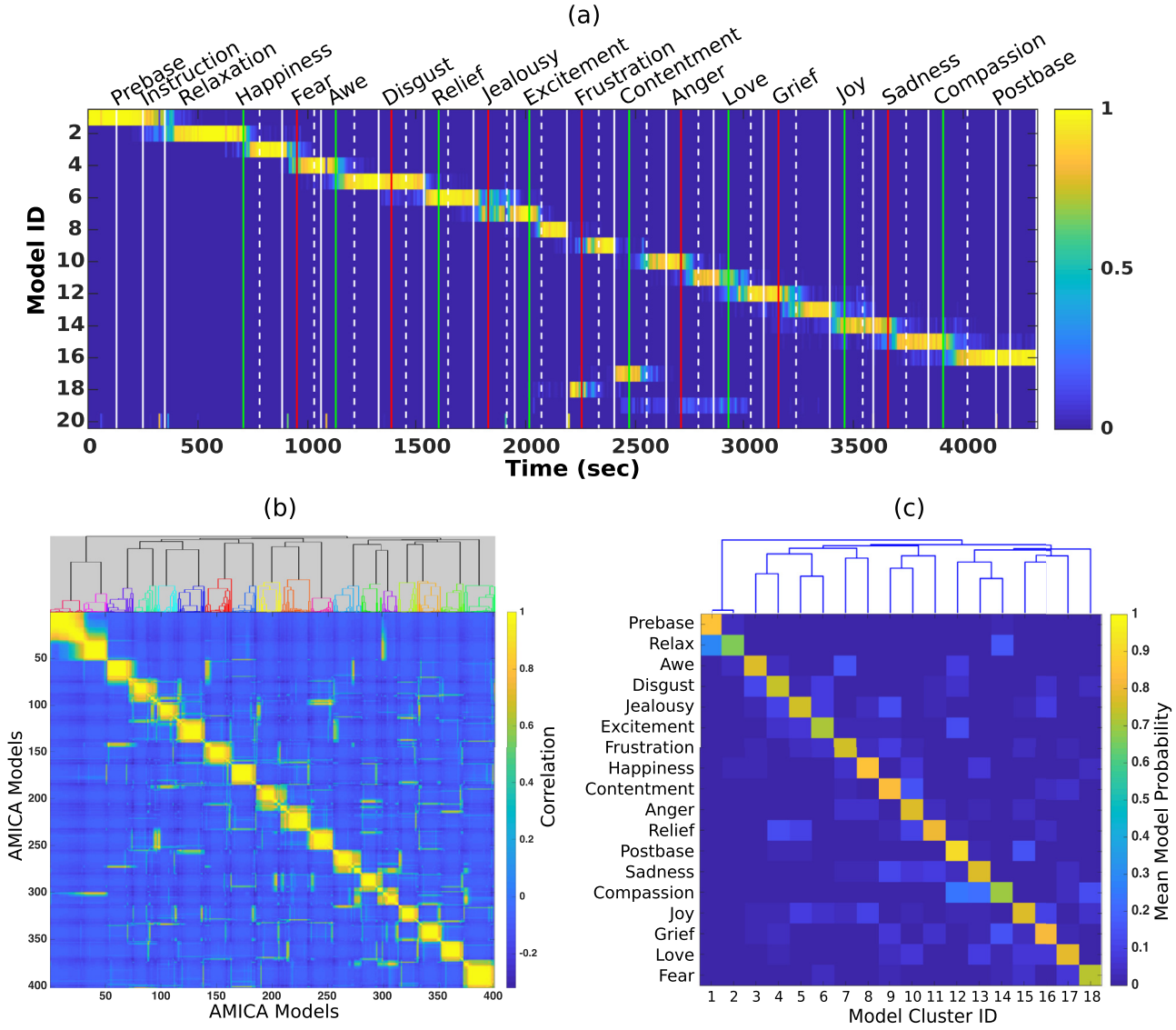

Fig. 2. (a) Probability time series of 20 models learned by AMICA for a single subject. Model ordered by the period of maximum likelihood. Colored vertical lines indicate the start of each emotion imagination trial with positive (green) and negative (red) emotions. Dashed and solid white lines mark first button presses that indicate the subject had begun to feel the emotion, and final button presses when the subject no longer felt the emotion. (b) Pairwise correlation coefficients between all AMICA models from all participants using model likelihood patterns, i.e., the mean model likelihood in each of the 18 task periods (15 emotion periods plus relaxation and two baseline periods), sorted according to the results of hierarchical clustering using 18 model clusters. A dendrogram for the clustering result is shown at the top; here colors represent individual model clusters. (c) Mean model likelihood patterns for each model cluster, with the dendrogram representing inter-cluster distances. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the imagined emotions given that their occurrence was approximately locked to the start of the audio narrative or the button press. The percentages ranged from 27% for “contentment” to more than 70% for “grief” and “happiness”. On average, 22% of models were Type III, suggesting the EEG activities learned by the models during the emotion imagination trials resembled those during the preceding rest period. Finally, 13% of models were of Type IV, and thus unlikely to be related to the specific emotion, given that they were also active during the previous emotion (of opposite valence). Possibly either the participants continued to feel the previous emotion, were not or no longer feeling that emotion when the trial ended, or the new emotion was not successfully elicited. Therefore, models of Types III and IV were rejected from the succeeding source distribution analyses.

Fig. 3 d summarizes, for each emotion, the distribution across the four Types of model dominance transitions, showing the median values of model likelihood across all models in the same emotion cluster for each 5 s window during the emotion imagination trials. Grand average model likelihood time series for all emotions are shown as gray curves. Models with Type IV dominance profiles were excluded as they did not likely capture specific emotion-related activities.

For all emotions, similar inverted U-shaped likelihood time courses were found, rising and falling near the beginning, middle, or end of the emotion imagination trials. The onset, slope and duration of the rising phase, the plateau, and the falling phase of the dominance period varied across emotions. Small increases in model likelihood began as early as the beginning of the rest period following the previous emotion trial (gray line). The rising phase continued, and for some emotions the slope increased during the audio narrative (red line). Model likelihood reached near peak values when the participants pressed a button to indicate they were feeling the emotion (dashed black line), and then plateaued until a final button press indicated the feeling had waned (dash gray line). Model likelihood decreased immediately after partic-

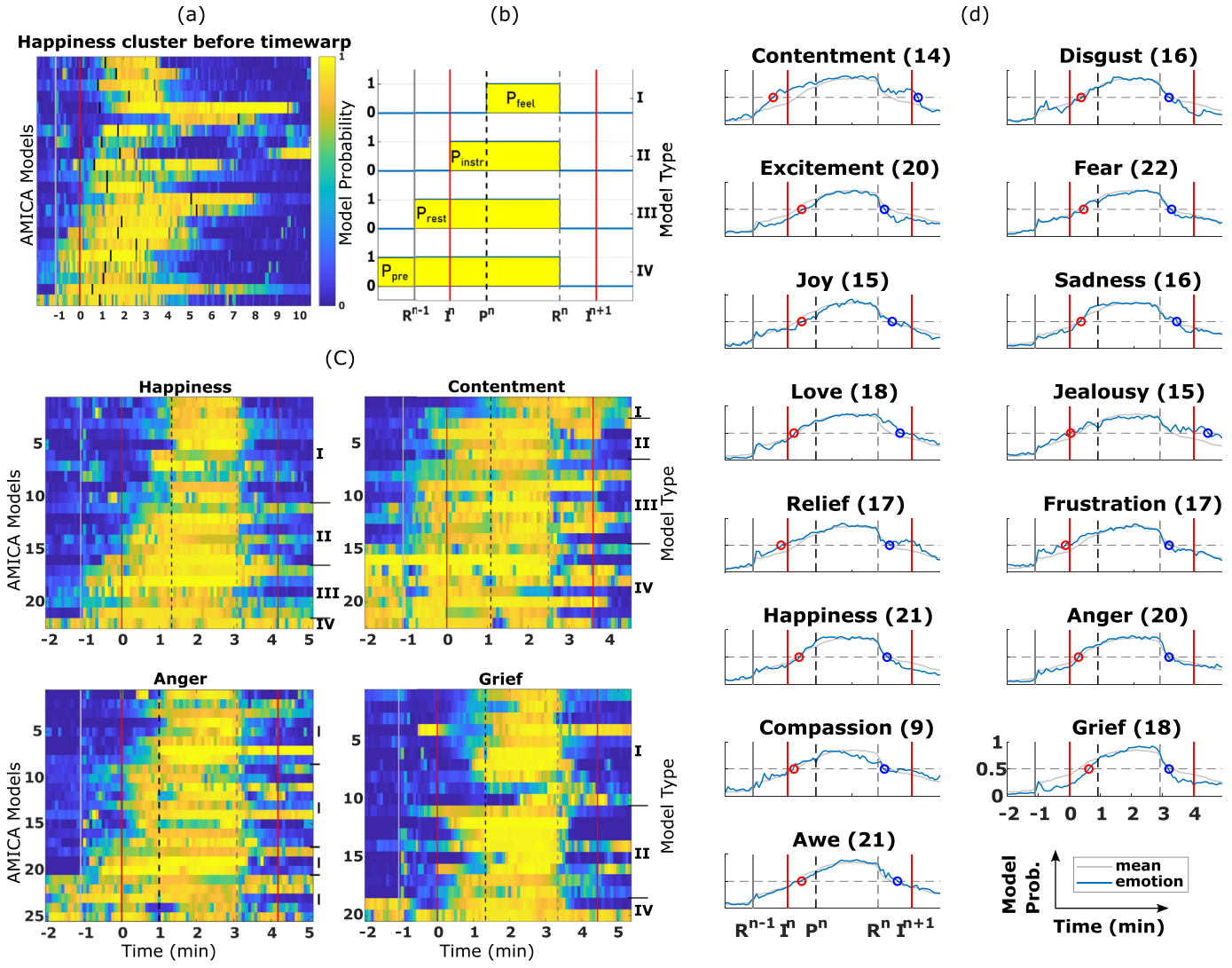

Fig. 3. (a) Model likelihood time series for all the AMICA models for 22 participants in the model cluster dominant during periods of imagined happiness. Vertical lines mark the end of the previous emotion period (white, R^{n-1}), the onset of the "happiness narrative (red, I^n), the first button press signaling that the subject was feeling the emotion (black, P^n), and the end of the imagination period as signaled by a second button press (gray, R^n). (b) Illustration of four model types (I–IV) based on the timing of their dominance periods (detailed criteria as defined in Section 2.3.3). (c) Model-likelihood time-series for "happiness", "contentment", "anger", and "grief" imagination periods. The times of the first "feeling it" button press (dashed black line) and the end of the imagination period (dash gray line) were time-warped to the median first-press latencies from all participants contributing a model to the cluster. Models were then sorted according to the four dominance pattern types. See Table 1 for detailed information. Results for all 15 emotions are shown in Fig. and in the Supplemental Materials. (d) Across participants median model-likelihood time series for each emotion cluster (blue curve) and mean likelihood time series across all emotions (gray curve), time-locked to the start of the interstitial resting period (gray vertical line) and the onset of the new emotion narrative (red line). The first and second button press times, time-warped to their median values across participants, are shown with dashed black and gray lines respectively. The number of models in each cluster is indicated in parentheses. Models related to positive (left column) versus negative (right column) emotions are sorted by their median imagination period lengths. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ipants after this trial end moment, as the participant entered the ensuing guided relaxation and rest period. These dominance cycles lasted 4–6 min on average.

Comparing temporal dynamics of model dominance across emotions, we found that the rising phases of the likelihood time series between I^n and P^n were more steeply sloping for happiness, sadness, anger, and grief, compared to the model mean profile (gray curves). The likelihood time series before P^n and after R^n were lower than average for grief and excitement and were higher for emotions including contentment and relief. For happiness, excitement, and most of the negative emotions, model likelihoods dropped faster than average immediately following the emotion period (R^n); while for emotions including love and contentment, model dominance tended to descend more gradually.

3.4. AMICA models learned a high percentage of dipolar brain ICs

Among all the models in the emotion dominance clusters, Table 2 shows that on average 24.4±8.6% of the 128 ICs from each model were classified by ICLabel as Brain ICs, while only 6.5±3.8% and 3.1±2.8% were classified as muscle-related or eye-related, respectively (Pion-Tonachini et al., 2019). The remaining ICs were mostly labeled Other (44.6±8.5%).

To validate the quality of IC decompositions produced by multi-model AMICA, we examined the percentage of dipolar ICs for each model (i.e., ICs whose projection patterns (scalp maps) matched the projection of a single equivalent dipole located in the template brain volume, a metric suggested by Delorme et al. (2012). We found that on

Table 2
(Top) Mean and standard deviation percentage of ICs in each category as classified by ICLabel across all AMICA models. (Bottom) Percent dipolar ICs (with r.v.<15%) within each class and among all dipolar ICs.

Percentage (%)	Brain	Muscle	Eye	Heart	Line Noise	Channel Noise	Other
Mean Percent	24.4	6.5	3.1	0.3	13.8	7.3	44.6
STD	8.6	3.8	2.8	0.4	5.2	4.4	8.5
Dipolar ICs	92.2	31.5	43.0	62.7	74.1	29.5	48.1
Total dipolar ICs account for	37.3	3.5	2.5	0.4	16.8	3.7	35.8

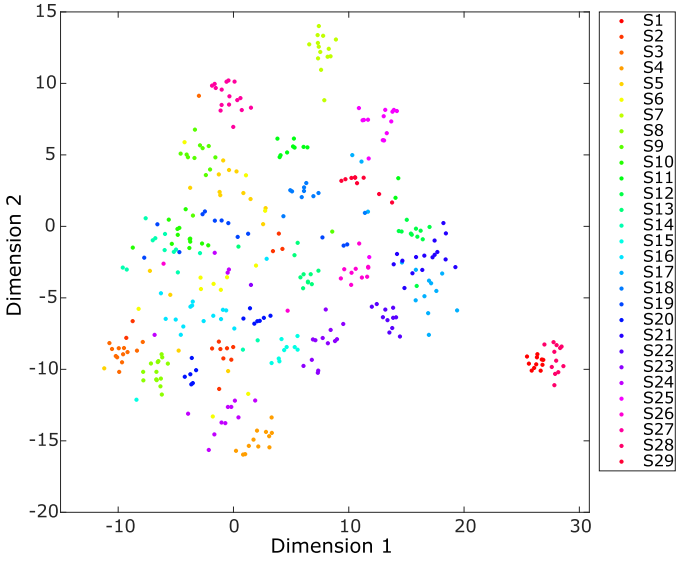

Fig. 4. Projecting dipole densities of all AMICA models onto a 2-dimensional feature space using t-SNE, color-coded by subject index.

average 60% of the ICs were dipolar, here defined as having less than 15% residual variance. Of the dipolar ICs, 37% of the ICs were labeled as Brain ICs, while only 3.5% and 2.5% were labeled as Muscle and Eye ICs.

We also found that the AMICA models rejected from the clustering analysis had a lower percentage of Brain ICs (20%) and a higher percentage of Other ICs (49%). However, there were no significant differences in the percentage of Brain ICs between models in different emotion clusters.

3.5. Dipole densities of brain ICs in the AMICA models

To quantitatively compare the equivalent dipole locations of brain effective source ICs across AMICA models, we computed dipole density of each AMICA model in the 3D template head brain model, excluding non-dipolar ICs (residual variance of the single equivalent dipole model >15%) and non-Brain ICs (identified by ICLabel). See Section 2.3.5 for details. Fig. in Supplemental Materials shows the normalized dipole density of dipolar brain ICs from all AMICA models during baseline and relaxation periods. We found the highest IC location density was in the bilateral occipital and central parietal regions.

The t-distributed stochastic neighborhood embedding (t-SNE) method was used to visualize (dis)similarities between dipole densities of different AMICA models by projecting the 91 × 109×91-dimension dipole-density features down to two dimensions. Fig. 4 illustrates that IC source location distributions were far more similar across emotion models for the same participant than for the same emotion across participants.

3.6. Spatial distribution of brain sources dominant during emotion imagination

When we first compared the equivalent dipole densities of AMICA model clusters between pairs of emotions, we found that the differences could be biased by which participants' models were in the clusters, because of variability in dipole density across participants (as shown in Fig. 4) and the limited number of models in each cluster (e.g., the "contentment" dominant cluster only had six Type I and II models, Table 1). To address this subject selection bias, we pooled model dominance clusters for the eight positive emotions (97 models) and compared the aggregated equivalent dipole density with the mean density for model clusters dominant during the seven negative emotions (96 models) using bootstrapping (see Section 2.3.6). No significant equivalent dipole density difference was found between the positive and negative emotion models.

Finally, we pooled together AMICA models from all 15 emotion clusters ("Emotion cluster") and compared the aggregated equivalent dipole density with mean density for model clusters dominant in the three baseline and relaxation periods ("Baseline cluster"). Fig. 5 shows lower dipole density in the Emotion cluster compared to the Baseline cluster in the premotor cortex (Brodmann Area (BA) 6), primary somatosensory cortex (BA 1), and primary motor cortex (BA 4), with the right hemisphere showing a broader effect than the left hemisphere. Additionally, Fig. 5 shows slight reductions in dipole density in the Emotion cluster in the left dorsolateral prefrontal cortex (BA 9), left anterior prefrontal cortex (BA10), and right insula (BA13). In contrast, the Emotion cluster showed significant increases in dipole density in the associative visual cortex (BA 19), left angular gyrus (BA 39), and ventral posterior cingulate cortex (BA 23).

4. Discussion

This study aims to explore EEG dynamics during emotional experiences using a unique "top-down approach. We applied unsupervised multi-model AMICA decomposition to high-density EEG data recorded during a self-paced emotion imagination experiment. The experiment used a guided imagination approach to include fifteen active emotions. This enabled us to investigate: (1) Could AMICA identify EEG dynamic models that became dominant only during emotion imagination? (2) How do such models relate to the natures of the 15 emotions (e.g., would the relationship be dimensional or categorical)? (3) When do relevant changes in spatiotemporal EEG dynamics occur during narrative induced, self-paced imagination of emotional scenes or scenarios, and how do they vary across emotions and participants?, and (4) What neurophysiological sources become active (or, e.g., available to discovery by ICA decomposition) during emotion imagination?

4.1. Data-driven multi-model AMICA decomposition recognizes changes in EEG dynamics during rest and imagination of different emotions

Here we first showed that decomposition of 1–2 h of high-density EEG data by multi-model AMICA was able to reliably resolve EEG models dominant in 10 or more of 18 distinct task periods throughout an

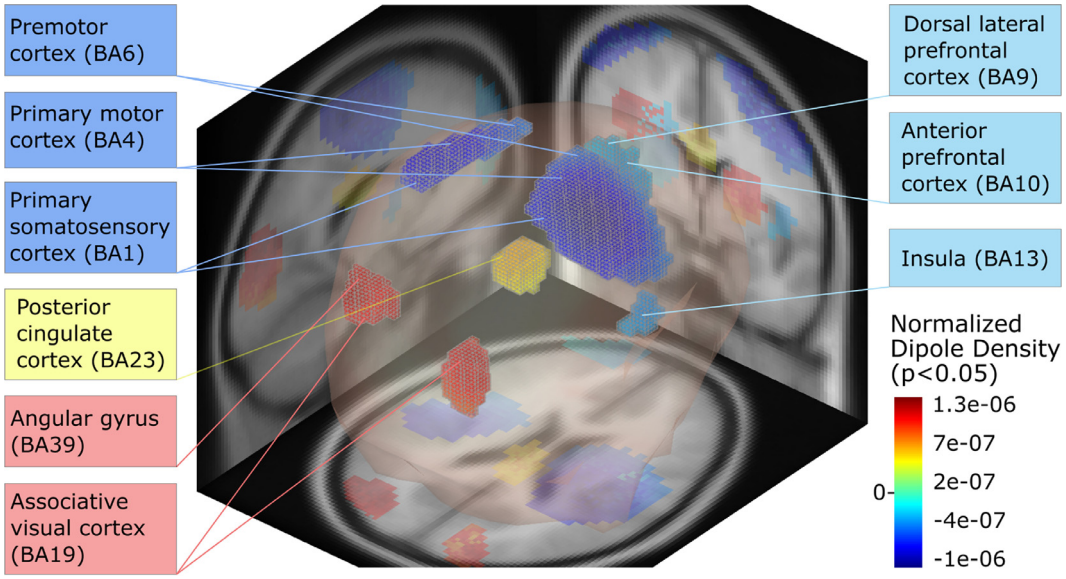

Fig. 5. The difference in normalized equivalent dipole density of dipolar brain ICs in AMICA models dominant in emotion periods versus models dominant during baseline periods (i.e., pre- and post-session baseline and relaxation periods), masked by the result of a significance test ($p < .05$) using bootstrapping. The result is superimposed on a 3D brain model with selected sagittal, coronal and axial slices of a template MR image shown for reference. Here Brodmann areas (BA) were identified by their MNI brain template coordinates.

emotion imagination experiment across 29 participants. EEG activities in each state could be separately modeled by one of 20 ICA models for partitions of the data learned by AMICA (Fig. 2a). Further, unsupervised segmentation of these states clearly corresponded to the beginning and end of each emotion imagination period.

Interestingly, the EEG models did not cluster according to important affective dimensions including valence or arousal. Rather, for each of the 18 experiment periods AMICA returned a distinct cluster of models across participants in which cluster models were only dominant (i.e., most likely to model the data) in that period (Fig. 2c). The clustered models were near-dominant in more than one experiment period in only a few cases such as during “prebase” and “relax” or in “sadness” and “compassion”. The latter conforms to self-reports by some participants that to experience “compassion” they imagined witnessing a sad scene (an accident victims suffering, for example).

While our results show that for most participants AMICA detected differences in EEG dynamics during imagination of most of the imagined emotions, models dominant during different emotions might have many shared dynamic aspects. We thus conducted an exploratory analysis, applying AMICA decomposition to the same dataset using different numbers of models. When 5 models were learned (Fig.), in some cases several consecutive emotion imagination trials shared the same dominant model (e.g., pre-session baseline, relaxation and awe for Model 1; compassion, fear, contentment, and jealousy for Model 5). When the number of learned models was increased to 10, several models were active in two successive trials of the same valence (e.g., frustration and anger for Model 2, joy and happiness for Model 3, sadness and grief for Model 4, love, relief and excitement for Model 5). When the number of models was increased to 15, model distinctions between individual emotion imagination periods became clearer, and when the number of models returned to 20, nearly every emotion period was distinctly separated into different dominant models. This combined evidence suggests that even though EEG activities in each of the emotion imagination periods could be distinctly separated when enough models were available, the periods for emotions with the same valence might share more similar EEG activities that could be account by one model when fewer models were available. We did not extend this exploration to other participants data.

4.2. Temporal dynamics of emotion-period models and their variability across emotions

Our earlier report showed that multi-model AMICA decomposition can assist in characterizing changes in EEG spatiotemporal dynamics with sub-second resolution (Hsu et al., 2018a). Here we found that transitions between best-fitting models could be well aligned to emotion imagination periods (Fig. 3). A third of the models dominant in each emotion imagination period became dominant as the participant listened to the emotion induction narrative (Type II models); another third became dominant as the participant pressed a thumb button to indicate they were actively feeling the suggested emotion (Type I models) (Table 1). For most of these models, model dominance “faded out” when the participant ended the period with a final button press and listened to the ensuing relaxation narrative. Thus, dominant model likelihood waxed and waned at the beginning and end of one emotion imagination period, providing strong evidence that the models separated EEG activities uniquely occurring (or co-occurring) during just one emotion period.

Comparing the temporal dynamics of different emotional-period models, we found that emotion trials such as grief, excitement, and happiness induced more distinctive activity shifts detected by AMICA, evidenced by a higher percentage of Type I and II models (Table 1 and Fig. 3). Some emotion-dominant models (during Love, Contentment, and Relief trials) remained dominant longer after the end of the imagination trial. These temporal features might be used as new measures to characterize differences in EEG supporting different emotions.

We also found that one-third of the models became dominant during the rest period prior to the emotion imagination trial (Type III models) or even during the previous emotion-imagination trial (Type IV models). Jealousy and Compassion trials had a higher percentage of Type III and IV dominant models (Table 1). Perhaps here the new emotion targets did not produce a change in the EEG model as the participant had become fatigued or otherwise failed to feel the requested emotion. In some cases, the EEG source dynamics prevailing during the new emotion may have been similar to those in the preceding resting state (Type III) or even the previous emotion (Type IV). In particular, the imagined experience of “Contentment” might well be supported by dynamics pre-

vailing during the preceding rest state; indeed, Contentment trials had the highest percentage of Type III and IV models (73%).

4.3. Emotions with less distinctive state changes in EEG activities during imagination were classified less accurately

In other cases, distinctive emotional states might not always have been achieved during emotion imagination. This is also consistent with the relatively low accuracy in emotion classification obtained in a previous study using this dataset (Kothe et al., 2013). There, Kothe et al. (2013) applied a machine learning model to data from a 12-subject subset given the goal of predicting the (positive or negative) valence of the emotion imagined emotion. They found that valence accuracies were the highest (above 90%) for Grief and Happiness and lowest for Awe, Jealousy, Contentment, and Frustration (below 65%) (see Fig. 1 in Kothe et al., 2013). Their valence classification accuracies for different emotions correlates well with the percentages of Type I and II models for each emotion in this study (Table 1). Thus, in this study emotions that were not successfully induced and/or those whose EEG dynamics were less distinguishable from relaxation (as evidenced by a lower percentage of Type I and II models) were the emotion that gave close to chance level accuracy in the Kothe et al. (2013) study.

These results pose AMICA as a promising tool for assessing whether and when emotion-relevant states are induced in emotion studies (e.g., studies comparing emotion induction paradigms). Possibly, using multi-model AMICA to select appropriate time windows for classification might improve emotion classification performance.

4.4. Neurophysiological sources activated during emotion imagination and inter-subject variability

Toward a neurophysiological interpretation of the dominant AMICA models and model switches, we first examined the quality of the AMICA decomposition. On average, 60% of ICs were dipolar (i.e., their scalp maps could be fit with a single equivalent dipole located in the template brain volume with less than 15% residual variance). In Delorme et al. (2012) comparing blind decomposition methods to 71-channel data from another study, a higher percentage of dipolar ICs was associated with a more successful ICA decomposition method, i.e., produced more overall mutual information reduction. In their study, AMICA gave the most dipolar components and highest mutual information reduction of 22 blind decomposition methods. Our (60% dipolar) result is consistent with the AMICA results in that report. This suggests that the decomposition quality of AMICA did not degrade significantly when learning a large number (20) of models.

Further, ICLabel (Pion-Tonachini et al., 2019) revealed that 37% of the dipolar ICs were brain-related ICs (Table 2), suggesting AMICA models were able to identify multiple independent brain processes. Notably, we have observed a higher percentage of “Others ICs (45%) than in single-model ICA decomposition (29%), which is likely due to the high number of ICs (128×20) found in the data by multi-model AMICA decomposition.

Dipolar and brain ICs were used to compute the spatial distribution of dipole density in the brain model. A sample dipole density plot of the dipolar brain ICs in the Baseline dominant model clusters (Fig.) shows dipolar brain sources were located most densely in the occipital and parietal cortices.

Visualizing dipole densities of all AMICA models in a 2D t-SNE feature space (Fig. 4) revealed that the dipole distributions were more highly similar across emotion-dominant models for the same participant rather than for the same emotion across participants. That is, each participant appears to have a core set of brain ICs whose locations in the brain model did not change significantly across AMICA models. It is, however, in IC probability density functions (pdfs) that AMICA uses to assign data to models it is not clear whether small differences in the locations of model equivalent dipoles in different models for the same participant represent inherent (noise) variability or actual small changes in cortical generating area (and source pdfs) in different emotions.

The minor differences in dipole distributions between models dominant for different emotions did not rise to significance across participants. Further, when we pooled together model dipole clusters dominant in positive emotions and compared them to those dominant during negative emotions, we found no significant dipole density difference ($p > .05$). This result is consistent with a review paper on neuroimaging evidence of brain activity supporting emotion (Lindquist et al., 2012), in which meta-analytic analysis found “little evidence that discrete emotion categories can be consistently and specifically localized to distinct brain regions.”

However, it is important to point out another explanation to our finding of non-significant dipole-density differences between emotions. Brain activity during the emotion imagination periods might be too individualized to exhibit consistent IC expressions across participants. For example, it might be determined by specific scenarios the participant imagined rather than by the overall emotion experienced by the participant.

We did find consistent differences across participants in dipole density of models dominant during periods of emotion imagination compared to models accounting for intervening periods of guided relaxation. A possible interpretation is that some brain areas began to or no longer produced the same local spatially coherent EEG signals when the participant shifted from self-absorbed relaxation to active imagination of emotionally charged scenarios. The primary somatomotor and premotor cortex had lower dipole density (Fig. 5) during emotion imagination than during relaxation. This may relate to absence, during emotion imagination periods, of specific “idling rhythm (e.g., circa 10-Hz mu rhythm) activities in motor areas during guided relaxation. Right dorsolateral prefrontal cortex (DLPFC, BA9), left rostrolateral prefrontal cortex (RLPFC, BA10), and right insula (BA13) also exhibited lower model IC dipole density during emotion imagination.

Previous studies have reported that the right DLPFC is associated with behavioral inhibition and self-control (Aupperle et al., 2012; Kerestes et al., 2012; Morawetz et al., 2016; Ray and Zald, 2012; Shackman et al., 2009; Viviani, 2014). The left RLPFC is generally believed to be related to memory recall and coordination of information processing (Ramnani and Owen, 2004), and is found to be activated during emotion-regulation tasks (Bramson et al., 2018; Koch et al., 2018). Recent studies found that the anterior insula could be related to emotion recognition and emotional awareness (Craig, 2009; Motomura et al., 2019). The BOLD signal results could be consistent with our findings if the activity in these areas shifted from predominantly lower-frequency (sub-gamma) activity to low-amplitude high-frequency activity, which is more difficult for ICA to resolve into dipolar sources.

Areas with higher dipole density in models accounting for emotion imagination than intervening relaxation were posterior: bilateral associative visual cortex (BA19), which BOLD studies have shown to be activated during both actual and imagined visual information processing (Fink et al., 2018; Kaas et al., 2010), left angular gyrus, associated with self-referential memory retrieval (Seghier, 2013), and posterior cingulate cortex (PCC), known to be active during focused attention and memory retrieval (Leech and Sharp, 2014; Rolls, 2019). In a review paper, Lindquist et al. (2012) reported a set of brain regions commonly involved during emotion experience across discrete emotion categories. This includes the motor cortex supporting language and executive attention, the visual cortex connecting with areas involved in core affect like the amygdala, and the prefrontal cortex and medial posterior group involved with conceptualization (Kober et al., 2008).

Taken together, the emotion-related brain areas identified by AMICA decomposition are thus largely consistent with those identified in many neuroimaging studies. Further analysis of these results might study what types of independent EEG sources in these areas produced during emotion imagination.

4.5. Limitations and future work

Although AMICA successfully characterized brain-state changes during the emotion imagination experiment, here we did not find ways in which the resulting models are consistent within emotions rather than participants. In particular, we found no difference in propensity for ICA decomposition to find independent components localizing to the bilateral temporal areas shown by Onton and Makeig (2009) in these same data to exhibit clear valence-related activity (stronger high-frequency broadband activity to more positive emotions), a result since replicated in several fMRI BOLD signal studies using positive, neutral and negative valence stimuli (music, speech, etc.). Evidently, distinct EEG source activity in these areas was no more likely to be found by AMICA models accounting for positive versus negative emotion imagination periods, nor for guided relaxation versus self-produced emotion imagination periods. Future work with these data could explore more completely dynamic (rather than spatial) differences between the different model clusters. Future replication of this experiment might include a control condition presenting motor or sensory imagery in a perceptual task, to test whether emotion imagination trials and the sensorimotor perception conditions would share the same AMICA models.

Since the sub-selected 128-channel EEG montage for training AMICA models covered the whole scalp and partial facial and neck areas (i.e., “whole montage”), it could be argued that the facial and neck muscle artifacts significantly contributed to the separation of EEG activities between different emotion-imagination trials. To test the hypothesis, we sub-selected a subset of 128 channels only from electrodes placed above the ears (i.e., “scalp montage”), the AMICA decompositions showed consistent results in terms of separating EEG segments that corresponded to emotion periods. Besides, the IC classification of the 128-channel whole montage using ICLabel shows that the numbers of artifactual ICs such as eye, heart, and muscle components were much smaller than the number of brain ICs (Table 2). The evidence suggests that the unsupervised segmentation of AMICA was not solely due to the contributions of muscle and eye artifacts, especially from facial and neck muscles.

For a neurophysiological interpretation of the AMICA models, we used dipole density to map the high-dimensional AMICA parameters of each model onto the same brain space. This enables systematic comparison and statistical testing of spatial distributions across AMICA models. However, this mapping involves user-defined thresholds for selecting dipolar ICs and types of ICs, which would affect the resulting dipole density. Alternative approaches to systematically compare and quantify the distance between the AMICA models in a high-dimensional parameter-space would advance AMICA results’ interpretability and provide further neurophysiological insights.

As discussed in the previous section, AMICA models – albeit able to separate different emotions – seem to be individualized and specific to each participant. This could be the result of over-fitting, given that there were approximately 338K parameters to learn (for 20 AMICA models, each with 128 ICs) with 1.38M data samples (for a 90 min recording). In fact, we have tried to further sub-select 64 channels for training AMICA models. The empirical result showed that the AMICA models with 64 ICs could still consistently segment the EEG data, but transitions between segments (i.e., changes in which model likelihood was dominant) were not as distinct compared with the results from 128- and 250-IC AMICA.

An important future direction could be testing whether the AMICA models associated with emotion imagination can be found again in the EEG activity of a separate test session, or in sessions using emotion-eliciting approaches. Although training AMICA models imposes a relatively heavy computational burden, models trained on a first dataset could be used to make statistical inferences on test data in a near-real-time fashion, enabling online emotion decoding.

For future advancement of emotion classification, it is important to compare and even leverage the insights gained from unsupervised methods with those from supervised methods. For example, training a supervised classifier after unsupervised AMICA decomposition would enable emotion identification that may better generalize across participants. Comparing source locations from multi-model AMICA decomposition with those from single-model ICA applied to individual emotion periods in a supervised manner, we could examine the validity of subjective emotion labels and separate emotion activities from those not related to emotion.

5. Conclusions

Our results demonstrate that multi-model decomposition of high-density (128-channel) EEG data by AMICA can detect shifts in EEG dynamics associated with differences in task (here guided rest versus active emotional scenario imagination), even when the data are to be segregated into as many as twenty models. As in our earlier report (Hsu et al., 2018a), the timing of model transitions can accurately reflect timing of changes in task orientation, and differences in timing of model transitions (as studied here) may reveal details about the timing and degree of shifts in mental goals and focus.

Brain regions with the biggest difference in dipole density during emotion imagination compared to rest were identified in the left dorsal lateral and anterior prefrontal cortex, posterior cingulate cortex, right insula, motor cortex, and visual cortex, consistent with previous emotion studies. Yet, no significant difference in dipole densities was found between positive and negative emotions. The spatial distributions of brain-localizable ICs showed higher similarity within-subject across emotions than within-emotion across participants. The results presented here suggest that continued work to characterize essential differences in AMICA models dominant in different time periods could be valuable. This study provides and validates a framework, i.e., AMICA and post-AMICA analyses, for data-driven discovery of brain state dynamics in an emotion imagination experiment, shedding light on the neurophysiological underpinnings of emotional experiences, thereby improving the performance of emotion decoding for EEG-based affective computing and advancing our understanding of emotion.

Acknowledgments

This work was supported in part by the U.S. National Science Foundation (IIP-1719130) and the Army Research Laboratory (W911NF-10-2-0022). Dr. Makeig’s participation was funded by a grant from the U.S. National Institutes of Health (R01 NS047293-13A1) and by a gift to UCSD from The Swartz Foundation (Old Field, NY).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.neuroimage.2022.118873.

References

Acar, Z.A., Makeig, S., 2010. Neuroelectromagnetic forward head modeling toolbox. J. Neurosci. Methods 190 (2), 258–270.

Aupperle, R.L., Allard, C.B., Grimes, E.M., Simmons, A.N., Flagan, T., Behrooznia, M., Cissell, S.H., Twamley, E.W., Thorp, S.R., Norman, S.B., et al., 2012. Dorsolateral prefrontal cortex activation during emotional anticipation and neuropsychological performance in posttraumatic stress disorder. Arch. Gen. Psychiatry 69 (4), 360–371.

Barrett, L.F., 1998. Discrete emotions or dimensions? the role of valence focus and arousal focus. Cognit. Emot. 12 (4), 579–599.

Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M., Pollak, S.D., 2019. Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. Psychol. Sci. Public Interest 20 (1), 1–68.

Bigdely-Shamlo, N., Mullen, T., Kreutz-Delgado, K., Makeig, S., 2013. Measure projection analysis: a probabilistic approach to eeg source comparison and multi-subject inference. Neuroimage 72, 287–303.

Bramson, B., Jensen, O., Toni, I., Roelofs, K., 2018. Cortical oscillatory mechanisms supporting the control of human social emotional actions. J. Neurosci. 38 (25), 5739–5749. doi:10.1523/JNEUROSCI.3382-17.2018.

Chang, C.Y., Hsu, S.H., Pion-Tonachini, L., Jung, T.P., 2019. Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel eeg recordings. IEEE Trans. Biomed. Eng..

Coan, J.A., Allen, J.J., 2004. Frontal eeg asymmetry as a moderator and mediator of emotion. Biol. Psychol. 67 (1–2), 7–50.

Cowen, A.S., Keltner, D., 2017. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. Proc. Natl. Acad. Sci. 114 (38), E7900–E7909.
Craig, A., 2009. How do you feel–now? the anterior insula and human awareness. Nat. Rev. Neurosci. 10, 59–71. doi:10.1038/nrn2555.
Damasio, A.R., 1999. The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Houghton Mifflin Harcourt.
Delorme, A., Makeig, S., 2004. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21.
Delorme, A., Palmer, J., Onton, J., Oostenveld, R., Makeig, S., 2012. Independent eeg sources are dipolar. PloS One 7 (2), e30135.
Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 34 (4), 1443–1449.
Ekman, P., 1993. Facial expression and emotion. Am. Psychol. 48 (4), 384.
Fink, A., Rominger, C., Benedek, M., Perchtold, C.M., Papousek, I., Weiss, E.M., Seidel, A., Memmert, D., 2018. EEG alpha activity during imagining creative moves in soccer decision-making situations. Neuropsychologia 114, 118–124.
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 101060.
Hsu, S.H., Pion-Tonachini, L., Palmer, J., Miyakoshi, M., Makeig, S., Jung, T.P., 2018a. Modeling brain dynamic state changes with adaptive mixture independent component analysis. Neuroimage.
Hsu, S.H., Zi, Y., Wu, Y.C., Jackson, P.M., Jung, T.P., 2018b. Exploring mental state changes during hypnotherapy using adaptive mixture independent component analysis of EEG. In: Proceedings of the IEEE Biomedical Circuits and Systems Conference (BioCAS). IEEE, pp. 1–4.
Jung, T.-P., Makeig, S., Lee, T.-W., McKeown, M.J., Brown, G., Bell, A.J., Sejnowski, T.J., 2000. Independent component analysis of biomedical signals. In: Proceeding of the International Workshop on Independent Component Analysis and Signal Separation. Citeseer, pp. 633–644.
Kaas, A., Weigelt, S., Roebroeck, A., Kohler, A., Muckli, L., 2010. Imagery of a moving object: the role of occipital cortex and human Mt/V5+. Neuroimage 49 (1), 794–804.
Kerestes, R., Ladouceur, C.D., Meda, S., Nathan, P.J., Blumberg, H.P., Maloney, K., Ruf, B., Saricicek, A., Pearlson, G.D., Bhagwagar, Z., Phillips, M.L., 2012. Abnormal prefrontal activity subserving attentional control of emotion in remitted depressed patients during a working memory task with emotional distracters. Psychol. Med. 42 (1), 29–40. doi:10.1017/S0033291711001097.
Kim, M.K., Kim, M., Oh, E., Kim, S.-P., 2013. A review on the computational methods for emotional state estimation from the human EEG. Comput. Math. Methods Med. 2013.
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 (2), 998–1031.
Koch, S.B., Mars, R.B., Toni, I., Roelofs, K., 2018. Emotional control, reappraised. Neurosci. Biobehav. Rev. 95 (March), 528–534. doi:10.1016/j.neubiorev.2018.11.003. https://doi.org/10.1016/j.neubiorev.2018.11.003
Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I., 2011. Deap: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. 3 (1), 18–31.
Kothe, C.A., Makeig, S., Onton, J.A., 2013. Emotion recognition from eeg during self-paced emotional imagery. In: Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, pp. 855–858.
Kothe, C. A. E., Jung, T.P., 2016. Artifact removal techniques with signal reconstruction. US Patent App. 14/895,440.
Lee, T.-W., Lewicki, M., Sejnowski, T., 2000. ICA mixture models for unsupervised classification of non-Gaussian classes and automatic context switching in blind signal separation. IEEE Trans. Pattern Anal. Mach. Intell. 22 (10), 1078–1089. doi:10.1109/34.879789.
Leech, R., Sharp, D.J., 2014. The role of the posterior cingulate cortex in cognition and disease. Brain 137 (1), 12–32. doi:10.1093/brain/awt162.
Liang, Z., Oba, S., Ishii, S., 2019. An unsupervised eeg decoding system for human emotion recognition. Neural Netw. 116, 257–268.
Lin, Y.-P., Wang, C.-H., Jung, T.-P., Wu, T.-L., Jeng, S.-K., Duann, J.-R., Chen, J.-H., 2010. EEG-based emotion recognition in music listening. IEEE Trans. Biomed. Eng. 57 (7), 1798–1806.
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 (3), 121.
van der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605. http://www.jmlr.org/papers/v9/vandermaaten08a.html
Mauss, I.B., Robinson, M.D., 2009. Measures of emotion: a review. Cognit. Emot. 23 (2), 209–237.
Morawetz, C., Bode, S., Baudewig, J., Kirilina, E., Heekeren, H.R., 2016. Changes in effective connectivity between dorsal and ventral prefrontal regions moderate emotion regulation. Cereb. Cortex 26 (5), 1923–1937. doi:10.1093/cercor/bhv005.
Motomura, K., Terasawa, Y., Natsume, A., Iijima, K., Chalise, L., Sugiura, J., Yamamoto, H., Koyama, K., Wakabayashi, T., Umeda, S., 2019. Anterior insular cortex stimulation and its effects on emotion recognition. Brain Struct. Funct. 224 (6), 2167–2181. doi:10.1007/s00429-019-01895-9. https://doi.org/10.1007/s00429-019-01895-9
Mullen, T.R., Kothe, C.A., Chi, Y.M., Ojeda, A., Kerth, T., Makeig, S., Jung, T.-P., Cauwenberghs, G., 2015. Real-time neuroimaging and cognitive monitoring using wearable dry EEG. IEEE Trans. Biomed. Eng. 62 (11), 2553–2567.
Nie, D., Wang, X.-W., Shi, L.-C., Lu, B.-L., 2011. Eeg-based emotion recognition during watching movies. In: Proceedings of the 5th International IEEE/EMBS Conference on Neural Engineering. IEEE, pp. 667–670.
Onton, J.A., Makeig, S., 2009. High-frequency broadband modulation of electroencephalographic spectra. Front. Hum. Neurosci. 3, 61.
Palmer, J.A., Makeig, S., Kreutz-Delgado, K., Rao, B.D., 2008. Newton method for the ICA mixture model. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2008. IEEE, pp. 1805–1808.
Picard, R.W., 2000. Affective computing. MIT Press.
Pion-Tonachini, L., Kreutz-Delgado, K., Makeig, S., 2019. Iclabel: an automated electroencephalographic independent component classifier, dataset, and website. Neuroimage 198, 181–197.
Ramnani, N., Owen, A.M., 2004. Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. Nat. Rev. Neurosci. 5 (3), 184–194. doi:10.1038/nrn1343.
Ran, S., Hsu, S.-H., Jung, T.-P., 2020. Examining the relationship between EEG dynamics and emotion ratings during video watching using adaptive mixture independent component analysis. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, pp. 1491–1497.
Ray, R.D., Zald, D.H., 2012. Anatomical insights into the interaction of emotion and cognition in the prefrontal cortex. Neurosci. Biobehav. Rev. 36 (1), 479–501. doi:10.1016/j.neubiorev.2011.08.005.
Rolls, E.T., 2019. The cingulate cortex and limbic systems for emotion, action, and memory. Brain Struct. Funct. 224 (9), 3001–3018. doi:10.1007/s00429-019-01945-2. https://doi.org/10.1007/s00429-019-01945-2
Safont, G., Salazar, A., Vergara, L., Gomez, E., Villanueva, V., 2017. Probabilistic distance for mixtures of independent component analyzers. IEEE Trans. Neural Netw. Learn. Syst. 1–13. doi:10.1109/TNNLS.2017.2663843.
Salazar, A., Vergara, L., Miralles, R., 2010. On including sequential dependence in ICA mixture models. Signal Process. 90 (7), 2314–2318. doi:10.1016/j.sigpro.2010.02.010.
Scherer, K.R., 2005. What are emotions? And how can they be measured? Soc. Sci. Inf. 44 (4), 695–729.
Seghier, M.L., 2013. The angular gyrus: multiple functions and multiple subdivisions. Neuroscientist 19 (1), 43–61. doi:10.1177/1073858412440596.
Sequeira, H., Hot, P., Silvert, L., Delplanque, S., 2009. Electrical autonomic correlates of emotion. Int. J. Psychophysiol. 71 (1), 50–56.
Shackman, A.J., McMenamin, B.W., Maxwell, J.S., Greischar, L.L., Davidson, R.J., 2009. Right dorsolateral prefrontal cortical activity and behavioral inhibition. Psychol. Sci. 20 (12), 1500–1506.
Sivagnanam, S., Majumdar, A., Yoshimoto, K., Astakhov, V., Bandrowski, A.E., Martone, M.E., Carnevale, N.T., 2013. Introducing the neuroscience gateway. IWSG 993.
Viviani, R., 2014. Neural correlates of emotion regulation in the ventral prefrontal cortex and the encoding of subjective value and economic utility. Front. Psychiatry 5 (September), 1–11. doi:10.3389/fpsyt.2014.00123.
Zheng, W.L., Zhu, J.Y., Peng, Y., Lu, B.L., 2014. Eeg-based emotion classification using deep belief networks. In: Proceedings of the IEEE International Conference on Multimedia and Expo (ICME). IEEE, pp. 1–6.
Zhuang, X., Rozgić, V., Crystal, M., 2014. Compact unsupervised eeg response representation for emotion recognition. In: Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). IEEE, pp. 736–739.