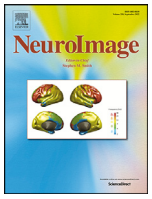
NeuroImage 258 (2022) 119374

Contents lists available at ScienceDirect

NeuroImage

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



# EEG decoding of multidimensional information from emotional faces

Yiwen Li [a,b,1], Mingming Zhang [a,b,1], Shuaicheng Liu [a,b], Wenbo Luo [a,b,*]

[a] Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian 116029, China
[b] Key Laboratory of Brain and Cognitive Neuroscience, Liaoning Province, Dalian 116029, China

ARTICLE INFO

*Keywords:*
Face recognition
Decoding
Emotion
Valence
Arousal

ABSTRACT

Humans can detect and recognize faces quickly, but there has been little research on the temporal dynamics of the different dimensional face information that is extracted. The present study aimed to investigate the time course of neural responses to the representation of different dimensional face information, such as age, gender, emotion, and identity. We used support vector machine decoding to obtain representational dissimilarity matrices of event-related potential responses to different faces for each subject over time. In addition, we performed representational similarity analysis with the model representational dissimilarity matrices that contained different dimensional face information. Three significant findings were observed. First, the extraction process of facial emotion occurred before that of facial identity and lasted for a long time, which was specific to the right frontal region. Second, arousal was preferentially extracted before valence during the processing of facial emotional information. Third, different dimensional face information exhibited representational stability during different periods. In conclusion, these findings reveal the precise temporal dynamics of multidimensional information processing in faces and provide powerful support for computational models on emotional face perception.

## 1. Introduction

The face is one of the most common stimuli in daily life and conveys multidimensional information that is essential for human survival and social communication (Little et al., 2011). For example, people can accurately extract an individual's age, gender, emotion, and identity within a few hundred milliseconds (Dima et al., 2018; Ambrus et al., 2019). However, we still do not know much about how this impressive ability occurs. The computational model of face perception assumes that different information from the face unfolds over time (Martinez, 2017). Therefore, it is fundamental to understand face processing by exploring which information is extracted and when it occurs.

Previous studies on the extraction of face information mainly performed univariate analysis of electroencephalogram (EEG) data and focused on specific event-related potentials (ERPs). For example, the classical N170 component is sensitive to facial emotion processing (Hinojosa et al., 2015). The N250 component has also been proposed as the marker of face identity recognition (Schweinberger et al., 2002; Schweinberger and Neumann, 2016). However, these components only involve limited channels, and channels that are not associated with such components have also been shown to facilitate the discrimination of different stimuli (Nemrodov et al., 2016). Therefore, univariate analysis of EEG data may not be well suited to describe face processing.

Multivariate pattern analysis (MVPA) of EEG data is more advantageous in addressing the temporal dynamics of information processing in the human brain (Grootswagers et al., 2017). In contrast to traditional univariate analysis techniques, MVPA considers the relationship among multiple variables (e.g., channels in EEG), which can capture the information that is not detectable in univariate analysis and improves the sensitivity of identifying differences among experimental conditions. For example, MVPA can detect signal differences during the early period (Cauchoix et al., 2012), and the results shown by MVPA may also vary from those found in the univariate analysis (Ritchie et al., 2015). In addition, emotions are represented in the brain in a complex and network-based manner (Hamann, 2012). Compared to univariate analysis, which explores the representation of emotions from local brain regions, MVPA can explore the representation of emotions from the whole brain.

MVPA studies have found that humans recognize faces from the background very early (approximately 100 ms; Cauchoix et al., 2014). There are also temporal variations in the processing of different types of face information (Dobs et al., 2019). Several MVPA studies have explored the temporal dynamics of facial identity and emotion separately and found that threatening emotional information of faces was processed 90 ms after stimulus appearance when performing a change detection task on faces (Dima et al., 2018). The identity information of

* Corresponding author: Wenbo Luo, Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, No. 850, Huanghe Road, Shahekou District, Dalian, 116029, China.
*E-mail address:* luowb@lnnu.edu.cn (W. Luo).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.neuroimage.2022.119374.
Received 7 January 2022; Received in revised form 3 June 2022; Accepted 10 June 2022
Available online 11 June 2022.

the faces was represented within 100 ms after the appearance of the face image and was modulated by gender differences and image similarities (Ambrus et al., 2019). However, these studies typically separate the processing of these two aspects and differ in their experimental tasks. Although some studies have found that the processing of emotion and identity occurs in different brain regions (O'Toole et al., 2002; Collins and Olson, 2014; Zhang et al., 2016), there are also interactions between these two aspects of information processing (Wegrzyn et al., 2015; Rivolta et al., 2016). Moreover, changes in the age (Folster et al., 2014) or gender (Ambrus et al., 2019) of faces and experimental tasks (Smith and Smith, 2019) will also affect the decoding of identity and emotion. Therefore, it is hardly possible for these studies to systematically reveal the processing of different face information.

For the processing of emotional information in faces, the circumplex model of affect (Russell, 1980) is widely accepted and suggests that all emotions could be represented by different valence and arousal levels. Valence and arousal are extracted separately from facial expressions during an early period (Calvo and Nummenmaa, 2016), and different neurophysiological systems process these two dimensions (Baucom et al., 2012). Relevant MVPA studies have also found that the processing of emotion follows the affective model of valence and arousal (Bush et al., 2018). Grootswagers et al. (2020) used representational similarity analysis (RSA) to compare different models of emotion and neural representations of stimuli in the brain and found that valence and arousal were better to explain neural representations, even after correcting for low-level image properties of the stimuli. However, these studies do not reveal the precise time when valence and arousal are represented.

In the present study, we aimed to integrate different dimensional face information, such as age, gender, emotion, and identity, and investigated the processing of these types of information in the same task. We performed multivariate EEG decoding and model-based RSA to systematically elucidate the processing of faces. In addition, we explored the relative time of valence and arousal represented during emotional face processing and would provide support for the computational models of emotional face perception.

## 2. Methods

### 2.1. Participants

As paid volunteers, 20 healthy college students participated in this study. One subject was excluded because of his/her poor sensitivity index *d'* (below 2.5 standard deviations of the mean). The other one was removed because of his/her excessive EEG artifacts, in which peak-to-peak deflection exceeding ± 80 $\mu$V accounted for more than 50% of the trials. Data from the remaining 18 subjects (11 females, mean age = 21.50 years, SD = 2.01 years) was used in all analyses throughout the study. All participants were right-handed and had normal or corrected to normal vision. They self-reported no diagnosed difficulties in recognizing faces (e.g., prosopagnosia) and other neurological or psychiatric problems. Informed consent was obtained prior to the experiment in accordance with procedures approved by the local ethics committee.

### 2.2. Stimuli

Materials consisted of 24 face images from the Tsinghua facial expression database (Yang et al., 2020). These images contained eight identities that varied orthogonally in age (half young, ≤ 24 years old; half old, ≥ 63 years old) and gender (half female, half male). Note that the gender here refers to the sex of the face images. For each identity, there were three different expressions (fearful, happy, and neutral). These original images were preprocessed with the following steps. First, each image was cropped into the shape of an ellipse using Adobe Photoshop CC 2015 software. Second, these images were converted into grayscale using the Shine toolbox (Willenbockel et al., 2010) in MATLAB 2018b (MathWorks, MA, USA) and were normalized to each other in terms of physical properties such as size, background, spatial frequency, contrast grade, and brightness.

An additional group of 22 paid healthy college students (15 female, mean age = 21.73 years, SD = 2.05 years) were recruited to rate the age, gender, valence, and arousal of each image we preprocessed using the mouse to scale on a slider with their right hands. Subjects were required to estimate age within a range of 10–80 years and to classify gender in a binary scheme (0 = male, 1 = female) when they did the age rating task and the gender discrimination task, respectively. The valence and arousal were both assessed using a 9-point rating scale (1 = negative or low arousal, 9 = positive or high arousal). These images used in this research differed significantly from each other in valence, $F(2, 63) = 134.36$, $p < 0.001$, $\eta_p^2 = 0.810$ ($M \pm SD$, happy: 7.45 ± 1.14; neutral: 4.19 ± 1.00; fearful: 2.30 ± 1.01) and also differed significantly in arousal, $F(2, 63) = 45.70$, $p < 0.001$, $\eta_p^2 = 0.592$; happy (6.74 ± 1.37, $p < 0.001$) and fearful (6.78 ± 1.56, $p < 0.001$) expressions had higher arousal than neutral ones (3.03 ± 1.55), while there was no significant difference between happy and fearful faces ($p = 1.000$). There was significant difference in age rating between young (26.90 ± 3.31) and old (59.98 ± 5.79) face images, $t(21) = 21.07$, $p < 0.001$. The significant difference in gender discrimination between male (0.01 ± 0.02) and female (0.99 ± 0.04) face images was also found, $t(21) = 110.63$, $p < 0.001$.

### 2.3. Procedure

During EEG recording, subjects viewed face images that appeared continuously on the screen (see Fig. 1). Stimuli were presented using the Psychtoolbox-3 (Brainard and Vision, 1997) in MATLAB 2018b (Mathworks, MA, USA). All face images were presented on a 30% gray background with a black fixation cross in the upper center of the screen (i.e., middle of the nose tip and the eyes). Subjects viewed the images from a seat 70 cm away from the screen, with a viewing angle of 3.84° × 2.54° for each image. In each trial, face image was first presented for 200 ms followed by a variable 800–1000 ms gray screen interstimulus interval (ISI). To maintain concentration, subjects were instructed to perform a 1-back task, during which they needed to press the space key with their right index finger as quickly and accurately as possible when a successive repetition of an identical image appeared. To avoid artifacts caused by blinking and eye movements, subjects were also instructed to fixate the black cross on the screen during the experiment and blink when making a button response, because these trials were not included in the EEG data analysis.

The formal experiment consisted of 10 blocks, each containing 240 trials. In each block, for the non-task trials, each of the 24 face images was presented eight times; for the task trials, each of the 24 face images was additionally and pseudo-randomly presented twice. These 48 task trials were used to ensure that subjects maintained their attention and were not included in the EEG data analysis. The subjects were allowed to rest between consecutive blocks, and the experiment lasted around 50 min. In total, there were 1920 trials included in the subsequent experimental analysis. To ensure that the subjects could better understand the experiment, one practice block of 120 trials was presented before the formal phase.

### 2.4. EEG recording and preprocessing

EEG data was continuously recorded by a 64-channel amplifier using a standard 10–20 system (Brain Products, Gilching, Germany). The EEG data was sampled at 1000 Hz and referenced online at FCz electrode, and all inter-electrode impedance was maintained below 5 kΩ. EEG data preprocessing was performed with EEGLAB (Delorme and Makeig, 2004) running in MATLAB 2018b (MathWorks, MA, USA). The removed channels were first interpolated and re-referenced offline to obtain a global average (Bentin et al., 1996), followed by bandpass-filtered between 0.1 and 30 Hz. Independent component analysis (Jung et al., 2000) was

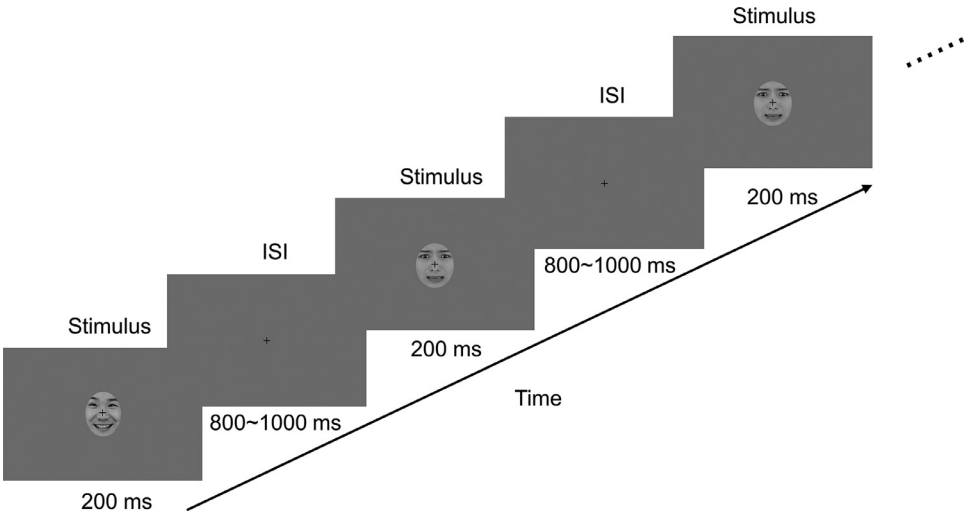



**Fig. 1.** Experimental paradigm. Subjects performed a 1-back task while viewing 24 face images. Each image was presented for 200 ms, followed by a variable interstimulus interval (ISI, 800–1000 ms).

used to correct ocular artifacts of the continuous data. Then epochs of −200–800 ms from each face image onset were created, and the same epoch length was used for all conditions. Rejection of the trials with excessive artifacts (i.e., peak-to-peak deflection exceeding ± 80 $\mu$V) was conducted, and 96.9% of the total trials were used for further data analysis. No other preprocessing steps were applied.

### 2.5. Data analyses

#### 2.5.1. Behavioral performance

The sensitivity index *d'* and mean response time (RT) in task trials were calculated according to signal detection theory. The extreme values that exceed ± 2.5 standard deviations of the mean RT were removed and accounted for 2.8% of the task trials. We then compared the RTs of age, gender, and emotion to test which kinds of faces were processed more efficiently than others. We used SPSS 25.0 to conduct a three-way repeated-measures analysis of variance (ANOVA) with age (young, old) × gender (female, male) × emotion (fearful, happy, neutral) for the RT of the hit trials in this experiment. Statistical effects that violated the sphericity test would be corrected for *p*-values using the Greenhouse-Geisser correction, and the Bonferroni correction was used for post-hoc comparisons. A further simple effect analysis would be performed when the significant interaction effect appeared.

#### 2.5.2. Decoding analysis

Multivariate time-series decoding was applied to each pair of face stimuli using support vector machine (SVM) classification, as well as functions from the CoSMoMVPA Toolbox (Oosterhof et al., 2016) and LIBSVM (Chang and Lin, 2011) in MATLAB 2018b (MathWorks, MA, USA). Classification was performed using the voltage values of all channels for each subject separately in a temporal-resolved manner (see Fig. 2A). In the following, we detailed one iteration of the multivariate classification procedure (see Fig. 2B): We first randomly assigned all trials of one condition to one of ten splits; there were approximately six to eight trials in each split when bad trials were removed (Isik et al., 2014). Supertrials were then created by averaging the trials in each split to improve the signal-to-noise ratio. In the second step, we divided these supertrials into training and testing data by randomly choosing one for testing and the others for training (i.e., ten-fold cross-validation). Last, binary classification was conducted on all pairwise comparisons between conditions (276 pairs in total; (24 - 1) × 24/2) to obtain decoding accuracies. This iteration procedure was repeated 100 times, and the average decoding accuracy between them resulted in a 24 × 24 decoding matrix, which is called the representational dissimilarity matrix (RDM). The RDM was symmetric, and its diagonal was not defined. The entire procedure was then repeated for each subject and time point, providing temporal-resolved estimates of EEG decoding accuracy (see Fig. 2C).

To measure the discrimination between each face image and all others in the EEG (i.e., image decoding), we averaged the lower off-diagonal of each RDM to evaluate the decoding accuracies at each time point. The time course of these decoding accuracies further served as a criterion for the time course of low-level image processing of the EEG data.

#### 2.5.3. Representational similarity analysis

We used RSA to investigate the representation of different face dimensions in the EEG data. Model RDMs were created for each face dimension using 24 × 24 binary matrices (see Fig. 3A). In the model RDMs of emotion and identity dimensions, 1 corresponded to a comparison between category stimuli (e.g., happy vs. fear for the emotion model), and 0 corresponded to a comparison within category stimuli (e.g., happy vs. happy). We also calculated the Euclidean distance between any two face images according to the ratings before the formal experiment, which resulted in four models of RDMs that corresponded to the age, gender, valence, and arousal dimensions of our stimuli. In addition, to exclude the contribution of the low-level features of the face images to the RSA results, a low-level feature model was built, in which the HMAX model described by Serre et al. (2005) was used. The features extracted from each of the 24 face images were computed in the C2 layer. We then calculated the Euclidean distance to measure the difference between the HMAX units of each pair of images, resulting in a 24 × 24 RDM (see Fig. 3B). Next, the lower off-diagonal of each matrix was extracted as vectors to calculate the Spearman rank correlations between each model and the EEG data. We calculated the partial correlation between the different model RDMs and the EEG RDMs at each time point for each subject after excluding other face models. Since some models were correlated (e.g., facial identity comparisons in age comparisons), excluding the other models allowed us to separate the contribution of these models from each other (Giordano et al., 2013; Cichy and Pantazis, 2017). For the arousal and valence dimensions, in order to explicitly compare these models to each other, arousal RDM would be excluded when computing a partial correlation between EEG RDM and valence RDM, and vice versa. These partial correlation coefficients served as an indicator of the time course of different face dimensions in the EEG data.

#### 2.5.4. Exploratory channel searchlight analysis

To explore the spatial source underlying the EEG decoding results (i.e., which channels drive decoding accuracy), we conducted a time-by-channel searchlight analysis. For each EEG channel, we selected the four closest neighboring channels to form a local cluster and then performed a decoding analysis of these channels (Oostenveld et al., 2011;

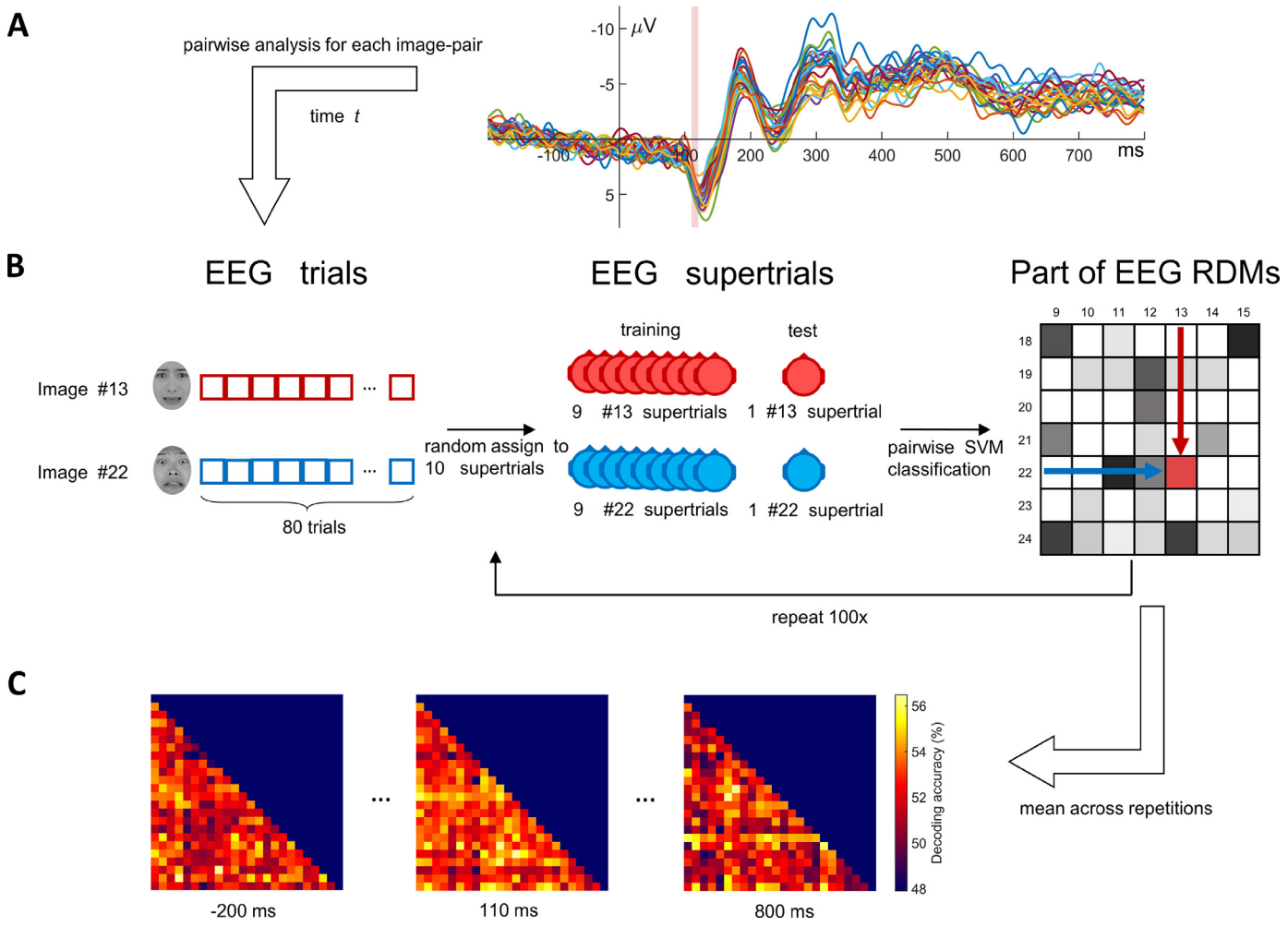



**Fig. 2.** Schematic for multivariate analyses of EEG data. A) Analyses were performed in a time-resolved manner from all EEG channels, separately for each subject. B) For each time point *t*, the responses of each pair of conditions were extracted to create supertrials, and pairwise cross-validated SVM classification was then performed to obtain the decoding accuracies. This process was repeated 100 times and finally the average decoding accuracy between conditions was placed in the corresponding positions in the 24 × 24 representational dissimilarity matrix (RDM). C) The decoding accuracies resulted in a 24 × 24 EEG RDM at each time point.

Oosterhof et al., 2016). The decoding accuracies were stored in the center channels of these clusters. This operation was repeated for each face dimension (image, age, gender, emotion, and identity) to generate a time-by-channel map of the decoding accuracies for each subject and dimension. We then calculated the multiclass decoding accuracies instead of the pairwise decoding accuracies to reduce calculation time. For better visualization, the time-by-channel searchlight maps were averaged every 20 ms for each face dimension.

### 2.5.5. Temporal generalization approach

To investigate whether different face dimension representations remained stable or changed over time, we extended the temporal generalization approach of MVPA (King and Dehaene, 2014). In this approach, a classifier is trained on data at a specific time point (i.e., training time) and tested at all other time points (i.e., generalization time). If the representation remains stable across time, the classifier should be able to distinguish two conditions not only at the trained time but also at later time points. If this distinction does not happen, then this indicates a change in the representation format. Like the previous classification analysis at identical time points, we repeated the same procedure described previously; however, instead of performing tests on the classifier only at a specific time point, we tested the identical classifier separately for all other time points. We performed a temporal generalization analysis for each subject and face dimension and then averaged the decoding accuracies of each subject. This generated five 1000 × 1000 matrices (−200–799 ms to stimulus onset) that captured the classifier generalization performance across time.

### 2.5.6. Statistical inference

We performed a non-parametric statistical approach for all EEG data analyses which did not depend on assumptions of the data distributions (Nichols and Holmes, 2002). We carried out permutation-based cluster-size inference to test the periods during which the decoding accuracy or correlation showed significant effects. The null hypothesis corresponds to 50% of decoding accuracy and a correlation coefficient 0. Significant temporal clusters were specified as adjacent time points that exceeded a statistical cut-off (i.e., cluster-inducing threshold). The conditional labels of the EEG data were permuted by random multiplication by +1 or −1 (i.e., sign permutation test), and this procedure was repeated 1000 times to acquire the *t*-value distribution of all possible permutations. The cutoff was determined through this sign permutation test according to these *t*-value distributions of all possible permutations. At each time point, the 95th percentile of the *t*-value distribution was used as the clustering induction threshold (equivalent to $p < 0.05$, one-sided). The 95th percentile of the maximum cluster size among all permutations served as the significant cluster (equivalent to $p < 0.05$, one-sided). This approach provided us with significant temporal clusters in which decoding accuracy or correlation showed significant effects.

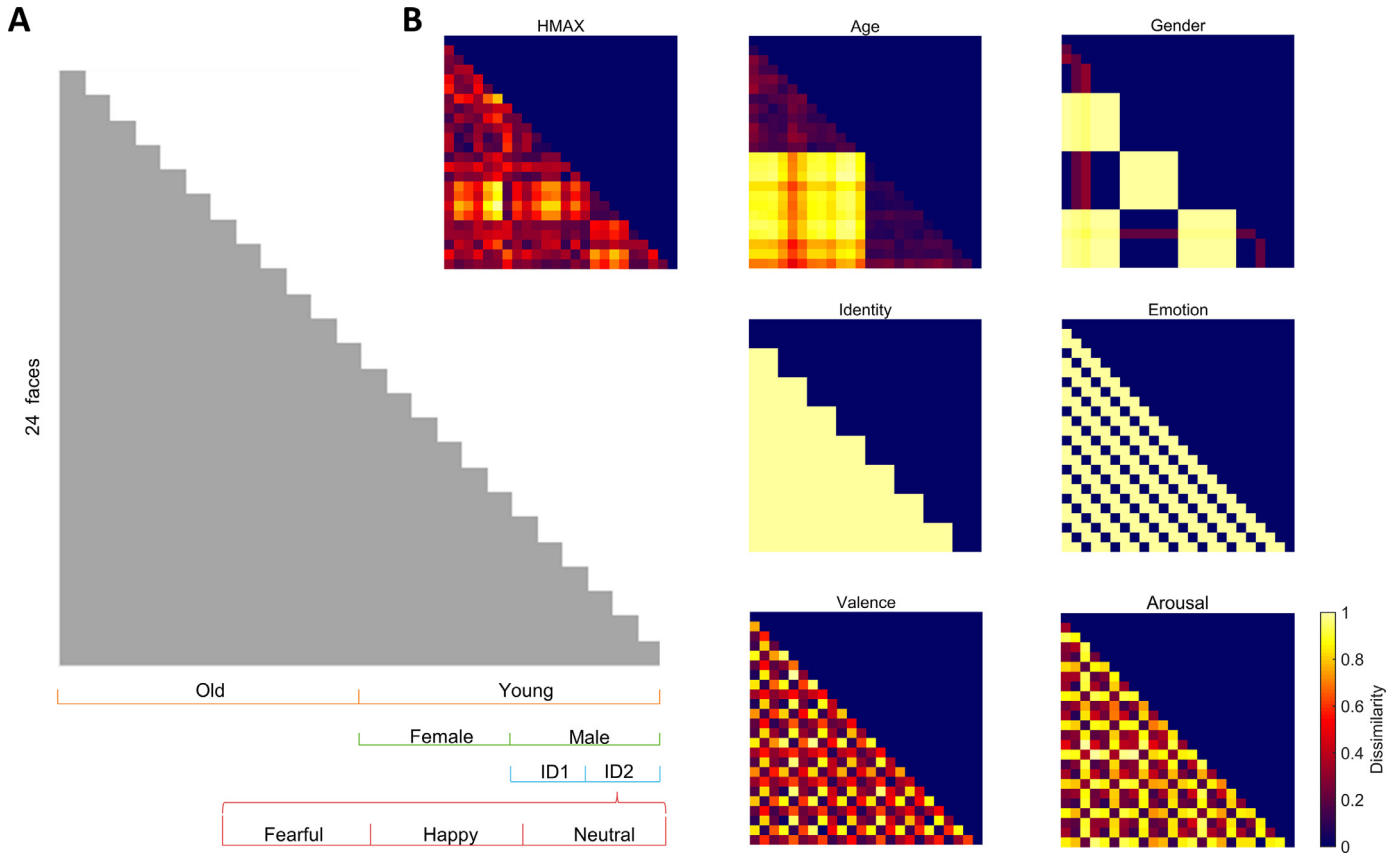



**Fig. 3.** Schematic for model representational dissimilarity matrix (RDM). A) To perform RSA, we constructed model RDMs for each face dimension (1 corresponding to between-category, 0 corresponding to within-category, respectively). B) Model RDMs for different dimensional face information.

For searchlight analysis, we first revealed the significant timepoints for each channel and then marked the significant channels on the topographic map at each time point. Note that channels would be marked only if they were significant at all time points within each time window of 20 ms. For temporal generalization matrices, we revealed significant clusters by extending the cluster-based approach to two-dimension. We only ran a subset of 1000 permutations selected randomly instead of replacement from all possible permutations, owing to computational limitations.

#### 2.5.7. Determining confidence intervals for latencies

Bootstrap tests were used to estimate the statistical differences and 95% confidence intervals (CI) of the onset, peak, and duration latencies between different face dimensions. The time courses (e.g., decoding accuracy or correlation coefficient) were bootstrapped 1000 times, resulting in an empirical distribution of the onset (i.e., the minimum significant time point after stimulus onset), peak (i.e., the time point at the maximum value between 120 and 220 ms after stimulus onset), and duration latencies (i.e., the maximum significant time point after stimulus onset). Since the analysis of peak latency was performed for the first peak that appeared after stimulus onset, the time window was restricted to 220 ms after stimulus onset to avoid the confounding of others owing to stimulus offset responses (Carlson et al., 2011). We defined the 95% confidence interval for these latencies as the 2.5th and 97.5th percentiles of these distributions. The differences between the latencies of any two dimensions in these 1000 bootstrap samples were then computed, causing an empirical distribution of latency differences. The *p*-value was defined as the number of samples whose differences were bigger or smaller than zero divided by the number of permutations (i.e., two-sided test). These *p*-values were corrected for multiple comparisons with a false discovery rate (FDR) of 0.05.

## 3. Results

### 3.1. Behavioral performance

Subjects were highly sensitive to the repetition of images (mean sensitivity index *d'* ± *SEM*: 3.35 ± 0.12) and responded quickly (mean RT ± *SEM*: 535 ± 17 ms after target stimulus onset). Repeated-measures ANOVAs revealed that RT was significantly affected by age ($F(1, 17) = 9.03$, $p = 0.008$, $\eta_p^2 = 0.347$) and gender ($F(1, 17) = 9.51$, $p = 0.007$, $\eta_p^2 = 0.359$). Subjects performed a shorter RT in the recognition of old faces (*M* ± *SE*, 530.7 ± 17.0 ms) than young faces (539.5 ± 17.0 ms) and had a shorter RT when identifying female faces (*M* ± *SE*, 531.2 ± 16.6 ms) than male faces (539.1 ± 17.3 ms). The age × gender × emotion interaction ($F(2, 34) = 6.79$, $p = 0.007$, $\eta_p^2 = 0.459$) indicated that RT was significantly shorter in the recognition of females (522.0 ± 16.7 ms, $p = 0.021$) than males (533.8 ± 16.8 ms) when old fearful faces were presented, with non-significant differences in the neutral and happy conditions (*ps* ≥ 0.511); while RT was shorter in the recognition of female (533.7 ± 19.4 ms, $p < 0.001$) compared to the male (550.3 ± 18.5 ms) when young neutral faces were presented, with non-significant differences in the fear and happy conditions (*ps* ≥ 0.126, see Fig. 4). These behavioral results indicate that old and female faces could be enhanced for processing compared to other faces, even though the task requires only image-level processing, not identity-level processing.

### 3.2. Time-resolved multivariate face image decoding

The time course of neural image decoding accuracy demonstrated that face information increased rapidly after the stimulus presentation, and the decoding accuracy reached significance at 79 ms ($t(17) = 1.785$,

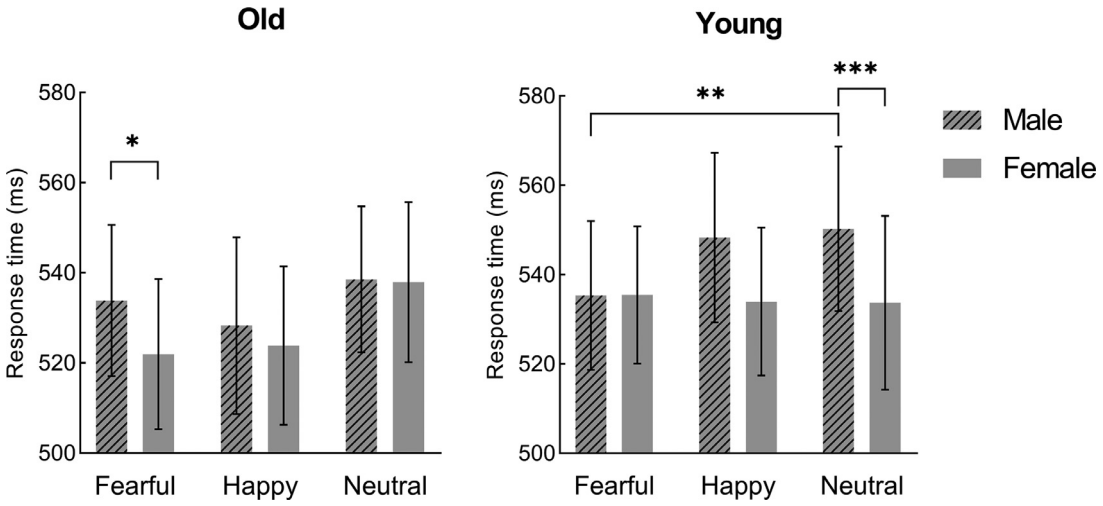



**Fig. 4.** Mean response time in behavioral performance. Error bars indicate *SEM*. $^{*}p < 0.05$, $^{**}p < 0.01^{***}$, $p < 0.001$.

Cohen's $d = 0.866$, see Fig. 5A, cluster-corrected sign permutation test, cluster definition threshold $p < 0.05$, cluster-corrected significance level $p < 0.05$). The previous studies with EEG (Cauchoix et al., 2014; Dziura and Thompson, 2020) and magnetoencephalography (MEG) decoding studies (Dima et al., 2018) as well as intracranial recording studies (Barbeau et al., 2008) have also found that the decoding of face information begins approximately in the first 100 ms. The image decoding accuracy peaked for the first time at 149 ms, with an average decoding accuracy of 54.5%. The information then remained significantly above the chance level and slowly decreased during the duration of the face stimulus disappearance and the stimulus interval (800 ms after stimulus onset). The results show how EEG responses are resolved at the level of individual face images and provide an important foundation for the following RSA, in which we will explore which dimensions of information in faces particularly contribute to the representation of these recognizable faces.

### 3.3. Representations of face dimensions revealed by RSA

The RSA results revealed the time course of the representations of age (see Fig. 5B, significant time points: 113–615 ms), emotion (114–559 ms and 566–799 ms), and identity (235–554 ms, all cluster-corrected sign permutation test, cluster definition threshold $p < 0.05$, cluster-corrected significance level $p < 0.05$). Age was first extracted from the EEG signal representation, followed by the extraction of emotion (see Fig. 5D, $p = 0.032$, two-sided bootstrap test, FDR corrected). Identity was finally extracted, approximately 120 ms later than emotion ($p < 0.001$, two-sided bootstrap test, FDR corrected). The processing of facial emotion began at about 120 ms and continued until about 800 ms after stimulus onset, suggesting that the processing of facial emotion might be a deliberate and persistent process in the human brain. Most dimensions of face information peaked around 170 ms (see Fig. 5E), which also coincides with the time point when the N170 component evoked by face stimuli appears in previous studies (Nemrodov et al., 2018). The duration of identity was significantly shorter than other dimensions (see Fig. 5F, *ps* $< 0.006$, two-sided bootstrap test, FDR corrected), which indicates that people could complete the extraction and processing of identity information within a short time after seeing faces.

In addition, we performed a similar cross-decoding analysis on the image dimensions, such as training age on some images and testing age on the rest of others (Cichy et al., 2014, 2017). Highly similar results were obtained for all dimensions (see Supplementary Material).

Emotion is critical throughout the face processing and can also be conceptualized as valence and arousal. The results of the RSA for these two dimensions showed the time course of representation for valence (see Fig. 5C, significant time point: 210–347 ms) and arousal (127–486 ms, and 669–729 ms; all cluster-corrected sign permutation test, cluster definition threshold $p < 0.05$, cluster-corrected significance level $p < 0.05$). The representation of arousal occurred significantly earlier than valence (see Fig. 5G, $p < 0.001$), peaked almost simultaneously with valence (see Fig. 5H, $p = 0.105$), and lasted significantly longer than valence (see Fig. 5I, $p = 0.001$, two-sided bootstrap test, FDR corrected). These results suggest that in the bottom-up processing of emotional faces, arousal might be extracted first and lasts for a long time, while valence would be processed next and has a short duration.

### 3.4. Searchlight analysis

The results illustrated the searchlight maps of five face dimensions from a 110–210 ms time window, and the time-by-channel searchlight map was averaged each 20 ms for the visualization (see Fig. 6). After 130 ms, all face dimensions except gender reached significance in most channels of the occipital region. Subsequently, some channels located in the temporal and parietal regions also reached significance. Interestingly, in the time window after 130 ms, the signal of emotion not only widely appeared in the right temporoparietal region, but also reached significance in most channels of the right frontal region. This result suggests that the right frontal region might be specific to the processing of emotional information in faces and that the processing of emotional information could be widely distributed in the brain.

### 3.5. Temporal generalization analysis

We found that the temporal generalization of the age dimension began at around 100–150 ms and lasted until about 650 ms, while subsequent stages (beginning at 350–450 ms) generalized about 400 ms away from their training time (see Fig. 7B). The temporal generalization of the gender dimension did not reach significance (see Fig. 7C). The temporal generalization of the emotion dimension appeared at about 150–250 ms, lasted until about 500 ms, and subsequent stages generalized about 400 ms away from their training time (see Fig. 7D). The temporal generalization of the identity dimension occurred at about 100–200 ms and generalized about 500 ms away from their training time (see Fig. 7E). For the images with more fine details, the neural representations were transient and had a shorter generalization time (see Fig. 7F). These results showed that different dimensional face information might be represented stable at different stages respectively. Interestingly, we found a lack of generalization between 200 and 230 ms in emotion, shown as non-significant off-diagonal areas between 200 and 230 ms; these neural representations were reactivated after 230 ms (see Fig. 7D).

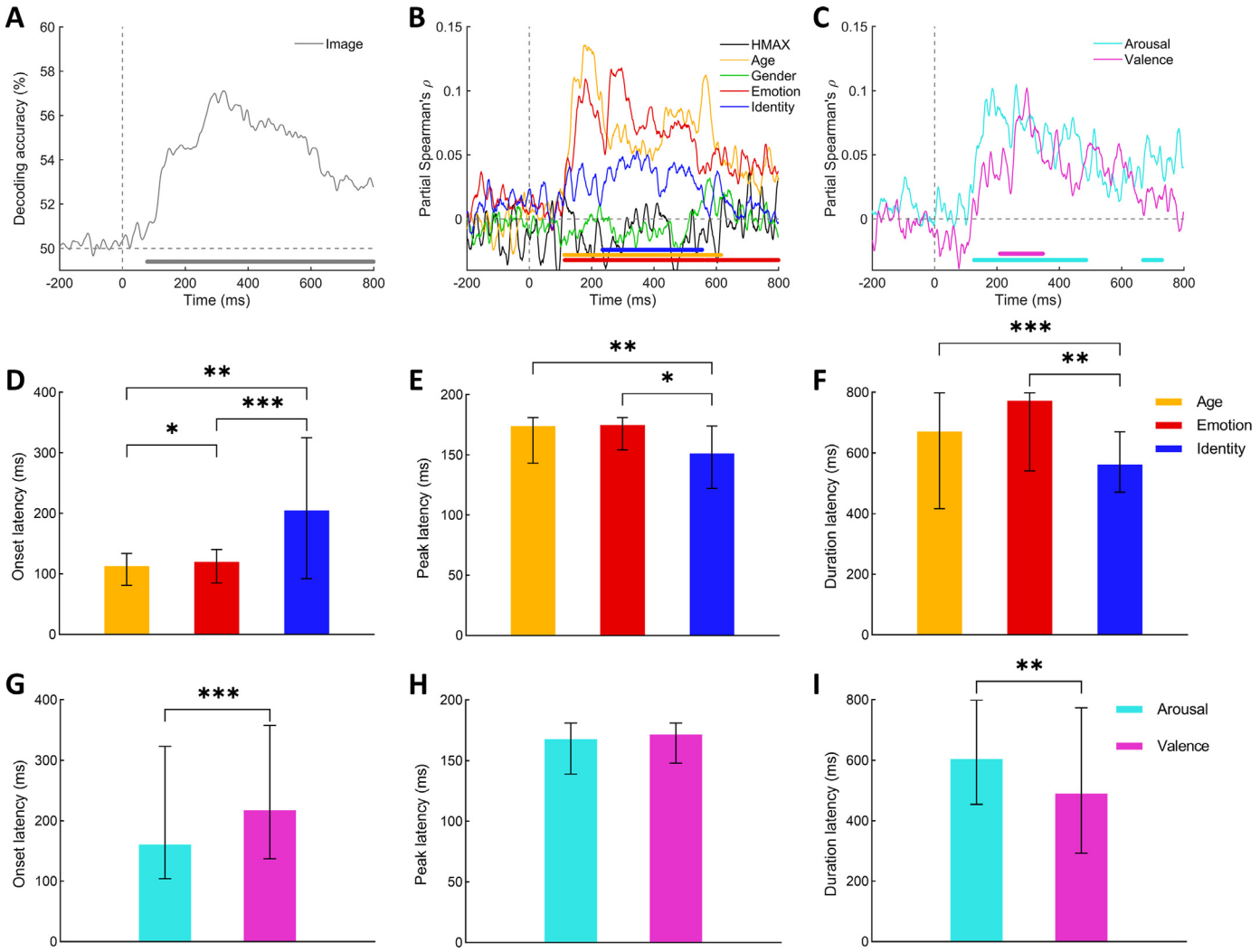



**Fig. 5.** Time-resolved EEG decoding of face images and other dimensions. A) Time course of image decoding where 0 indicates image onset. B) Time course of partial Spearman correlations between EEG RDMs and model RDMs for HMAX (black), age (orange), gender (green), emotion (red), and identity (blue). C) Time course of partial Spearman correlations for valence (purple) and arousal (cyan). Bold lines above the horizontal axis indicate significant time points using the cluster-corrected sign permutation test (cluster definition threshold $p < 0.05$, cluster-corrected significance level $p < 0.05$). D-F) Onset (D), peak (E), and duration (F) latencies for decoding age, emotion, and identity. G-I) Onset (G), peak (H), and duration (I) latencies for decoding valence and arousal. Error bars indicate bootstrapped 95% confidence intervals. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-sided bootstrap test, FDR-corrected). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Overall, we found that there were different periods of generalization for different facial neural representations, corresponding to the periods in the previous RSA results.

## 4. Discussion

This study used multivariate EEG decoding and model-based RSA to elucidate the neurodynamic pattern of multidimensional information processing in faces. First, we found that the extraction process of emotion in the face occurred before that of identity, had a longer duration, and was specific to the right frontal region. Second, arousal was extracted and processed first, and the duration was long during the processing of facial emotional information, whereas valence was extracted later and had a short duration. Third, different dimensional face information had representational stability during different periods. These findings reveal the temporal dynamics of emotional face processing and provide robust support for computational models of emotional face perception.

The present study showed that the age dimension of faces was processed first, followed by the detailed distinction of identity, indicating that face processing might follow a temporal trajectory from coarse to fine. This is consistent with Dobs et al. (2019), who found that individual characteristics, such as gender and age, were decoded early in the neural response to faces, while identity decoding was typically discovered later and appeared after 100 ms (Vida et al., 2017; Ambrus et al., 2019). Similarly, the neuronal activity of the temporal cortex in macaques responded to the integral information of the face in the initial stage, and then the processing of more refined information appeared, such as identity (Sugase et al., 1999; Cauchoix et al., 2012). However, these studies on the processing of faces in the human brain typically ignore the role of emotion. Previous most univariate studies have not observed a moderation of emotion in the P1 component (100 ms; Schupp et al., 2004; Frühholz et al., 2011), whereas the moderation of emotion occurred in the N170 component (170 ms; Hinojosa et al., 2015), which suggests that the decoding of emotions may occur between 100 and 170 ms. A multivariate study by Muukkonen et al. (2020) also found that expres-

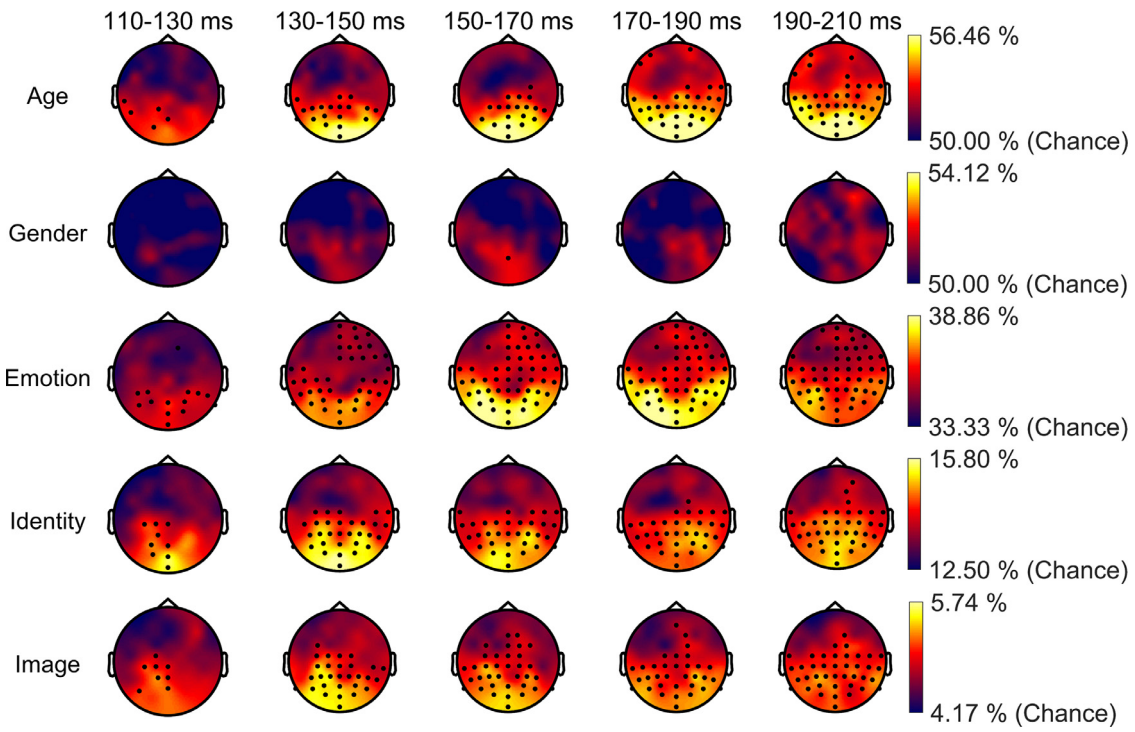



**Fig. 6.** Exploratory channel searchlight analysis. Plots show decoding accuracies as a topographic distribution. For visualization, time-by-channel searchlight maps were averaged across time every 20 ms. Rows show all five face dimensions, columns show five time windows, and the color scale indicates the decoding accuracy. The black dots in the topographic maps represent channels reaching significance ($p < 0.05$, cluster-corrected sign permutation test).

sions were decoded significantly after 120 ms, but none of these studies made inferences about the precise process of facial emotional representation. We incorporated emotion into the process of face recognition and found that the extraction process of emotion took place before the extraction of identity, beginning at about 120 ms and lasting a long time.

Interestingly, the searchlight analysis in the present study found that the right frontal region was specific to emotional processing. The right inferior frontal cortex processes visual or auditory emotional communication signals, and there is a hemispheric asymmetry in the inferior frontal cortex associated with emotional communication processing (Nakamura et al., 1999). The right medial prefrontal cortex is also involved in individual ratings of facial emotional arousal (Gerber et al., 2008). In addition, some injury studies have also found that patients with right hemisphere damage were significantly more impaired in perceiving facial emotion than those with left hemisphere damage (Mandal et al., 1999). Lesions in the right frontal region also interfere with valence recognition (Peper and Irle, 1997). Note that although the emotional effect was observed in the channels located in the right frontal region, the present study did not pinpoint the source of the signal due to volume conduction. In the future, we will explore the processing locations of different dimensional face information in the human brain by combining EEG and functional magnetic resonance imaging (fMRI). Overall, these findings are basically consistent with the prior literature and attempt to interpret the detailed properties of the human faces using EEG with high temporal resolution.

Arousal was preferentially extracted prior to valence during the processing of facial emotional information. Previous research on emotion and attention has shown that it is difficult to ignore the emotional information presented in sensory stimuli. For example, expressions that are independent of the task, can affect the matching performance of face identity (Van den Stock and de Gelder, 2014). In the current experimental design, it was necessary to discuss the role of emotion in face processing, and there were two general theories of emotion processing. Basic emotion theory holds that each emotion category is supported by a separate neural system. However, EEG research by Muukkonen et al. (2020) showed that the classification was not based on the emotion category or intensity. Instead, an increasing number of fMRI studies support the circumplex model of affect (Yuen et al., 2012; Bush et al., 2018), which states that different neurophysiological systems represent valence and arousal separately. Based on the circumplex model of affect, the present study revealed how valence and arousal are represented in the human brain, supporting the concept that different brain regions represent valence and arousal. During the processing of emotional information, arousal was extracted first and the duration was long, whereas valence was extracted later and the duration was short. Arousal representation rose faster in the about 200 ms, despite not representing the actual peak of the time course. Previous univariate studies have found that threatening face stimuli are processed preferentially (Schupp et al., 2004), which may be attributable to the negative valence involved. However, Balconi and Pozzoli (2012) provide a more reasonable explanation that threatening facial expressions require more attentional resources (Williams et al., 2005; Balconi and Pozzoli, 2008), which affects the level of arousal and is subsequently reflected in the corresponding ERP components. In conclusion, the present study suggests that arousal is preferentially extracted, providing an alternative explanation for the processing advantage of threatening faces.

The present study also found that different dimensions of the face have representational stability at different periods. This result illustrates the temporal variability in the representations of different facial dimensions in the human brain, rather than being stable over time (Wardle et al., 2020). In contrast, the neural representations of the image dimension were primarily transient and not persistent over time, suggesting a continuous transformation of low-level stimulus information. This conclusion corresponds to a lack of temporal generalization in the decoding of cross-temporal targets based on MEG data (Carlson et al., 2013; Isik et al., 2014; Dobs et al., 2019). Interestingly, there was a lack of generalization between 200 and 230 ms for emotion, with neural representations being reactivated after 230 ms, which may reflect the shift

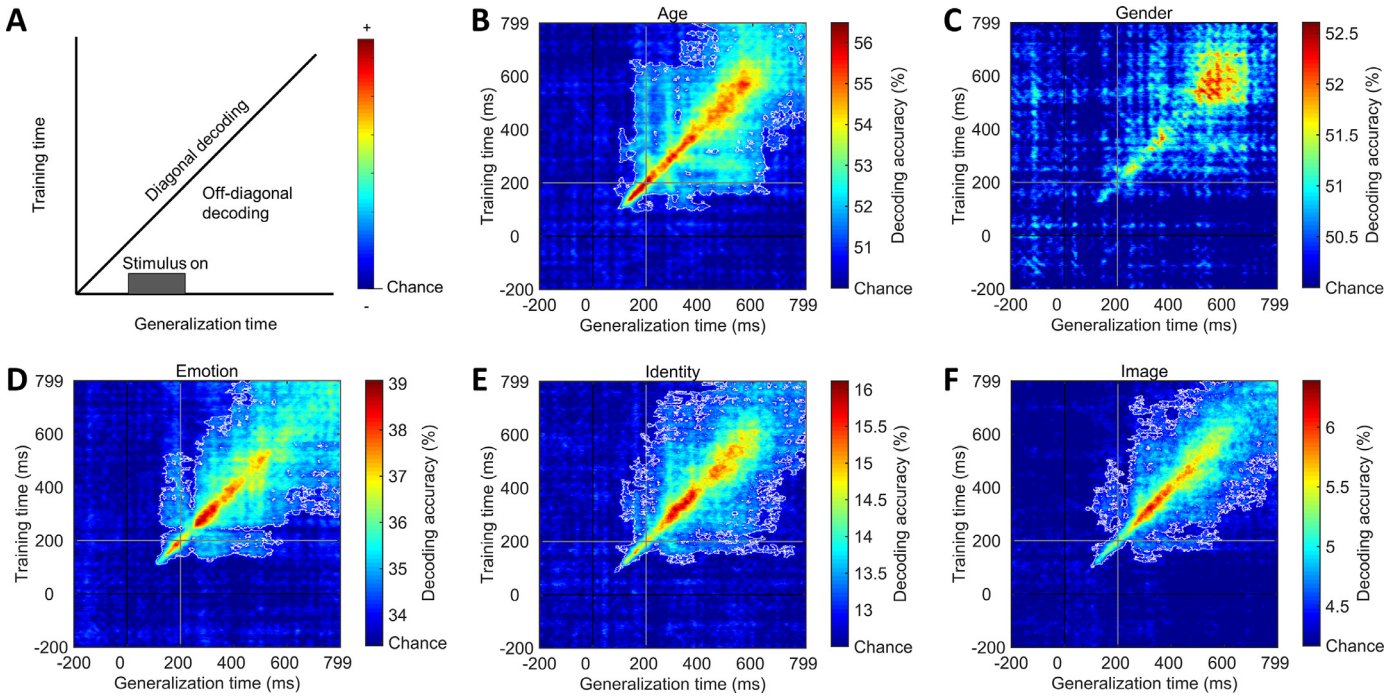



**Fig. 7.** Temporal generalization analysis. A) Schematic for temporal generalization analysis. B-F) Temporal cross-classification matrix for different face dimensions. The y-axis represents the classifier training time relative to stimulus onset, the x-axis represents the classifier generalization time, and the color scale indicates the cross-classification accuracy for each combination of training and generalization time. The white outline indicates significant clusters ($p < 0.05$, cluster-corrected sign permutation test).

from arousal to the valence of emotional processing. Overall, we found different periods of temporal generalization for different facial neural representations.

The results of this study further deepen our understanding of temporal dynamics in facial expression processing, but there are also several limitations. For example, to avoid low-level image confusion, the face images used in this study were cropped to retain only the core face region, which lost their ecological validity compared with natural images. The ears and hair of these facial images were cropped, resulting in the loss of gender information (Ergen and Abut, 2013). This may also explain why the gender dimension did not reach significance in this research. Furthermore, subjects performed a 1-back task to maintain attention in the present experiment. Although we have only selected non-task trials in the decoding analysis, it is still possible to introduce ERP modulation due to ongoing working memory (Shalchy et al., 2020). Therefore, the ERP for face recognition might be confounded with those working memory. In the future, some other experimental tasks could be used to eliminate the task-related effect on the current findings.

## 5. Conclusions

In summary, our study of multivariate EEG decoding demonstrates the precise temporal dynamics of multidimensional information processing in faces and highlights their stability of neural representations. Our results reveal that the representation of emotions in the human brain is according to the circumplex model of affect and that arousal is represented first, followed by valence. The present study provides computational models of emotional face perception with more clear and powerful support.

## Ethics statement

This work was approved by the Ethics Committee of Liaoning Normal University in accordance with the Declaration of Helsinki (1991). Informed consent was obtained from all subjects involved in the study.

## Data availability statement

The available data can be found at https://osf.io/pcvt5/.

## Declaration of Competing Interest

No potential conflict of interest was reported by the authors.

## Credit authorship contribution statement

**Yiwen Li:** Conceptualization, Writing – original draft, Writing – review & editing, Methodology, Investigation, Formal analysis, Visualization. **Mingming Zhang:** Writing – original draft, Writing – review & editing, Visualization, Supervision. **Shuaicheng Liu:** Investigation, Software. **Wenbo Luo:** Conceptualization, Supervision, Project administration, Funding acquisition.

## Acknowledgments

All authors approved the final version of the manuscript for submission. This work was supported by the National Natural Science Foundation of China (Grant Nos. 31871106).

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119374.

## References

Ambrus, G.G., Kaiser, D., Cichy, R.M., Kovacs, G., 2019. The neural dynamics of familiar face recognition. Cereb. Cortex 29, 4775–4784. doi:10.1093/cercor/bhz010.

Balconi, M., Pozzoli, U., 2008. Event-related oscillations (ERO) and event-related potentials (ERP) in emotional face recognition: A regression analysis. Int. J. Neurosci. 118, 1412–1424. doi:10.1080/00207450601047119.

Balconi, M., Pozzoli, U., 2012. Encoding of emotional facial expressions in direct and incidental tasks: An event-related potentials N200 effect. J. Neurother. 16, 92–109. doi:10.1080/10874208.2012.677659.

Barbeau, E.J., Taylor, M.J., Regis, J., Marquis, P., Chauvel, P., Liegeois-Chauvel, C., 2008. Spatio temporal dynamics of face recognition. Cereb. Cortex 18, 997–1009. doi:10.1093/cercor/bhm140.

Baucom, L.B., Wedell, D.H., Wang, J., Blitzer, D.N., Shinkareva, S.V., 2012. Decoding the neural representation of affective states. Neuroimage 59, 718–727. doi:10.1016/j.neuroimage.2011.07.037.

Bentin, S., Allison, T., Puce, A., Perez, E., McCarthy, G., 1996. Electrophysiological studies of face perception in humans. J. Cogn. Neurosci. 8, 551–565. doi:10.1162/jocn.1996.8.6.551.

Brainard, D.H., Vision, S., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436. doi:10.1163/156856897X00357.

Bush, K.A., Privratsky, A., Gardner, J., Zielinski, M.J., Kilts, C.D., 2018. Common functional brain states encode both perceived emotion and the psychophysiological response to affective stimuli. Sci. Rep. 8, 1–10. doi:10.1038/s41598-018-33621-6.

Calvo, M.G., Nummenmaa, L., 2016. Perceptual and affective mechanisms in facial expression recognition: An integrative review. Cogn Emot 30, 1081–1106. doi:10.1080/02699931.2015.1049124.

Carlson, T., Tovar, D.A., Alink, A., Kriegeskorte, N., 2013. Representational dynamics of object vision: The first 1000 ms. J. Vis. 13, 1–19. doi:10.1167/13.10.1.

Carlson, T.A., Hogendoorn, H., Kanai, R., Mesik, J., Turret, J., 2011. High temporal resolution decoding of object position and category. J. Vis. 11, 1–17. doi:10.1167/11.10.9.

Cauchoix, M., Arslan, A.B., Fize, D., Serre, T., 2012. The neural dynamics of visual processing in monkey extrastriate cortex: A comparison between univariate and multivariate techniques. In: Langslrina, G., Grosse-Wentrup, R., Murphy, B. (Eds.), Machine Learning and Interpretation in Neuroimaging. Springer, Sierra Nevada, Spain, pp. 164–171.

Cauchoix, M., Barragan-Jason, G., Serre, T., Barbeau, E.J., 2014. The neural dynamics of face detection in the wild revealed by MVPA. J. Neurosci. 34, 846–854. doi:10.1523/JNEUROSCI.3030-13.2014.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 1–27. doi:10.1145/1961189.1961199.

Cichy, R.M., Khosla, A., Pantazis, D., Oliva, A., 2017. Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. Neuroimage 153, 346–358. doi:10.1016/j.neuroimage.2016.03.063.

Cichy, R.M., Pantazis, D., 2017. Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. Neuroimage 158, 441–454. doi:10.1016/j.neuroimage.2017.07.023.

Cichy, R.M., Pantazis, D., Oliva, A., 2014. Resolving human object recognition in space and time. Nat. Neurosci. 17, 455–462. doi:10.1038/nn.3635.

Collins, J.A., Olson, I.R., 2014. Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. Neuropsychologia 61, 65–79. doi:10.1016/j.neuropsychologia.2014.06.005.

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, 9–21. doi:10.1016/j.jneumeth.2003.10.009.

Dima, D.C., Perry, G., Messaritaki, E., Zhang, J., Singh, K.D., 2018. Spatiotemporal dynamics in human visual cortex rapidly encode the emotional content of faces. Hum. Brain Mapp. 39, 3993–4006. doi:10.1002/hbm.24226.

Dobs, K., Isik, L., Pantazis, D., Kanwisher, N., 2019. How face perception unfolds over time. Nat. Commun. 10, 1258. doi:10.1038/s41467-019-09239-1.

Dziura, S.L., Thompson, J.C., 2020. Temporal dynamics of the neural representation of social relationships. J. Neurosci. 40, 9078–9087. doi:10.1523/JNEUROSCI.2818-19.2020.

Ergen, B., Abut, S., 2013. Gender recognition using facial images. Paper presented at the International Conference on Agriculture and Biotechnology.

Folster, M., Hess, U., Werheid, K., 2014. Facial age affects emotional expression decoding. Front Psychol. 5, 30. doi:10.3389/fpsyg.2014.00030.

Frühholz, S., Jellinghaus, A., Herrmann, M., 2011. Time course of implicit processing and explicit processing of emotional faces and emotional words. Biol. Psychol. 87, 265–274. doi:10.1016/j.biopsycho.2011.03.008.

Gerber, A.J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., Peterson, B.S., 2008. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 46, 2129–2139. doi:10.1016/j.neuropsychologia.2008.02.032.

Giordano, B.L., McAdams, S., Zatorre, R.J., Kriegeskorte, N., Belin, P., 2013. Abstract encoding of auditory objects in cortical activity patterns. Cereb. Cortex 23, 2025–2037. doi:10.1093/cercor/bhs162.

Grootswagers, T., Kennedy, B.L., Most, S.B., Carlson, T.A., 2020. Neural signatures of dynamic emotion constructs in the human brain. Neuropsychologia 145, 106535. doi:10.1016/j.neuropsychologia.2017.10.016.

Grootswagers, T., Wardle, S.G., Carlson, T.A., 2017. Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. J. Cogn. Neurosci. 29, 677–697. doi:10.1162/jocn_a_01068.

Hamann, S., 2012. Mapping discrete and dimensional emotions onto the brain: Controversies and consensus. Trends Cogn. Sci. 16, 458–466. doi:10.1016/j.tics.2012.07.006.

Hinojosa, J., Mercado, F., Carretié, L., 2015. N170 sensitivity to facial expression: A meta-analysis. Neurosci. Biobehav. Rev. 55, 498–509. doi:10.1016/j.neubiorev.2015.06.002.

Isik, L., Meyers, E.M., Leibo, J.Z., Poggio, T., 2014. The dynamics of invariant object recognition in the human visual system. J. Neurophysiol. 111, 91–102. doi:10.1152/jn.00394.2013.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758. doi:10.1016/S1388-2457(00)00386-2.

King, J.-R., Dehaene, S., 2014. Characterizing the dynamics of mental representations: The temporal generalization method. Trends Cogn. Sci. 18, 203–210. doi:10.1016/j.tics.2014.01.002.

Little, A.C., Jones, B.C., DeBruine, L.M., 2011. The many faces of research on face perception. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366, 1634–1637. doi:10.1098/rstb.2010.0386.

Mandal, M.K., Borod, J.C., Asthana, H.S., Mohanty, A., Mohanty, S., Koff, E., 1999. Effects of lesion variables and emotion type on the perception of facial emotion. J. Nerv. Ment. Dis. 187, 603–609. doi:10.1097/00005053-199910000-00003.

Martinez, A.M., 2017. Computational models of face perception. Curr. Dir. Psychol. Sci. 26, 263–269. doi:10.1177/0963721417698535.

Muukkonen, I., Olander, K., Numminen, J., Salmela, V.R., 2020. Spatio-temporal dynamics of face perception. Neuroimage 209, 116531. doi:10.1016/j.neuroimage.2020.116531.

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Kojima, S., 1999. Activation of the right inferior frontal cortex during assessment of facial emotion. J. Neurophysiol. 82, 1610–1614. doi:10.1152/jn.1999.82.3.1610.

Nemrodov, D., Niemeier, M., Mok, J.N.Y., Nestor, A., 2016. The time course of individual face recognition: A pattern analysis of ERP signals. Neuroimage 132, 469–476. doi:10.1016/j.neuroimage.2016.03.006.

Nemrodov, D., Niemeier, M., Patel, A., Nestor, A., 2018. The neural dynamics of facial identity processing: Insights from EEG-based pattern analysis and image reconstruction. eNeuro 5, 1–17. doi:10.1523/ENEURO.0358-17.2018.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum. Brain Mapp 15, 1–25. doi:10.1002/hbm.1058.

O'Toole, A.J., Roark, D.A., Abdi, H., 2002. Recognizing moving faces: A psychological and neural synthesis. Trends Cogn. Sci. 6, 261–266. doi:10.1016/S1364-6613(02)01908-3.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869. doi:10.1155/2011/156869.

Oosterhof, N.N., Connolly, A.C., Haxby, J.V., 2016. CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. Front Neuroinform. 10, 27. doi:10.3389/fninf.2016.00027.

Peper, M., Irle, E., 1997. The decoding of emotional concepts in patients with focal cerebral lesions. Brain Cogn. 34, 360–387. doi:10.1006/brcg.1997.0913.

Ritchie, J.B., Tovar, D.A., Carlson, T.A., 2015. Emerging object representations in the visual system predict reaction times for categorization. PLoS Comput. Biol. 11, e1004316. doi:10.1371/journal.pcbi.1004316.

Rivolta, D., Puce, A., Williams, M.A., 2016. Editorial: Facing the other: Novel theories and methods in face perception research. Front Hum. Neurosci. 10, 32. doi:10.3389/fnhum.2016.00032.

Russell, J.A., 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39, 1161–1178. doi:10.1037/h0077714.

Schupp, H.T., Ohman, A., Junghofer, M., Weike, A.I., Stockburger, J., Hamm, A.O., 2004. The facilitated processing of threatening faces: An ERP analysis. Emotion 4, 189–200. doi:10.1037/1528-3542.4.2.189.

Schweinberger, S.R., Neumann, M.F., 2016. Repetition effects in human ERPs to faces. Cortex 80, 141–153. doi:10.1016/j.cortex.2015.11.001.

Schweinberger, S.R., Pickering, E.C., Burton, A.M., Kaufmann, J.M., 2002. Human brain potential correlates of repetition priming in face and name recognition. Neuropsychologia 40, 2057–2073. doi:10.1016/s0028-3932(02)00050-7.

Serre, T., Wolf, L., Poggio, T., 2005. Object recognition with features inspired by visual cortex. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Ieee, pp. 994–1000.

Shalchy, M.A., Pergher, V., Pahor, A., Van Hulle, M.M., Seitz, A.R., 2020. N-back related ERPs depend on stimulus type, task structure, pre-processing, and lab factors. Front. Hum. Neurosci. 14, 549966. doi:10.3389/fnhum.2020.549966.

Smith, F.W., Smith, M.L., 2019. Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks. Neuroimage 195, 261–271. doi:10.1016/j.neuroimage.2019.03.065.

Sugase, Y., Yamane, S., Ueno, S., Kawano, K., 1999. Global and fine information coded by single neurons in the temporal visual cortex. Nature 400, 869–873. doi:10.1038/23703.

Van den Stock, J., de Gelder, B., 2014. Face identity matching is influenced by emotions conveyed by face and body. Front. Hum. Neurosci. 8, 53. doi:10.3389/fnhum.2014.00053.

Vida, M.D., Nestor, A., Plaut, D.C., Behrmann, M., 2017. Spatiotemporal dynamics of similarity-based neural representations of facial identity. Proc. Natl. Acad. Sci. U. S. A. 114, 388–393. doi:10.1073/pnas.1614763114.

Wardle, S.G., Taubert, J., Teichmann, L., Baker, C.I., 2020. Rapid and dynamic processing of face pareidolia in the human brain. Nat. Commun. 11, 4518. doi:10.1038/s41467-020-18325-8.

Wegrzyn, M., Riehle, M., Labudda, K., Woermann, F., Baumgartner, F., Pollmann, S., Bien, C.G., Kissler, J., 2015. Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. Cortex 69, 131–140. doi:10.1016/j.cortex.2015.05.003.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: The SHINE toolbox. Behav. Res. Methods 42, 671–684. doi:10.3758/BRM.42.3.671.

Williams, M., Moss, S., Bradshaw, J., Mattingley, J., 2005. Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. Vis. Cogn. 12, 29–50. doi:10.1080/13506280444000193.

Yang, T., Yang, Z., Xu, G., Gao, D., Zhang, Z., Wang, H., Liu, S., Han, L., Zhu, Z., Tian, Y., Huang, Y., Zhao, L., Zhong, K., Shi, B., Li, J., Fu, S., Liang, P., Banissy, M.J., Sun, P., 2020. Tsinghua facial expression database - A database of facial expressions in Chinese young and older women and men: Development and validation. PLoS One 15, 1–14. doi:10.1371/journal.pone.0231304.

Yuen, K., Johnston, S., Martino, F., Sorger, B., Formisano, E., Linden, D., Goebel, R., 2012. Pattern classification predicts individuals' responses to affective stimuli. Transl. Neurosci. 3, 278–287. doi:10.2478/s13380-012-0029-6.

Zhang, H., Japee, S., Nolan, R., Chu, C., Liu, N., Ungerleider, L.G., 2016. Face-selective regions differ in their ability to classify facial expressions. Neuroimage 130, 77–90. doi:10.1016/j.neuroimage.2016.01.045.