TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214
DOI: 10.26599/TST.2018.9010000

Exploring the Neural Basis and Validity of Ordinal Emotion Representation through EEG

Xuyang Chen, Xin Xu, Dan Zhang, Quanying Liu, and Xinke Shen*

Abstract: Accurately measuring emotion is a major challenge in advancing the understanding of human emotion and developing emotional artificial intelligence. In many existing studies, participants' emotional ratings in interval scales are considered the true reflection of their emotional experiences. However, recent research suggests that ordinal annotations of emotions can more accurately capture the emotional expression process, providing a potential method for more precise emotion measurement. However, our understanding of the characteristics and validity of this new form of emotion representation is still relatively lacking. In particular, there is a lack of research using neural signals to explore the validity and neural basis of ordinal emotion representation. In this study, we used a video-elicited EEG dataset (n = 123) to identify the neural basis of ordinal emotion representation and demonstrate its validity from a neural perspective. Furthermore, we explored various characteristics of ordinal emotion representation, showing how it is superior to the interval form. First, we conducted inter-situation representational similarity analysis (RSA) and inter-subject RSA to test the degree to which ordinal representation captures both group commonalities and individual differences of emotion. Next, we investigated the characteristics of ordinal representation under different combinations of emotion items, including uni-variate and multivariate emotions, positive and negative emotions. Our results show that both group commonalities and inter-subject variations in EEG features are better explained by ordinal emotion representations than by interval ones. Multivariate ordinal representations showed better inter-subject reliability and higher representational similarity with EEG features compared to univariate counterparts, highlighting the co-occurrence nature of human emotions. Compared to negative emotions, ordinal representation showed greater improvements for positive emotions, suggesting that the complexity of positive emotions is well captured by ordinal representations. Taken together, these findings demonstrate that multivariate ordinal emotion ratings provide a more accurate measure of real emotional experience, which is crucial for enabling machines to precisely understand and express human emotions.

Key words: Ordinal Emotion Representations; EEG; Representational Similarity Analysis; Emotional Artificial Intelligence.

1 Introduction

Grasping human emotions and measuring them with precision remains a core challenge in psychological research and the advancement of emotional artificial intelligence (AI). Emotional AI seeks to empower machines to recognize, interpret, and respond to human emotions, showing great potential for enhancing human-computer interactions, advancing mental health diagnostics, and developing more empathetic technologies. Achieving accurate emotion measurement is critical for advancing these areas, as it ensures emotional AI can effectively comprehend and engage with human emotional states.

However, current emotion measurement techniques

have significant limitations, mainly stemming from the widespread use of interval-based approaches [1]. Interval data poses challenges for capturing subjective constructs like emotions, as it may inaccurately reflect a person's actual emotional state. Such misrepresentation occurs because the externalization of emotion experience is inherently relative rather than absolute [2, 3]. This makes the absolute differences in interval scores unstable and noisy [2, 4], thus not suitable for describing emotions. Additionally, interval measurements often exhibit low reliability, especially in dynamic contexts where emotions fluctuate over time [3]. These limitations underscore the need for more precise methods of emotion annotation.

Recent research has proposed that ordinal representations may provide a more accurate reflection of emotional experiences [3]. Unlike interval ratings, ordinal scales are more aligned with how humans naturally perceive and express emotions, which are inherently comparative. Empirical evidence highlights several benefits of ordinal annotations: 1) Greater inter-rater reliability: Ordinal scales reduce subjective biases and inconsistencies commonly found with interval ratings, leading to more consistent results across different raters [3, 7]. 2) Stability and robustness: Ordinal measures capture the rank order of emotional intensity without assuming exact differences between emotions, making them less vulnerable to distortions caused by extreme ratings [5, 6]. 3) Better alignment with human perception theories: Ordinal data more accurately reflects the comparative nature of emotional experiences, avoiding the pitfalls of misrepresenting emotional states through fixed interval scales [3]. These benefits suggest that ordinal emotion representations may better capture the structure of emotional experiences and provide a promising alternative to traditional interval measurement methods [7].

• Xuyang Chen, Xin Xu, Quanying Liu and Xinke Shen are with the Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen 518055, China. E-mail: {chenxy9, 12211762}@mail.sustech.edu.cn; {liuqy, shenxk}@sustech.edu.cn

• Dan Zhang is with the Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing 100084, China. E-mail: dzhang@tsinghua.edu.cn

∗ To whom correspondence should be addressed.

Manuscript received: 2024-x-x; accepted: 2024-x-x

However, the validity of ordinal measurements of emotions is not fully investigated. Existing research on the validity of ordinal forms of emotion has been limited to peripheral physiological (e.g., skin conductance [8–10], heart rate [10]) and non-physiological signals such as speech [8]. Currently, there is a lack of validation based on neural signals, testing whether ordinal forms of emotion better corresponds to the neural activity in human brain. Validating the ordinal form of emotion using neural signals can demonstrate its neurophysiological substrates. Uncovering the neural basis of different emotion categories is crucial for understanding the underlying process of emotion externalization. Besides, the current understanding of the characteristics of this new form of emotion representation is still lacking — how is ordinal form superior, and what specific aspects of emotions it can better describe, remain largely unknown.

To address the above issues, we used a video-elicited emotional EEG dataset with 123 subjects to conduct validation based on neural signals and identified the neural representations of ordinal emotions. Next, we explored the performance of ordinal emotions in terms of group commonalities, individual differences, uni-variate/multivariate emotions, and positive/negative emotions. To be specific, first, recent studies have proposed inter-subject representational similarity analysis (RSA) and inter-situation RSA as powerful methods for studying individual differences and group commonalities in emotions [11, 12]. These methods can also be used to reveal how good the ordinal form describes individual differences and group commonalities in emotional experiences, in a way that aligns with the neural signal. Second, analyses under different combinations of emotional items can demonstrate the characteristics of the ordinal emotions on multiple aspects. For example, dividing emotions into positive and negative categories can show the relationship between these two primary emotional groups and the ordinal form; viewing emotions as multidimensional vectors can capture the complexity and co-occurrence of emotional experiences, allowing for a more nuanced characterization of emotions [13].

2 Methods

2.1 Data Acquisition and Processing

2.1.1 The FACED Dataset

To validate the proposed algorithm, we used data from the Finer-grained Affective Computing EEG Dataset (FACED) [14]. The FACED dataset aims to provide fine-grained emotional computing data, with a balanced classification of both positive and negative emotional aspects.

This dataset contains 32-channel EEG recordings from 123 participants, each of whom watched 28 emotionally evocative video clips categorized into nine distinct emotions: amusement, inspiration, joy, tenderness, anger, fear, disgust, sadness, and neutral. Each negative and positive emotion category included three video clips, while the neutral category featured four. The clips had an average length of about 66 seconds, ranging from 34 to 129 seconds. After viewing each clip, participants were asked to report their subjective emotional experiences across 12 items, covering eight specific emotions (anger, fear, disgust, sadness, amusement, inspiration, joy, tenderness) and four broader dimensions (valence, arousal, liking, familiarity). Ratings were provided on a continuous scale from 0 to 7, where 0 represented "very negative" for valence and "not at all" for the other dimensions, while 7 represented "very positive" or "very much." The analysis focused on multivariate ratings across the eight specific emotion categories.

2.1.2 EEG Data Preprocessing

The EEG signals were downsampled to 125 Hz and bandpass filtered between 0.5 to 40 Hz using a fourth-order zero-phase-shift Butterworth filter. Noisy channels were identified manually by assessing variance and outlier values, with an average of 0.06±0.26 noisy channels per trial. These noisy channels were interpolated using the average of three neighboring channels. Independent component analysis (ICA) was applied to eliminate ocular and muscle artifacts, with 2-4 components being removed per participant. The data were then re-referenced to the average of the two mastoid channels. Preprocessing steps utilized Noisetools [15] and Fieldtrip toolboxes[16].

2.2 The Computation of Distance and Ordinal Representations

The core idea behind ordinal representation is that individuals evaluate emotions relatively. Thus, we transformed the emotional annotations for each participant individually. To extract ordinal information from the original interval annotations (Fig. 1a), we first calculated the relative differences within each dimension to derive a distance representation (Fig. 1b). Next, the magnitude of the differences, which is irrelevant in ordinal representation, was discarded, leaving only the sign (indicating the greater/lesser relationship in ordinal representation, Fig. 1c). The specific computational process is outlined below.

2.2.1 Distance Representation

For a participant's annotations on E emotions across all N situations, $Y_1, Y_2, ..., Y_n$, where $Y_i \in \mathbb{R}^E$, all annotations forms a network with N nodes and $N \times (N-1)/2$ edges. Rather than expressing the emotion of each trial in its original emotional space, we represent emotions based on their relative positions within the network. In this network, the position of each node is determined solely by its relative relationship to other nodes, making it independent of absolute coordinates. These relative positions are described through a set of distance vectors, and this representation is referred to as the distance representation. Specifically, for Y_i, the distance representation consists of a set of distance vectors $\{Dist_{ij} | j = 1, 2, ..., i, ..., N\}$, where:

$$Dist_{ij} = Y_i - Y_j \tag{1}$$

The vector $Dist_{ij} \in \mathbb{R}^E$ represents the distance between node Y_i and Y_j. By concatenating the one-dimensional distance vectors, we generate the distance representation matrix $D_i \in \mathbb{R}^{E \times N}$.

2.2.2 Ordinal Representation

In ordinal measurement, the specific numerical differences between two annotations hold no intrinsic meaning. Therefore, we convert the distances into ordinal categories—greater than, less than, or equal to. This transformation shifts the focus from the actual magnitude of differences to the relative ordering of the nodes, capturing the ranking relationships instead. This process is represented as follows:

$$O_{i,e,n} = \begin{cases} -1, & \text{if } D_{i,e,n} < 0 \\ 1, & \text{if } D_{i,e,n} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$
$$e = 1, 2, ..., E; n = 1, 2, ..., N$$

Where the values 1, 0, and -1 are used to represent "greater than," "equal to," and "less than" in ordinal measurement. The transformed matrix $O_i \in \mathbb{R}^{E \times N}$captures only the size relationship information from the distance representation, while discarding the numerical values. This ordinal representation retains

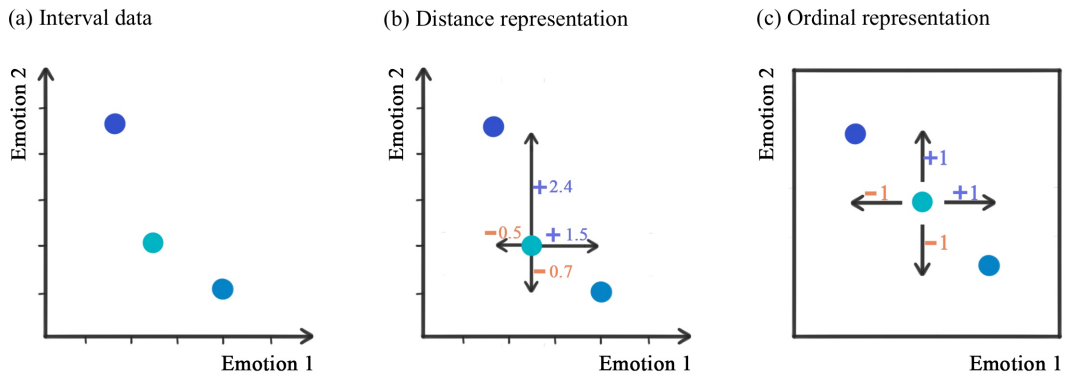


Fig. 1 General idea of extracting ordinal representation from interval annotations. (a) the original annotation data in interval scale. (b) Considering the relative positions on each emotion dimension. (c) Only preserving the rank information yields the ordinal representation.

only the relative ranks of the annotations, making it equivalent to ordinal measurement.

2.3 Inter-rater Reliability

Inter-rater reliability evaluates the level of agreement among a group of raters [17] and is widely used to assess the reliability of emotion annotations, and also an important indicator of whether an emotion measure is reasonable. Previous research has shown that ordinal emotions have higher inter-rater reliability than interval emotions. Here, we further examine the reliability improvements provided by ordinal representations, considering both uni-variate/multivariate and positive/negative emotions. We assess inter-rater reliability by measuring the consistency of the inter-situation similarity matrix across participants. A higher consistency indicates that participants rate emotions in a similar manner across different situations, as compared to how other participants rate them. In this study, we use the four-dimensional vector composed of ”joy,” ”tenderness,” ”inspiration,” and ”amusement” to represent positive emotions, and the four-dimensional vector composed of ”anger,” ”disgust,” ”fear,” and ”sadness” to represent negative emotions. On the other hand, an eight-dimensional vector composed of all eight emotion items is used to represent multivariate emotions, while each single emotion item represents uni-variate emotions. We then calculate each participant’s inter-situation emotion similarity matrix based on both ordinal and non-ordinal emotion representations, and compute the consistency across participants.

2.3.1 Inter-situation Similarity Matrix of Ratings for Each Subject

First, for each participant, we calculated the inter-situation similarity matrix of their ratings. Here, we denote $S_i^{sit_rating} \in \mathbb{R}^{v\times v}$ as the inter-situation similarity matrix of emotion annotations for participant i. It is computed based on the annotations across all situations for that participant, R_i. When using multivariate emotion, $R_i \in \mathbb{R}^{v\times E}$ includes subject i’s ratings of v situations (i.e., elicitation videos) on all E emotion items; When using uni-variate emotion, $R_i \in \mathbb{R}^{v}$ includes subject i’s ratings of v situations on one specific emotion item; When using positive emotion, $R_i \in \mathbb{R}^{v\times E_{pos}}$ includes subject i’s ratings of v situations on E_{pos} positive emotion items; When using negative emotion, $R_i \in \mathbb{R}^{v\times E_{neg}}$ includes subject i’s ratings of v situations on E_{neg} negative emotion items.

For the computation process, R_i is L2-normalized along the situation dimension to obtain normalized annotations $R_i^{(1)}$. Next, the annotations are multiplied by their own transpose to obtain the inter-situation similarity matrix of emotion annotations:

$$S_i^{sit_rating} = R_i^{(1)}(R_i^{(1)})^T, i = 1, 2, ..., u \quad (3)$$

where u is the total number of subjects. Each element of the similarity matrix represents the cosine similarity between the emotion annotations of two specific situations.

2.3.2 Inter-rater Consistency of Inter-situation Similarity Matrices

After obtaining the inter-situation similarity matrix for each participant’s ratings, which is $\{S_1^{sit}, S_2^{sit}, ..., S_u^{sit}\}$, the next step is to compute

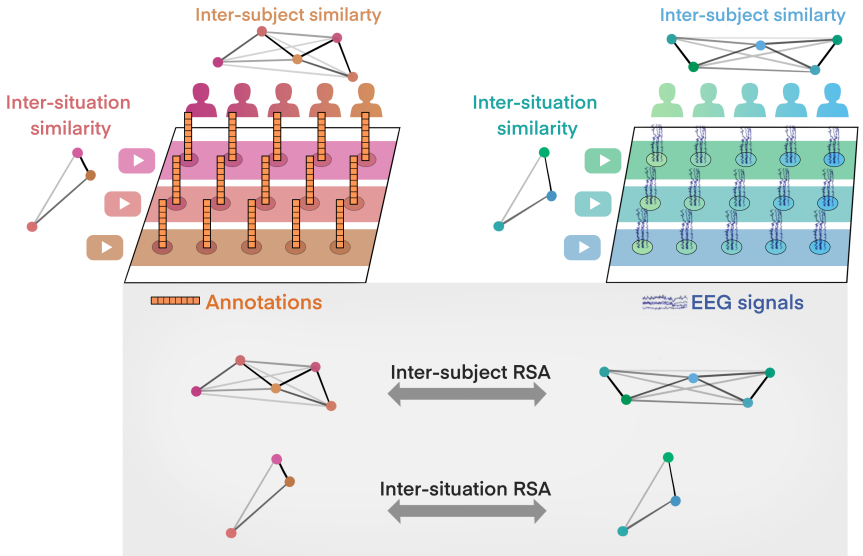


Fig. 2 The overall illustration of inter-subject representational similarity analysis (RSA) and inter-situation RSA.

their inter-rater consistency. We enumerate all possible pairings between all participants and calculate the Pearson's correlation of their similarity matrices. The resulting distribution of all $u \times (u-1)/2$ r-values demonstrates the inter-rater consistency of the inter-situation similarity matrices. By definition, this reflects the inter-rater reliability of the emotion ratings.

2.4 Representational Similarity Analysis

Representational Similarity Analysis (RSA) is a commonly used multivariate analysis method in cognitive neuroscience, designed to quantify and compare information representations in the brain or computational models. RSA evaluates representations in a brain region or computational model by constructing a distance matrix of response patterns, highlighting which differences between stimuli are accentuated and which are minimized [18]. In this study, we perform two types of RSA (Fig. 2): inter-situation RSA and inter-subject RSA. Inter-situation RSA involves comparing the similarity patterns of neural signals across different situations with the similarity patterns of emotion ratings (where "situations" refer to the emotion-eliciting materials). This analysis aims to test how well the situation-varying patterns in emotion ratings shared by all individuals are aligned with those of the neural signals. Inter-subject RSA, on the other hand, compares the similarity patterns of neural signals across participants with the similarity of emotion ratings across those participants. This analysis aims to test how well the individual difference of emotion ratings align with that of the neural signals, considering the inter-subject discrepancy across all situations.

2.4.1 Inter-situation RSA

In inter-situation RSA, for each participant, we generate a set of inter-situation similarity matrices based on their emotion annotations across all situations. Similarly, another set of inter-situation similarity matrices is generated using their neural signals across all situations [11]. The RSA process is conducted between the inter-situation similarity matrices of emotion ratings and the inter-situation similarity matrices of neural signals (Fig. 3a).

(1) Inter-situation Similarity Matrix of Emotion Ratings: For each subject i, the inter-situation similarity matrix of emotion ratings $S_i^{sit} \in \mathbb{R}^{v \times v}$ is computed based on all E emotion items, according to formula (3).

(2) Inter-situation Similarity Matrix of Neural Signals: The Inter-situation similarity matrix of neural signals for participant i, $S_i^{sit_neural} \in \mathbb{R}^{v \times v \times M \times B}$ is computed from the EEG features of that participant in all v, situations, denotes as $X_i \in \mathbb{R}^{v \times M \times B}$, where M is the number of channels and B is the number of frequency bands.

Specifically, the signals are filtered into five frequency bands: Delta: 0.5–3 Hz, Theta: 4–7 Hz, Alpha: 8–13 Hz, Beta: 14–29 Hz, Gamma: 30–40 Hz . Next, frequency domain features are extracted, including Differential Entropy (DE) and Power Spectral Density (PSD). Difference between

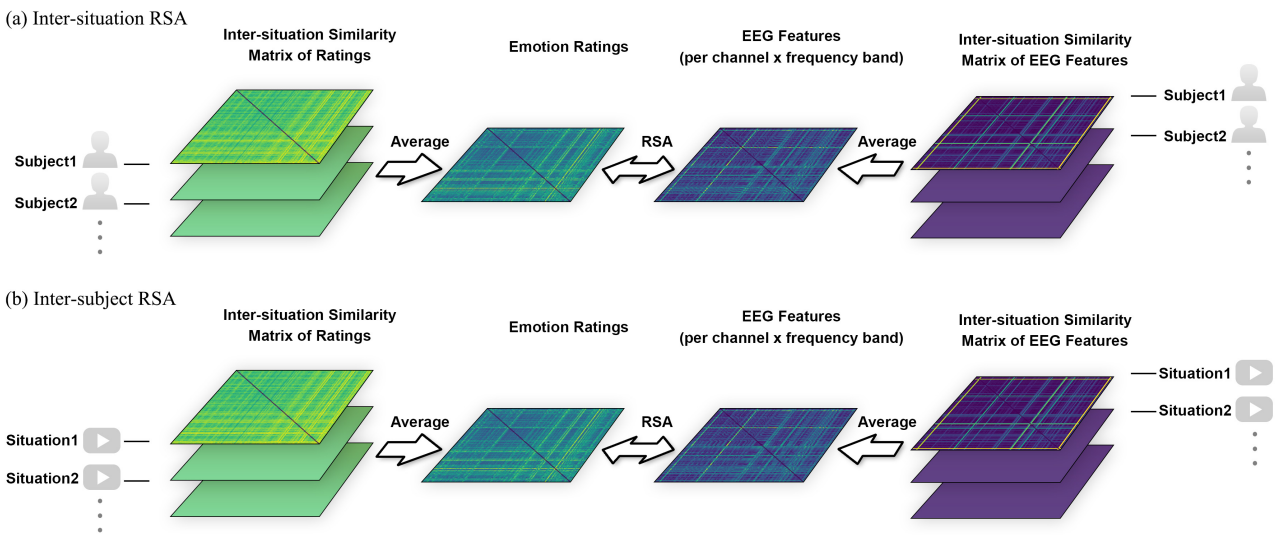


Fig. 3 Pipeline of Inter-situation RSA (above) and inter-subject RSA (below). Violet cubes stands for similarity matrices of EEG, while green squares stands for similarity matrices of annotation.

each feature dimension is computed to obtain the inter-situation dissimilarity matrix of neural signals $\Delta^{sit_neural} \in \mathbb{R}^{u\times v\times v\times M\times B}$:

$$\Delta^{sit_neural}_{i,v_1,v_2,m,b} = |X_{i,v_1,m,b} - X_{i,v_2,m,b}|$$
$$i = 1, 2, ..., u; m = 1, 2, ..., M \quad (4)$$
$$b = 1, 2, .., B; v_1, v_2 = 1, 2, ..., V$$

where $\Delta^{sit_neural}_{i} \in \mathbb{R}^{v\times v\times M\times B}$ represents the inter-situation difference patterns of participant i. Each element of corresponds to the difference of an EEG feature between two situations. Then, the dissimilarity matrix Δ^{sit_neural} was converted into similarity matrix S^{sit_neural} by normalizing it and subtracting it from an all-ones matrix.

$$S^{sit_neural} = 1 - normalize(\Delta^{sit_neural}) \quad (5)$$

Here we compute the average inter-situation similarity matrix of ratings across all participants, obtaining $\bar{S}^{sit_rating} \in \mathbb{R}^{vv}$, and the average inter-situation similarity matrix of EEG features for all participants, obtaining $\bar{S}^{sit_neural} \in \mathbb{R}^{vv\times M\times B}$. For each channel m and frequency band b, then compute the correlation coefficient between its neural similarity matrix and rating similarity matrix:

$$r^{sit}_{m,b} = pearson_corr(\bar{S}^{sit_rating}, \bar{S}^{sit_neural}_{:,:,m,b})$$
$$m = 1, 2, .., M; b = 1, 2, ..., B \quad (6)$$

where $r^{sit} \in \mathbb{R}^{MB}$ encompasses the inter-situation neural-rating consistency across all channels and frequency bands, representing the neural substrates of different emotion experiences.

2.4.2 Inter-Subject RSA

The computation of inter-subject RSA follows a structure similar to inter-situation RSA, with the key difference being that the similarity is calculated across subjects rather than situations. In inter-subject RSA, for a specific situation, a set of inter-subject similarity matrices is calculated based on the emotional annotations from all participants. Another set of inter-subject similarity matrices is generated using the neural signals from all subjects [12]. The correlation between these matrices is then used to assess the relationship between the inter-subject differences in emotion ratings and neural signals. (Fig. 3b). (1) Inter-Subject Similarity Matrix of Emotional Annotations: The inter-subject similarity matrix of situation j is denoted as $S^{sub_rating}_{j} \in \mathbb{R}^{uu}$, where u is the number of subjects. This is calculated based on the ratings of all subjects in this situation R_j. Same as the Inter-situation similarity matrix described in equation (3), when using multivariate emotion, $R_j \in \mathbb{R}^{u\times E}$; When using univariate emotion, $R_j \in \mathbb{R}^{u}$; When using positive emotion, $R_j \in \mathbb{R}^{u\times E_{pos}}$; And when using negative emotion, $R_j \in \mathbb{R}^{u\times E_{neg}}$. R_j is then L2-normalized along the subject dimension to obtain the normalized annotation $R^{(1)}_{j} \in \mathbb{R}^{uE}$. Next, the annotation is

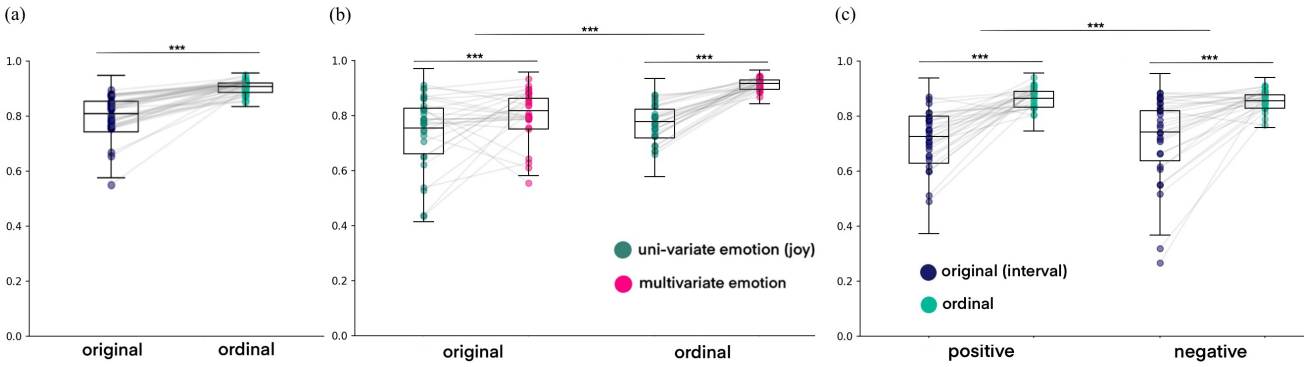


Fig. 4 The results of reliability test. (a) The inter-rater reliability demonstrated by pairwise similarity of all subjects' inter-situation similarity matrices using original interval annotation (left) and ordinal representation (right). (b) The inter-rater reliability of single-variate and multivariate emotions. This is demonstrated by the pairwise similarity of inter-situation similarity matrices of original and ordinal data, using single-variate emotions and multivariate emotion. Shown in figure is the when 'joy' is chosen as the single-variate. Under any single-variate emotions, the results are consistently similar to the shown case. (c) The inter-rater reliability of positive and negative emotions. This is demonstrated by the pairwise similarity of inter-situation similarity matrices of positive (joy, tenderness, inspiration, amusement) and negative (anger, disgust, fear, sadness) emotions, using original interval data and ordinal representation respectively.

multiplied by its transpose to obtain the inter-subject similarity matrix of emotional annotations:

$$S_j^{sub_rating} = R_j^{(1)}(R_j^{(1)})^T, j = 1, 2, ..., v \quad (7)$$

where $S_j^{sub_rating} \in \mathbb{R}^{uu}$ represents the inter-subject similarity pattern for situation j. Each element of $S_j^{sub_rating}$ represents the cosine similarity of the annotation vectors from two subjects for situation j. $S^{sub_rating} \in \mathbb{R}^{vuu}$ was averaged over the situation dimension to obtain the similarity matrix of ratings $\bar{S}^{sub_rating} \in \mathbb{R}^{uu}$.

(2) Inter-Subject Similarity Matrix of Neural Signals: The inter-subject similarity matrix of neural signals across all situations, $\bar{S}^{sub_neural} \in \mathbb{R}^{uuMB}$, is obtained from the EEG features of all subjects in all situations, denoted as $X \in \mathbb{R}^{uvMB}$. For situation j and subjects i_1, i_2, the computation is as follows:

$$\Delta_{i_1,i_2,j,:,:}^{sub_neural} = |X_{i_1,j,:,:} - X_{i_2,j,:,:}|$$
$$i_1, i_2 = 1, 2, ..., u; j = 1, 2, ..., v \quad (8)$$

Here, $\Delta_{i_1,i_2}^{sub_neural} \in \mathbb{R}^{v\times M\times B}$ represents the neural dissimilarity between subjects i_1, i_2 calculated regarding v situations, M channels and B bands. Then, the dissimilarity matrix Δ^{sub_neural} was converted into similarity matrix S^{sub_neural} by normalizing it and subtracting it from an all-ones matrix.

$$S^{sub_neural} = 1 - normalize(\Delta^{sub_neural}) \quad (9)$$

Then average $S^{sub_neural} \in \mathbb{R}^{u\times u\times v\times M\times B}$ along the third dimension (situations), yielding the average neural similarity matrix $\bar{S}^{sub_neural} \in \mathbb{R}^{u\times u\times M\times B}$.

The matrices $\bar{S}^{sub_rating} \in \mathbb{R}^{uu}$ and $\bar{S}^{sub_neural}$ reflect the inter-subject similarity patterns across all situations in neural activity or ratings, i.e., inter-subject differences. Here, for each channel m and frequency band b, we calculate the correlation coefficient between its neural similarity matrix and the rating similarity matrix:

$$r_{m,b}^{sub} = pearson_corr(\bar{S}^{sub_rating}, \bar{S}_{:,:,m,b}^{sub_neural})$$
$$m = 1, 2, .., M; b = 1, 2, ..., B \quad (10)$$

$r^{sub} \in \mathbb{R}^{MB}$ represents the correlation of inter-subject ratings' difference and EEG features' difference.

3 Results

3.1 Inter-rater Reliability

For the inter-rater reliability test, ordinal representations showed significantly higher consistency between subjects than interval ratings (Fig. 4a), which is consistent with previous research.

When comparing multivariate and univariate emotions, ordinal representation resulted in a notable reliability improvement for both cases (Fig. 4b). On the other hand, multivariate emotion representations exhibit higher inter-subject reliability compared to single-variate representations, regardless of which single-variate emotion rating is used and whether ordinal representation is applied.

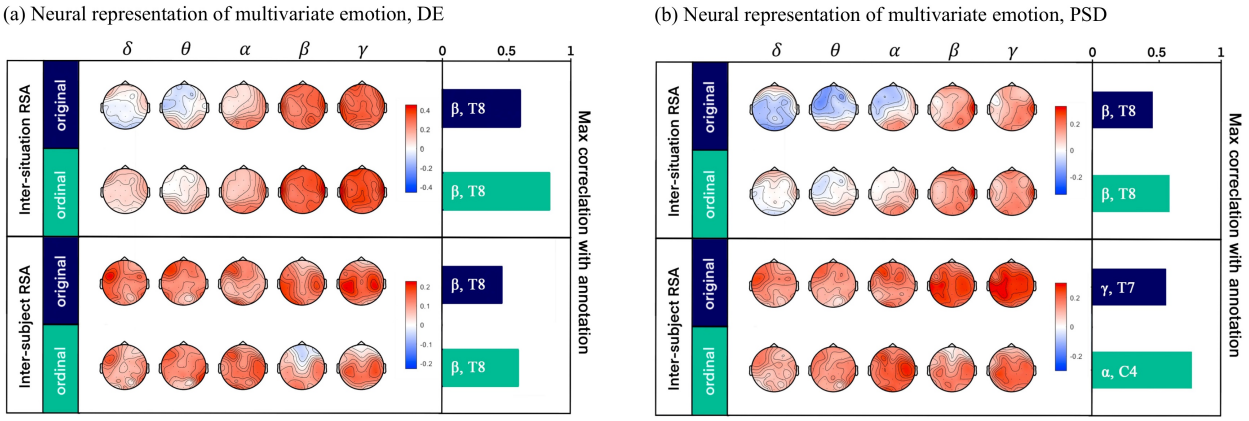


Fig. 5 The neural representation of multivariate emotion (all items). The bars at right indicates the levels of validity, represented by maximum r. The corresponding frequency bands and channels are shown in the bars. (a) Using DE as the EEG feature. (b) Using PSD as the EEG feature.

Using two-way ANOVA, we found a significant interaction between multivariate/uni-variate and ordinal/non-ordinal emotion representations (for 'joy', $F(3, 30008) = 1874.3, p < 0.001$. Other items are also significant, see supplementary). This suggests that ordinal representation is more superior when combined with multivariate emotion.

When comparing positive and negative emotions, the ordinal representation improved the reliability of both (Fig. 4c, positive: $t(7503) = 138.9, p < 0.001$, negative: $t(7503) = 104.6, p < 0.001$), with a greater increase for positive emotions (two-way ANOVA: $F(3, 30008) = 143.2, p < 0.001$).

3.2 Validity and Neural Substrates of Ordinal Emotion Representations

Ordinal representations had a higher correlation with EEG features than original annotations in both inter-situation RSA and inter-subject RSA (Fig. 5), indicating ordinal representations can better reflect both the inter-situation difference and inter-subject difference of neural activities. Regardless of the feature set used (DE or PSD), or the type of representation (ordinal or interval), the brain regions most correlated with emotional representations were concentrated in the temporal lobe at higher frequency bands (beta and gamma). Moreover, the correlation in this region and frequency band was stronger under ordinal representation compared to interval representation. The ordinal representation brought validity improvements for both uni-variate and multivariate emotions. Interestingly, multivariate emotion representations consistently exhibit higher correlation with neural signals (Fig. 6). This further confirms that significant emotional characteristics of individuals may be hidden within multivariate emotional structures, and multivariate ordinal representation is a promising approach to measure the emotion states.

While ordinal representation of both positive and negative emotions shows higher validity, the improvement brought by the ordinal representation for positive emotions is significantly greater than for negative emotions (Fig. 7, 6). Fig. 7 shows the validity described by the maximum correlation coefficient between emotion annotations and EEG components (DE features and PSD features), where positive emotions consistently show a greater increase when using the ordinal representation. Fig. 6 illustrates the validity of individual uni-variate emotion items under ordinal and non-ordinal conditions, where positive emotion items exhibit more stable and larger improvements. The EEG encoding of positive and negative emotions based on DE are shown in Fig. 8. For the result based on PSD, see supplementary materials.

4 Discussion

In this study, we used a large-scale emotional EEG dataset to uncover the neural correlates of ordinal emotional representations. Our findings revealed that ordinal representations not only exhibit higher inter-rater reliability (consistent with previous studies [3,19]) but also better capture the commonalities and individual differences in EEG features. These results emphasize

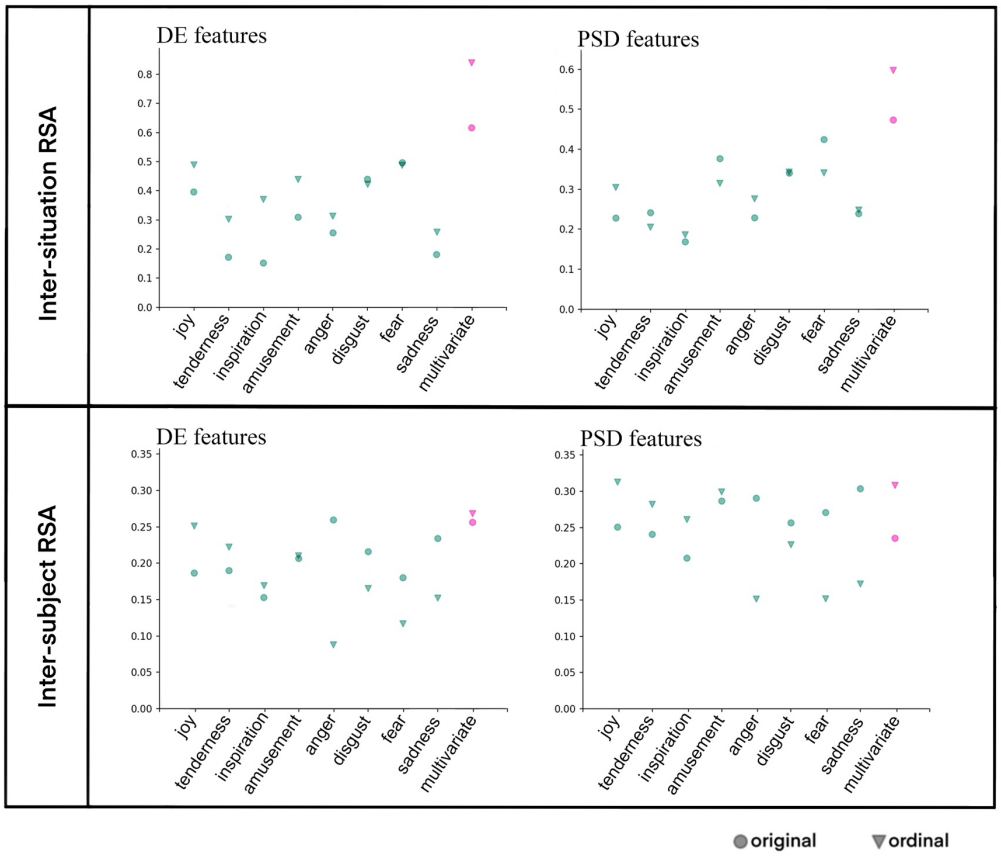


Fig. 6 The validity demonstrated by maximum correlation coefficient between EEG signals and ratings, using single-variate and multivariate emotion, and two different forms (original interval form in circle, ordinal form in triangle). Above shows the result based on inter-situation RSA, demonstrating the validity in capturing group emotional commonalities. Below shows the result based on inter-subject RSA, demonstrating the validity in capturing individual emotional differences.

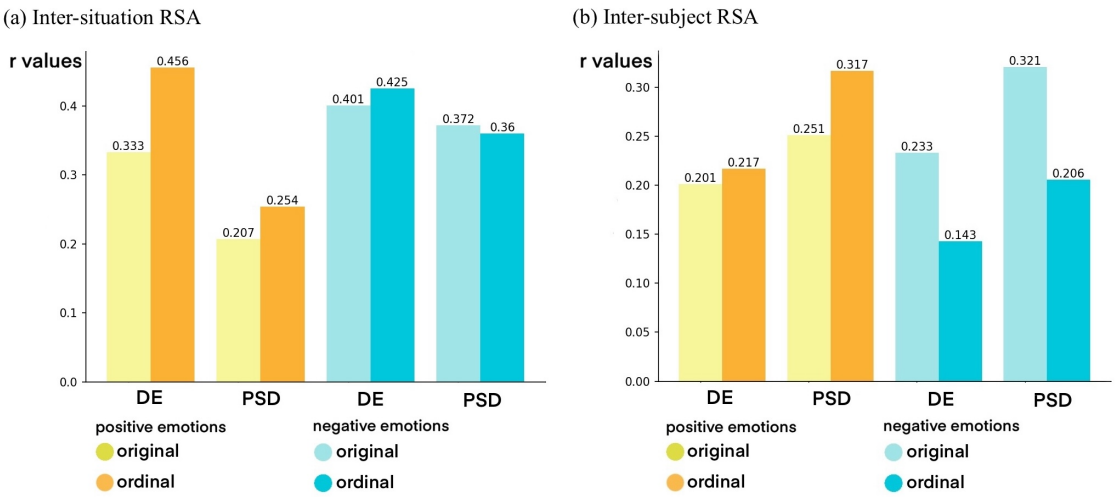


Fig. 7 The validity demonstrated by maximum correlation coefficient between EEG signals and ratings, using positive and negative emotions, and two different forms. Ordinal representation brings greater improvements for positive emotions than negative emotions. (a) validity in inter-situation RSA (b) validity in inter-subject RSA.

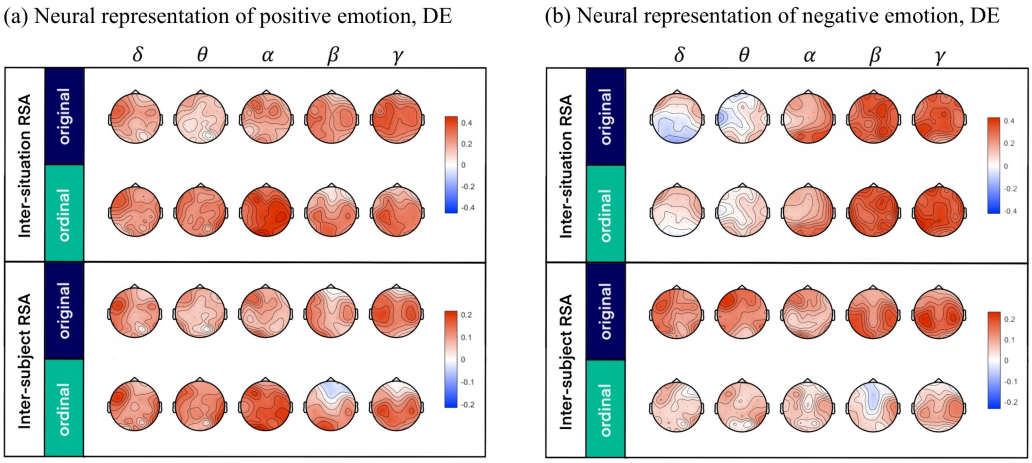


Fig. 8 The neural representation of positive emotions (joy, tenderness, inspiration, amusement) and negative emotions (anger, disgust, fear, sadness). (a) Positive emotions. (b) Negative emotions. Shown in the figure are the results based on DE features. For results based on PSD features, see supplementary materials.

the value of ordinal scales in emotion measurement, providing a more detailed understanding of the neural mechanisms underlying human emotions.

One of the most compelling discoveries was the superior alignment of multivariate ordinal emotion representations with EEG signals, compared to their uni-variate counterparts. This indicates that considering the complexity and co-occurrence of multiple emotional states provides a more robust measure of emotional experiences. This finding aligns with the understanding that human emotions are rarely experienced in isolation but rather as complex, intertwined states [20, 21].

Moreover, we found the ordinal representation brought greater improvements in reliability and validity for positive emotions compared to negative emotions. Positive emotions have more complex dynamics and co-occurrence patterns [22–24]. Hence, it is more challenging to represent positive emotions than negative emotions reliably with classical interval ratings [25]. The superior performance of ordinal representations in describing positive emotions may indicate that positive emotions possess a clearer structure in the ordinal space. It is possible that the high-dimensional, co-occurring nature of positive emotions leads to much misinterpretation when they are mapped into the interval space, and ordinal measurement has great potential to address this issue.

Additionally, the analysis revealed that EEG features in temporal regions, particularly in high-frequency bands such as beta and gamma, showed a stronger correlation with ordinal emotion representations. This is consistent with previous research suggesting that these brain regions and frequency bands play a critical role in processing complex emotional stimuli [26–28].

5 Conclusion

Our study demonstrates that ordinal emotion representations offer a more effective approach to capturing the complexity of human emotions. Compared to interval-based measurements, ordinal representations show higher inter-rater reliability and better alignment with both group commonalities and individual differences in neural activity. Notably, ordinal representations enhance the alignment with neural signals particularly when applied to multivariate and positive emotions. This highlights the co-occurrence of multiple emotions and intricate patterns of positive emotions are better reflected by ordinal scales. In summary, ordinal representations provide a robust framework for measuring emotions, ensuring greater inter-rater reliability and stronger grounding in neurophysiological data. These findings are essential for accurately representing human emotions in machines, improving machines' capacity to comprehend human emotions.

Acknowledgment

The authors wish to thank Zongsheng Li for his valuable suggestions. This work was funded in part by the National Key R&D Program of

China (2021YFF1200804), Shenzhen Excellent Youth Project (RCYX20231211090405003), Shenzhen Science and Technology Innovation Committee (RCBS20231211090748082, 2022410129, KCXFZ20201221173400001, KJZD20230923115221044).

References

[1] R. Cowie, C. Cox, J. C. Martin, A. Batliner, D. Heylen, and K. Karpouzis, "Issues in emotion recognition," 2011.

[2] M. Kim and V. Pavlovic, "Structured Output Ordinal Regression for Dynamic Facial Emotion Intensity Prediction," in Computer Vision – ECCV 2010, K. Daniilidis, P. Maragos, and N. Paragios, Eds. Berlin, Heidelberg: Springer, 2010, vol. 6313, pp. 649-662. https://doi.org/10.1007/978-3-642-15558-1-47.

[3] G. N. Yannakakis, R. Cowie, and C. Busso, "The ordinal nature of emotions: An emerging approach," IEEE Transactions on Affective Computing, vol. 12, no. 1, pp. 16-35, 2021.

[4] V. Sethu, E. M. Provost, J. Epps, C. Busso, N. Cummins, and S. Narayanan, "The ambiguous world of emotion representation," arXiv preprint arXiv:1909.00360, 2019.

[5] Y. Neuman and Y. Cohen, "Predicting change in emotion through ordinal patterns and simple symbolic expressions," Mathematics, vol. 10, no. 13, p. 2253, 2022.

[6] M. Mitsios, G. Vamvoukakis, G. Maniati, N. Ellinas, G. Dimitriou, K. Markopoulos, P. Kakoulidis, A. Vioni, M. Christidou, J. Oh, G. Jho, I. Hwang, G. Vardaxoglou, A. Chalamandaris, P. Tsiakoulis, and S. Raptis, "Improved Text Emotion Prediction Using Combined Valence and Arousal Ordinal Classification," arXiv:2404.01805 [cs.LG], 2024. https://doi.org/10.48550/arXiv.2404.01805.

[7] G. N. Yannakakis and H. P. Martínez, "Grounding truth via ordinal annotation," presented at the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, 2015, pp. 574-580. https://doi.org/10.1109/ACII.2015.7344627.

[8] H. P. Martínez, G. N. Yannakakis, and J. Hallam, "Don't Classify Ratings of Affect; Rank Them!," IEEE Transactions on Affective Computing, vol. 5, no. 3, pp. 314-326, 2014. https://doi.org/10.1109/TAFFC.2014.2352268.

[9] P. Lopes, G. N. Yannakakis, and A. Liapis, "RankTrace: Relative and unbounded affect annotation," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 158-163. https://doi.org/10.1109/ACII.2017.8273594.

[10] D. Melhart, K. Sfikas, G. Giannakakis, and G. Y. Liapis, "A Study on Affect Model Validity: Nominal vs Ordinal Labels," in *Proceedings of IJCAI 2018 2nd Workshop on Artificial Intelligence in Affective Computing*, W. Hsu and H. Yates, Eds. PMLR, 2020, vol. 86, pp. 27-34.

[11] J. Liu, X. Hu, X. Shen, S. Song, and D. Zhang, "Electrophysiological representations of multivariate human emotion experience," Cognition and Emotion, vol. 38, no. 3, pp. 378-388, 2024.

[12] X. Hu, F. Wang, and D. Zhang, "Similar brains blend emotion in similar ways: Neural representations of individual difference in emotion profiles," Neuroimage, vol. 247, p. 118819, 2022.

[13] A. Werner-Seidler, C. Hitchcock, E. Hammond, et al., "Emotional complexity across the life story: Elevated negative emodiversity and diminished positive emodiversity in sufferers of recurrent depression," Journal of Affective Disorders, vol. 273, pp. 106-112, Apr. 2020. https://doi.org/10.1016/j.jad.2020.04.060.

[14] J. Chen, X. Wang, C. Huang, X. Hu, X. Shen, and D. Zhang, "A large finer-grained affective computing EEG dataset," Scientific Data, vol. 10, no. 1, p. 740, 2023.

[15] A. de Cheveigne and J. Z. Simon, "Sensor noise suppression," J. Neurosci. Methods, vol. 168, no. 1, pp. 195-202, Feb. 2008.

[16] R. Oostenveld, P. Fries, E. Maris, and J. M. Schoffelen, "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," Computational Intelligence and Neuroscience, vol. 2011, no. 1, p. 156869, 2011.

[17] J. J. Bartko and W. T. Carpenter, "On the methods and theory of reliability," The Journal of Nervous and Mental Disease, vol. 163, no. 5, pp. 307-317, 1976.

[18] N. Kriegeskorte, M. Mur, and P. A. Bandettini, "Representational similarity analysis-connecting the branches of systems neuroscience," Frontiers in Systems Neuroscience, vol. 2, p. 249, 2008.

[19] J. Kim, K. P. Truong, V. Charisi, C. Zaga, V. Evers, and M. Chetouani, "Multimodal detection of engagement in groups of children using rank learning," in Human Behavior Understanding: 7th International Workshop, HBU 2016, Amsterdam, The Netherlands, October 16, 2016, Proceedings, vol. 7, Springer International Publishing, 2016, pp. 35-48.

[20] P. C. Ellsworth and C. A. Smith, "From appraisal to emotion: Differences among unpleasant feelings," Motivation and Emotion, vol. 12, no. 3, pp. 271-302, 1988. https://doi.org/10.1007/BF00993115.

[21] P. C. Ellsworth and C. A. Smith, "Shades of Joy: Patterns of Appraisal Differentiating Pleasant Emotions," Cognition and Emotion, vol. 2, no. 4, pp. 301-331, 1988. https://doi.org/10.1080/02699938808412702.

[22] J. G. Chipperfield, R. P. Perry, and B. Weiner, "Discrete Emotions in Later Life," *The Journals of Gerontology: Series B*, vol. 58, no. 1, pp. P23–P34, Jan. 2003.

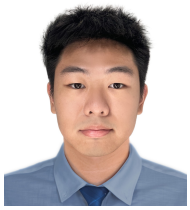
[23] M. N. Shiota, D. A. Sauter, and P. M. A. Desmet, "What are 'positive' affect and emotion?," *Current Opinion in Behavioral Sciences*, vol. 39, pp. 142-146, 2021.

[24] D. Trampe, J. Quoidbach, and M. Taquet, "Emotions in everyday life," *PloS One*, vol. 10, e0145450, 2015.

[25] B. L. Fredrickson and M. A. Cohn, "Positive emotions," in *Handbook of Emotions*, 3rd ed., 2008, pp. 777-796.

[26] L. Pessoa, "On the relationship between emotion and cognition," Nat. Rev. Neurosci., vol. 9, pp. 148-158, 2008. https://doi.org/10.1038/nrn2317.

[27] K. Yang, L. Tong, J. Shu, N. Zhuang, B. Yan, and Y. Zeng, "High gamma band EEG closely related to emotion: evidence from functional network," Front. Hum. Neurosci., vol. 14, p. 89, 2020.
[28] W. Zheng and B. Lu, "Investigating Critical Frequency Bands and Channels for EEG-based Emotion Recognition with Deep Neural Networks," IEEE Transactions on Autonomous Mental Development, vol. 7, no. 3, pp. 162-175, 2015.



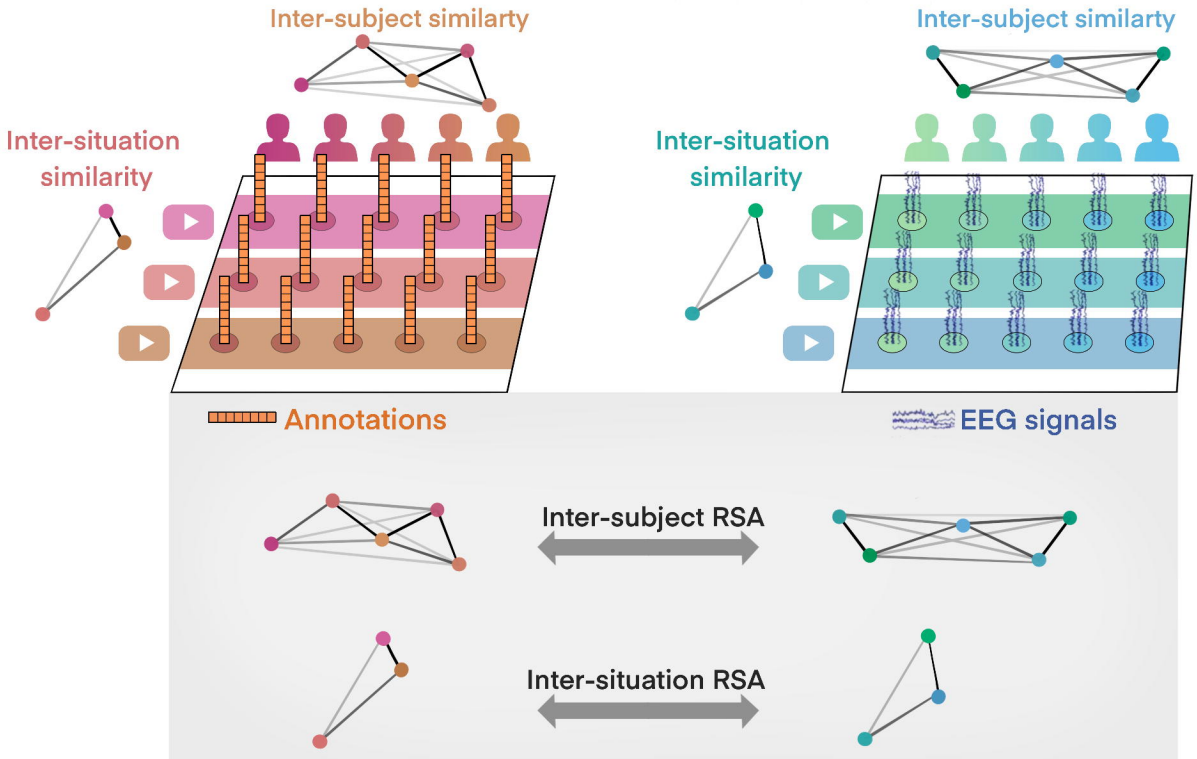
Xuyang Chen received the B.S. degree in bioinformatics from Southern University of Science and Technology, Shenzhen, China, in 2024. He is working as a research assistant in the Department of Biomedical Engineering, Southern University of Science and Technology. His current research interests focus on the neural mechanisms of emotion evaluation and emotion experiences.

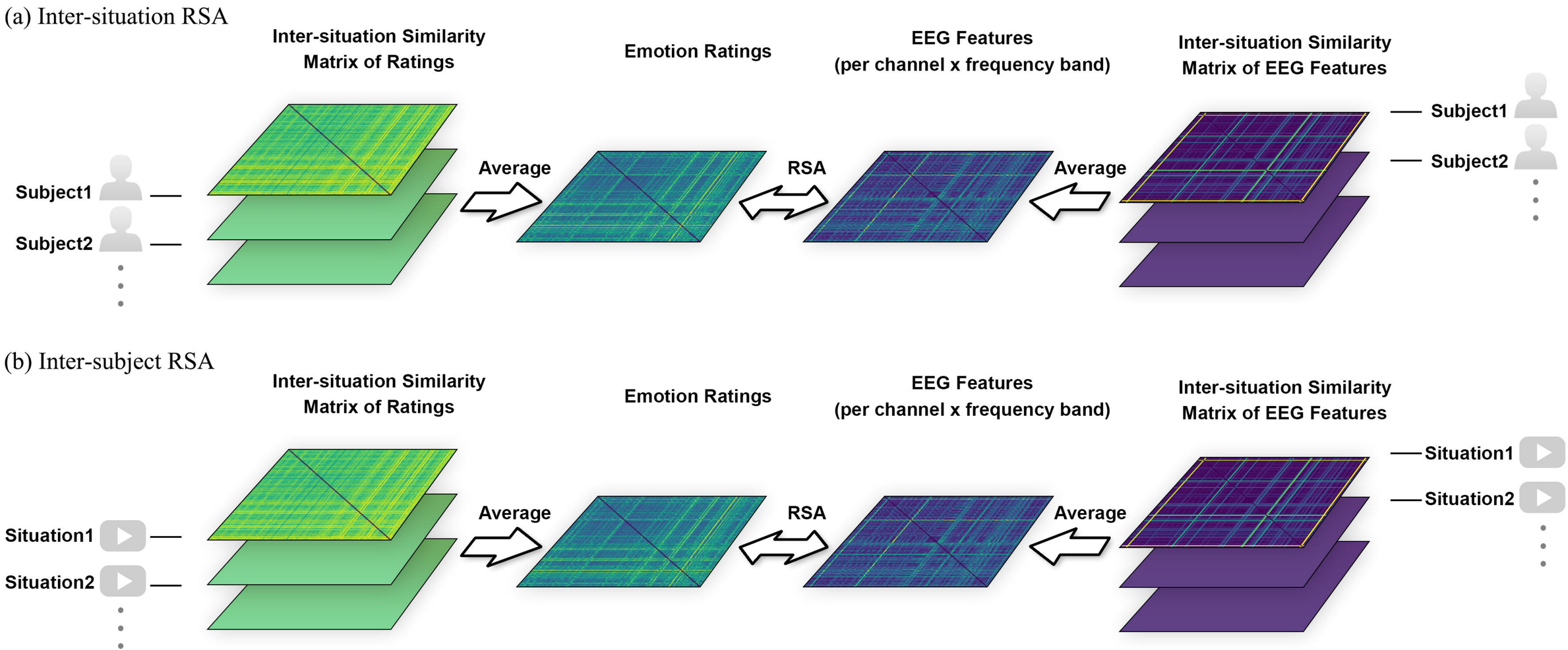
Xin Xu is an undergraduate student in the Department of Biomedical Engineering at Southern University of Science and Technology, enrolled in 2022. She is interested in affective computing, brain-computer interfaces and emotion-related neural representations.

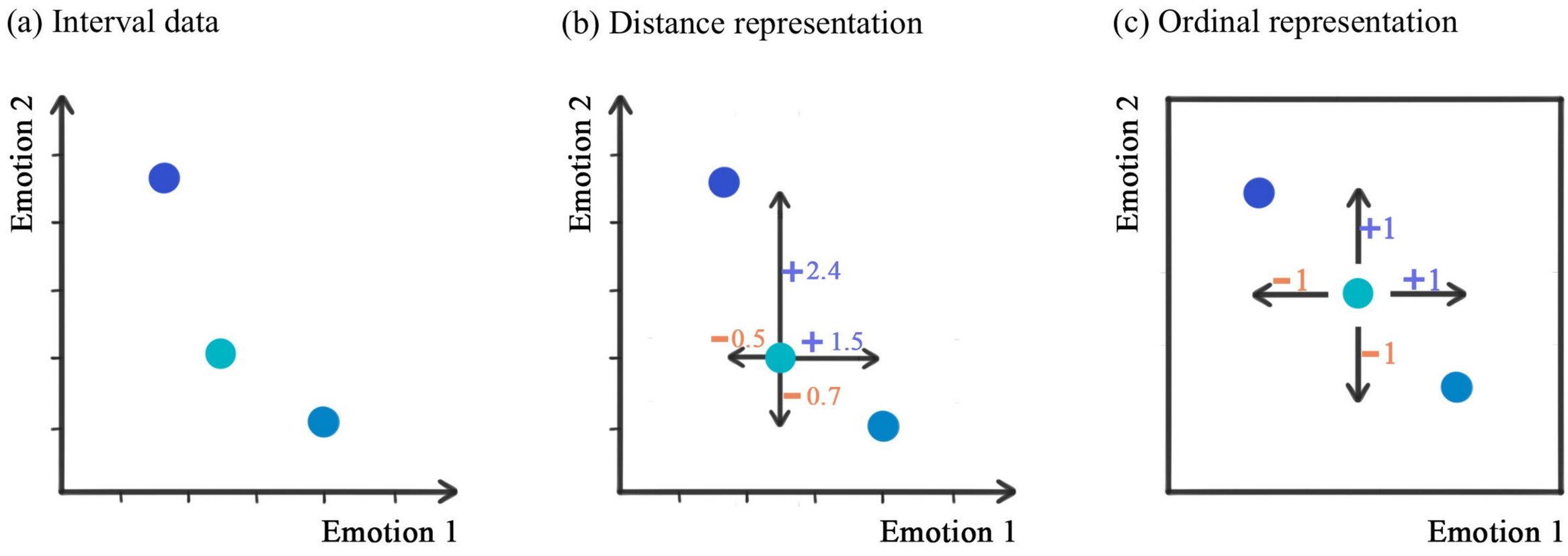
Dan Zhang received the B.E. degree in automation in 2005 and the Ph.D. degree in biomedical engineering in 2011, both from Tsinghua University, Beijing, China. He was a postdoctoral fellow in School of Medicine, Tsinghua University from 2011 to 2013. He is currently an associate professor at the Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing, China. His research interests include social neuroscience, engineering psychology, and brain-computer interfaces. He is a member of the IEEE.

Quanying Liu received the B.S. degree in electrical engineering and the M.S. degree in computer science at Lanzhou University in 2010 and 2013. After, she received her Ph.D. degree from ETH Zurich and received postdoctoral training at Caltech. She is an Assistant Professor at the Department of Biomedical Engineering, Southern University of Science and Technology and the PI of Neural Computing and Control Laboratory (NCC lab). Her research focuses on interactions among neuroscience, machine learning and control theory, including multi-modal neural signal processing (EEG, sEEG, fMRI, DTI), AI for neuroscience (explainable AI to interpret the structure and function of the brain), and optimization for neuromodulation (tES, TMS, electrical stimulation). She is the associate editor of IEEE Journal of Translational Engineering in Health and Medicine (JTEHM).

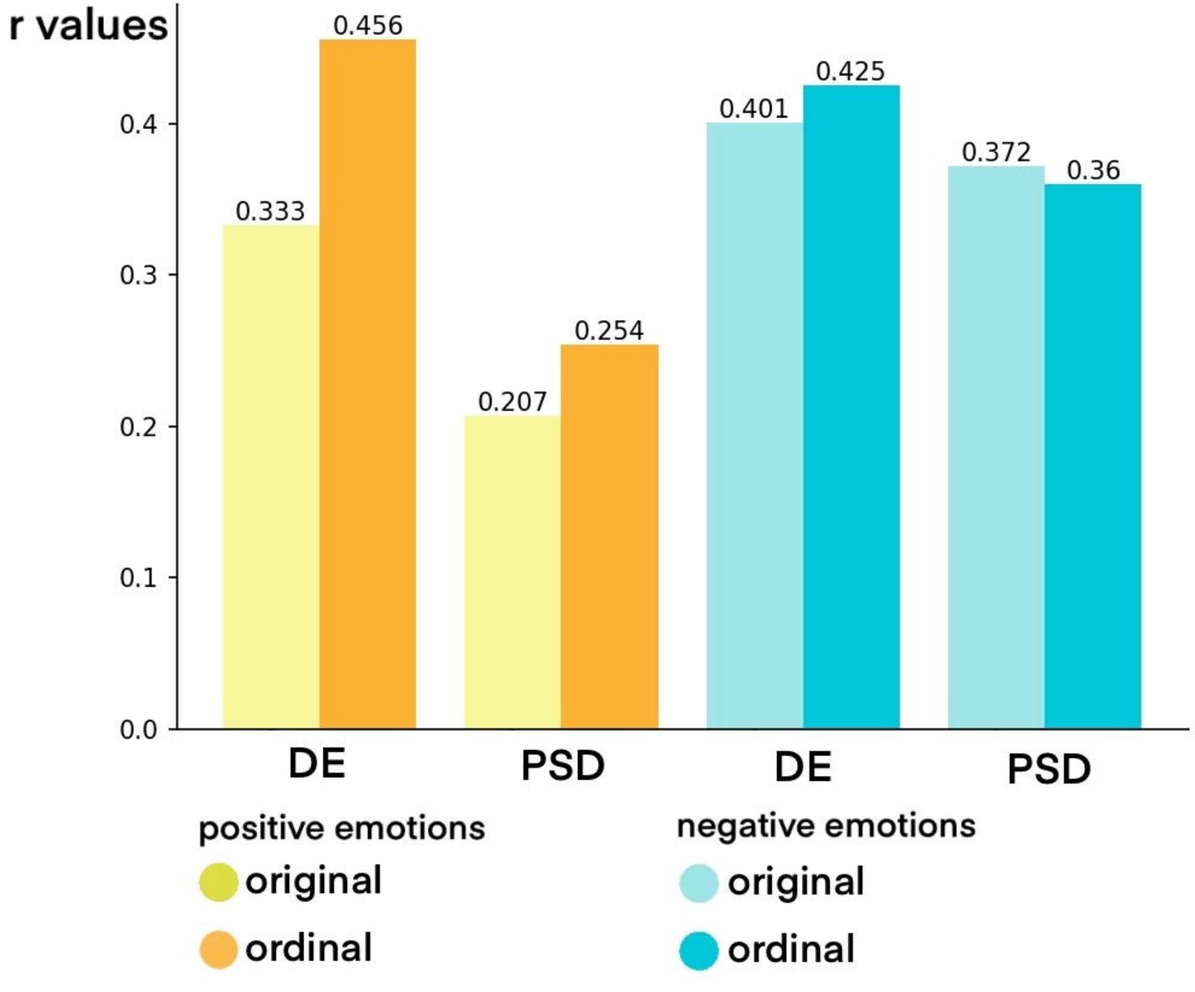
Xinke Shen received the B.E. degree in biomedical engineering from Beihang University, Beijing, China, in 2017 and the Ph.D. degree in biomedical engineering in Tsinghua University, Beijing, China, in 2023. He is currently a postdoctoral fellow in the Department of Biomedical engineering, Southern University of Science and Technology, Shenzhen, China. His research interests include affective computing and affective neuroscience.


Inter-subject similarty
Inter-situation similarity
Inter-subject similarty
Inter-situation similarity
Annotations
EEG signals
Inter-subject RSA
Inter-situation RSA

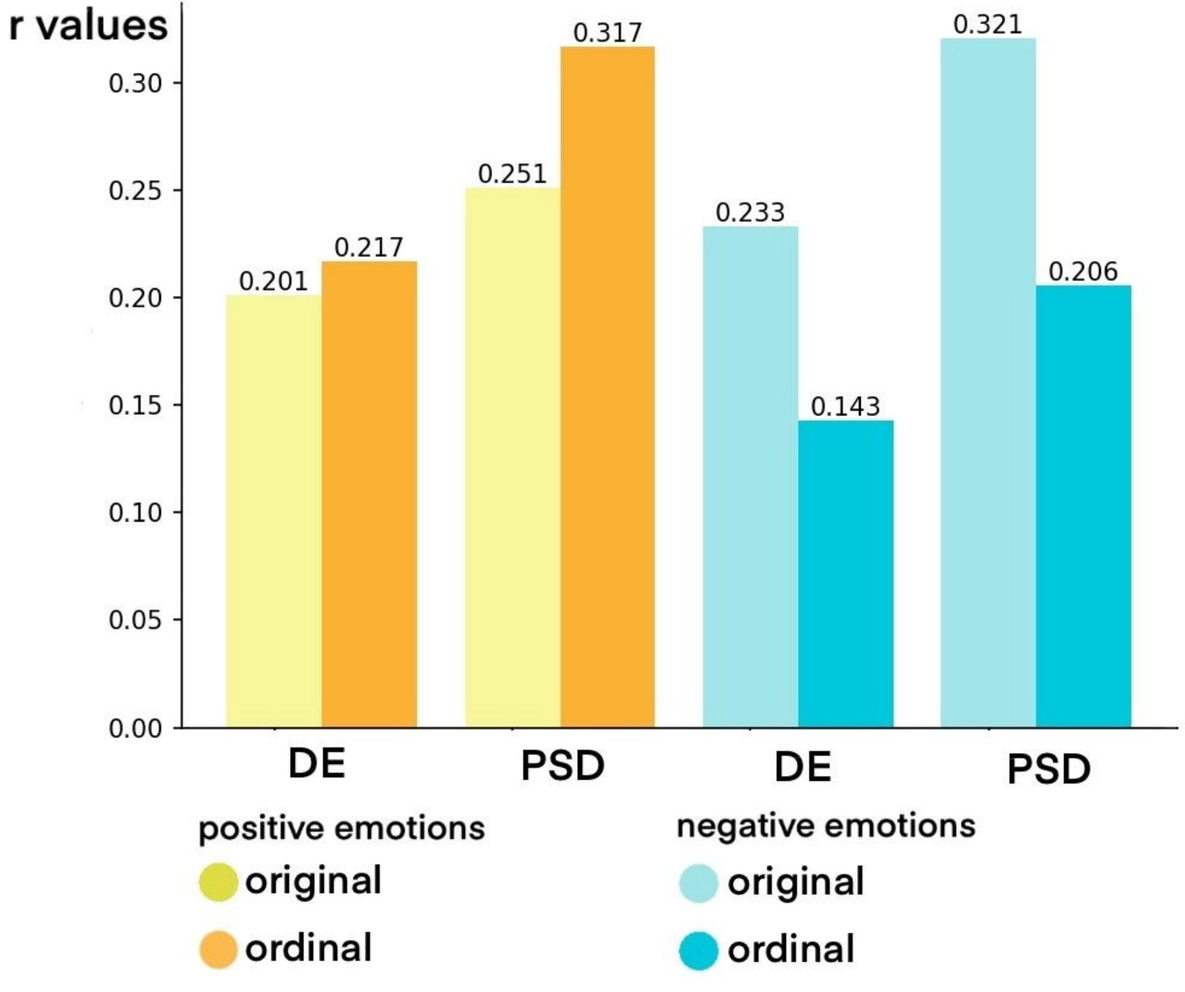

(a) Inter-situation RSA
Inter-situation Similarity Matrix of Ratings
Emotion Ratings
EEG Features (per channel x frequency band)
Inter-situation Similarity Matrix of EEG Features
Subject1
Subject2
Average
RSA
Average
Subject1
Subject2
(b) Inter-subject RSA
Inter-situation Similarity Matrix of Ratings
Emotion Ratings
EEG Features (per channel x frequency band)
Inter-situation Similarity Matrix of EEG Features
Situation1
Situation2
Average
RSA
Average
Situation1
Situation2


(a) Interval data
Emotion 2
Emotion 1
(b) Distance representation
Emotion 2
+2.4
−0.5
+ 1.5
− 0.7
Emotion 1
(c) Ordinal representation
Emotion 2
+1
−1
+1
−1
Emotion 1

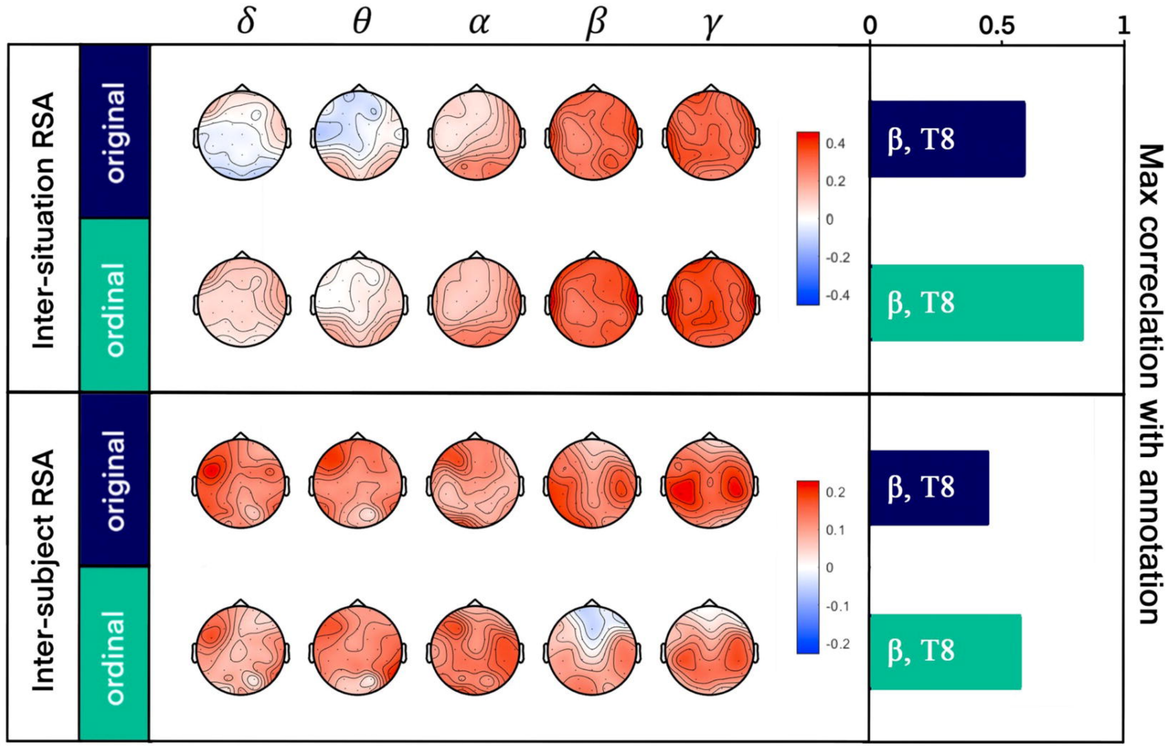
(a) Inter-situation RSA



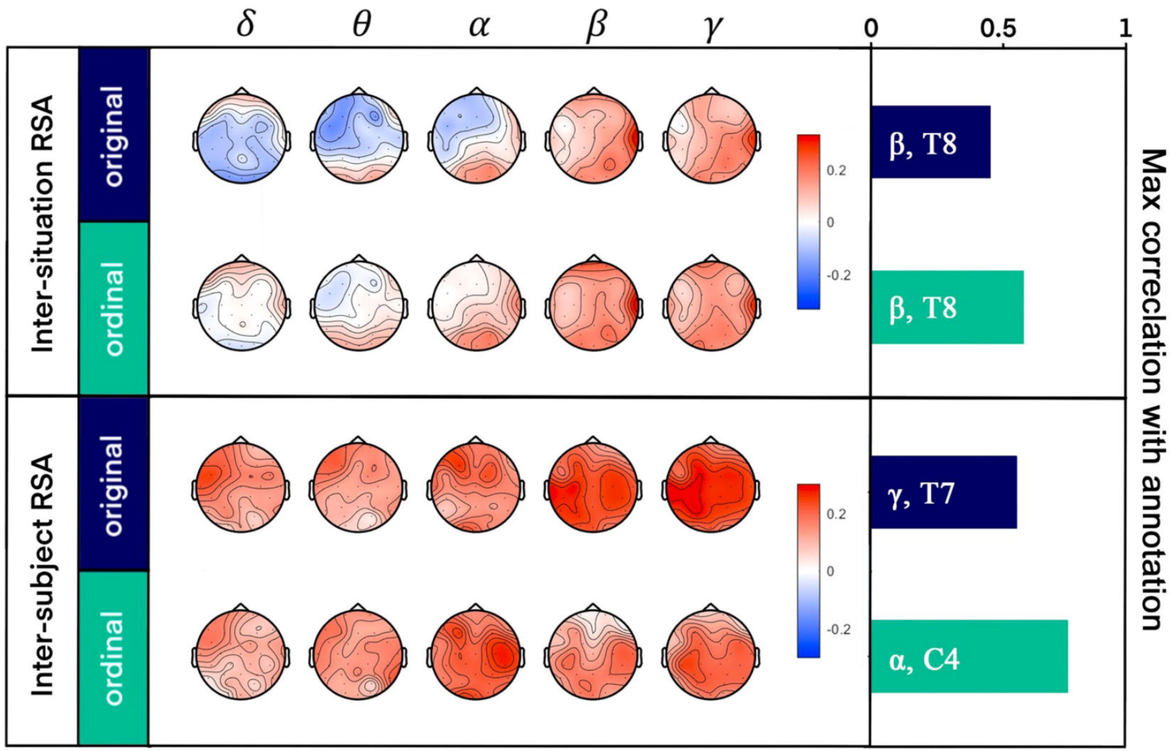
(b) Inter-subject RSA



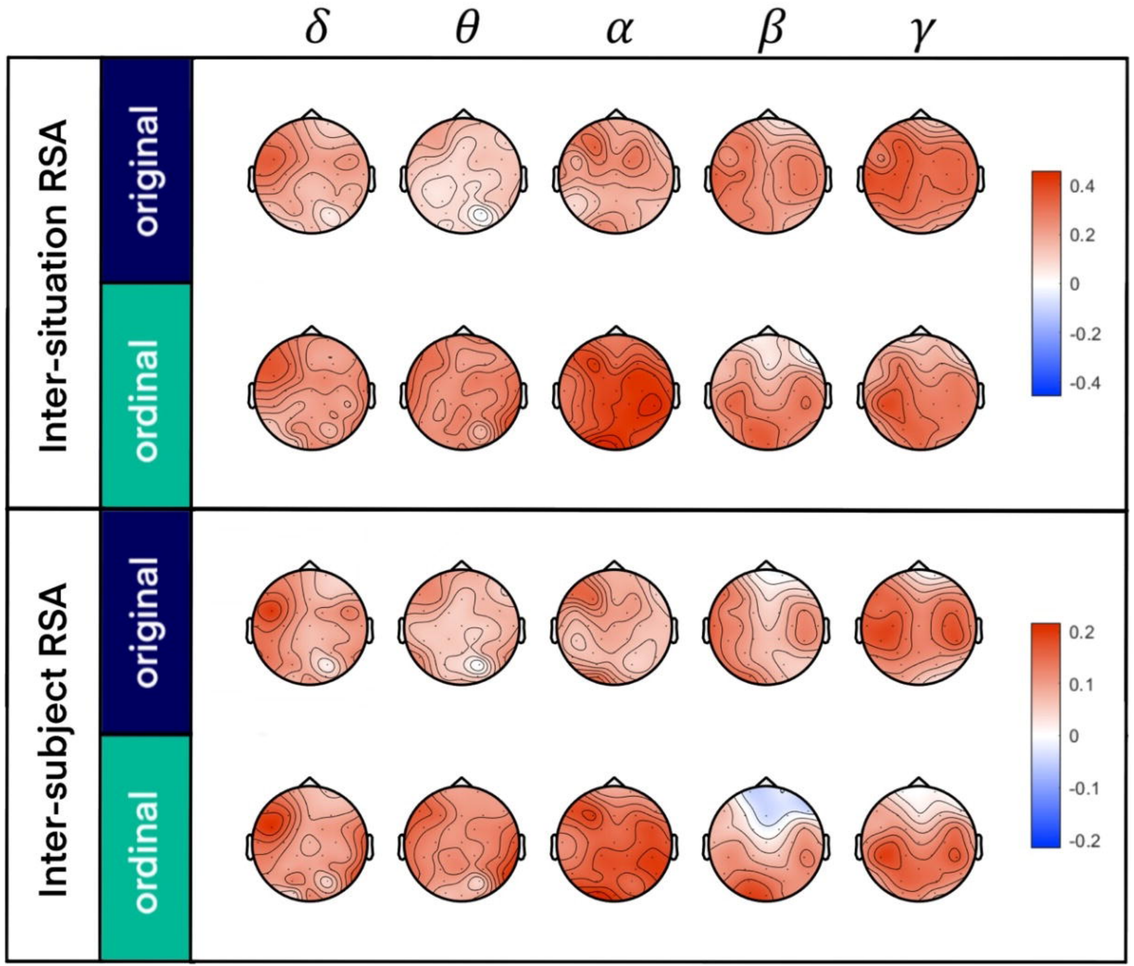
(a) Neural representation of multivariate emotion, DE



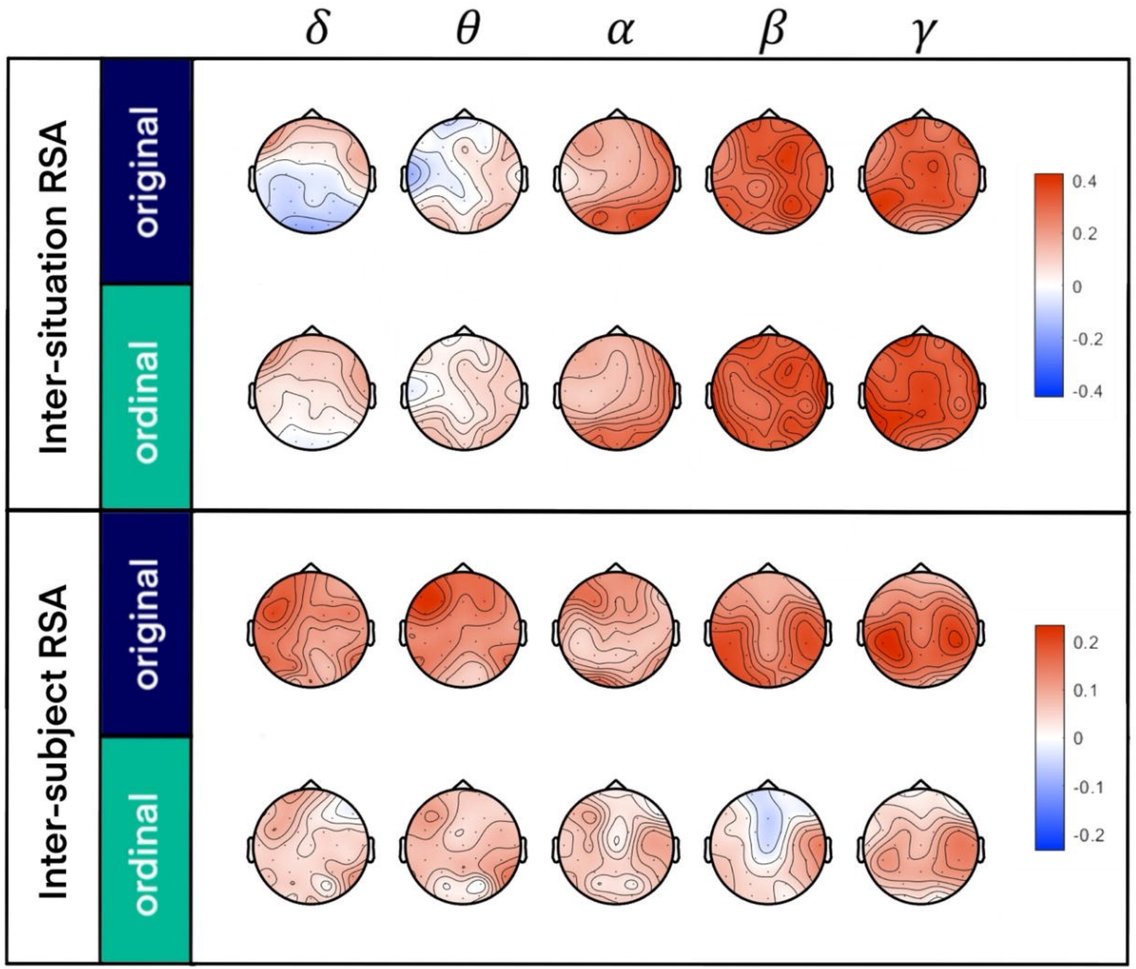
(b) Neural representation of multivariate emotion, PSD

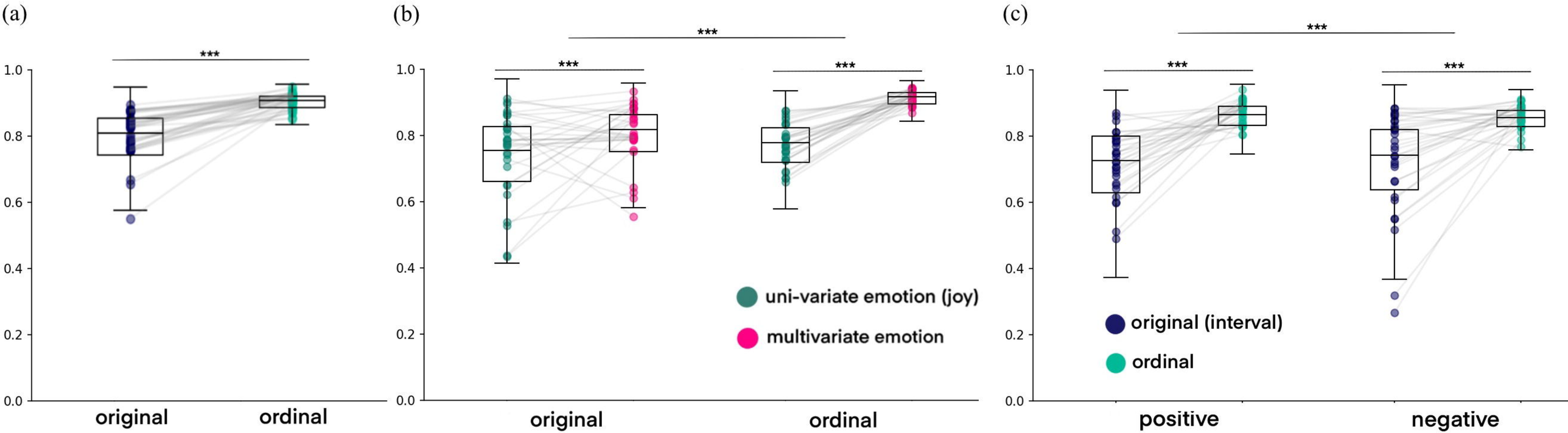


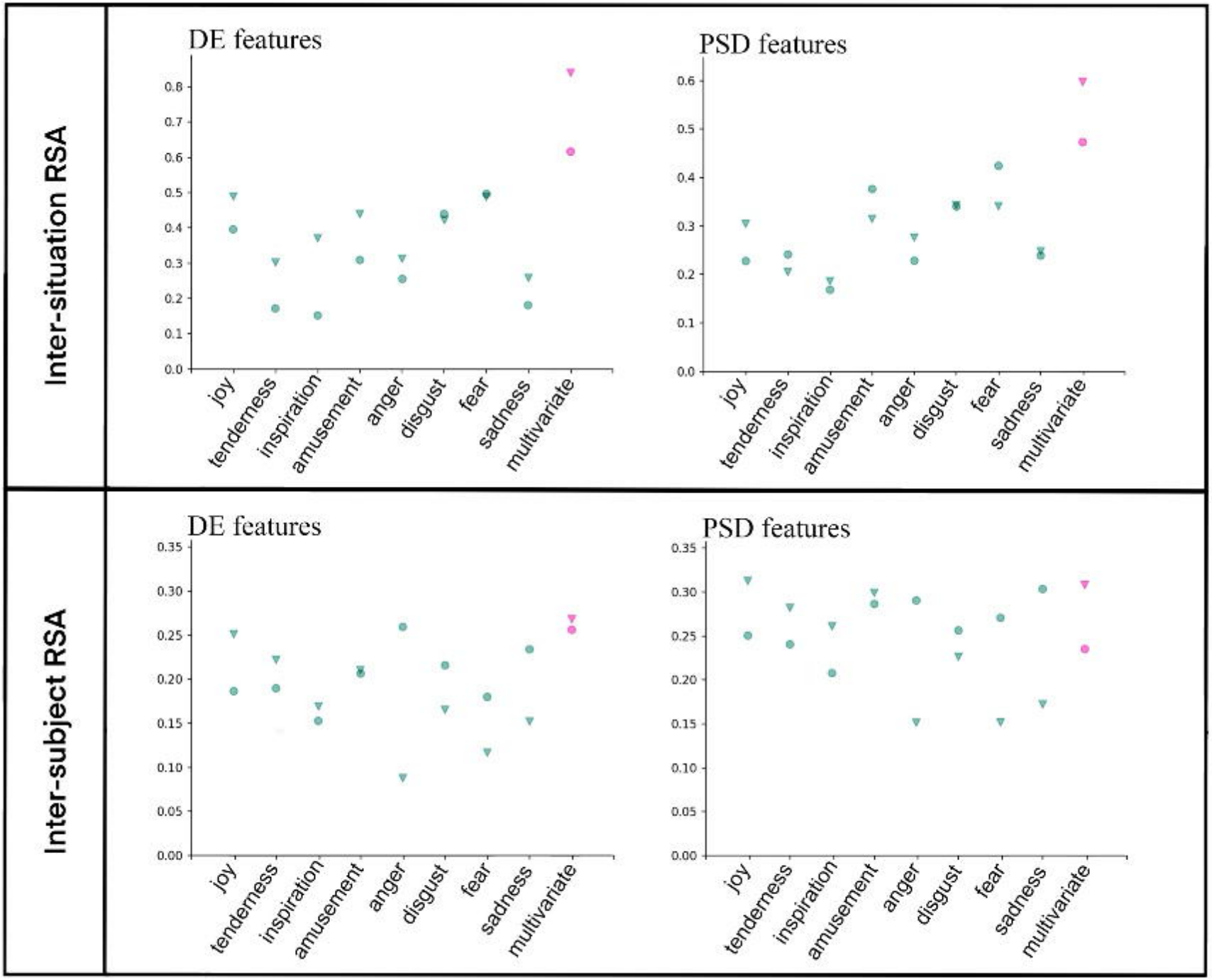
(a) Neural representation of positive emotion, DE



(b) Neural representation of negative emotion, DE




(a)
***
1.0
0.8
0.6
0.4
0.2
0.0
original
ordinal
(b)
***
***
***
1.0
0.8
0.6
0.4
0.2
0.0
uni-variate emotion (joy)
multivariate emotion
original
ordinal
(c)
***
***
***
1.0
0.8
0.6
0.4
0.2
0.0
original (interval)
ordinal
positive
negative


Inter-situation RSA
DE features
PSD features
Inter-subject RSA
DE features
PSD features
joy
tenderness
inspiration
amusement
anger
disgust
fear
sadness
multivariate
original
ordinal