OXFORD

Neural decoding dissociates perceptual grouping between proximity and similarity in visual perception

Lin Hua[1,2], Fei Gao[1], Chantat Leong[1,2], Zhen Yuan (iD)[1,2,*]

[1]Centre for Cognitive and Brain Sciences, N21 Research Building, University of Macau, Avenida da Universidade, Taipa, Macau SAR 999078, China,
[2]Faculty of Health Sciences, E12 Building, University of Macau, Avenida da Universidade, Taipa, Macau SAR 999078, China
*Corresponding author: Centre for Cognitive and Brain Sciences and Faculty of Health Sciences, E12 Building, University of Macau, Avenida da Universidade, Taipa, Macau SAR 999078, China. Email: zhenyuan@um.edu.mo

Unlike single grouping principle, cognitive neural mechanism underlying the dissociation across two or more grouping principles is still unclear. In this study, a dimotif lattice paradigm that can adjust the strength of one grouping principle was used to inspect *how*, *when*, and *where* the processing of two grouping principles (proximity and similarity) were carried out in human brain. Our psychophysical findings demonstrated that similarity grouping effect was enhanced with reduced proximity effect when the grouping cues of proximity and similarity were presented simultaneously. Meanwhile, EEG decoding was performed to reveal the specific cognitive patterns involved in each principle by using time-resolved MVPA. More importantly, the onsets of dissociation between 2 grouping principles coincided within 3 time windows: the early-stage proximity-defined local visual element arrangement in middle occipital cortex, the middle-stage processing for feature selection modulating low-level visual cortex such as inferior occipital cortex and fusiform cortex, and the high-level cognitive integration to make decisions for specific grouping preference in the parietal areas. In addition, it was discovered that the brain responses were highly correlated with behavioral grouping. Therefore, our study provides direct evidence for a link between the human perceptual space of grouping decision-making and neural space of brain activation patterns.

Key words: EEG; MVPA; perceptual grouping; visual perception.

Introduction

Perceptual grouping or organization depicts the process on how the human brain aggregates meaningless local sensory elements into meaningful global patterns. Interestingly, perceptual grouping is based on the cognitive mechanism associated with various principles, such as proximity, similarity, good continuation, closure, and common fate of visually ordered elements in our real-world experience (Wertheimer 1922, 1923). In particular, proximity and similarity as the earliest identified principles are recognized as the most fundamental processes of perceptual grouping (Wagemans et al. 2012). Previous studies have inspected the dynamic signatures and anatomical distributions associated with single grouping principle in vision science. More importantly, neuroimaging studies demonstrated that compared with the ungrouped local elements in space, the single perceptual grouping principle implicates an early visual processing (Nikolaev et al. 2008). Meanwhile, it was discovered that multiple brain activation regions were involved in the processes of single grouping principle, including areas V1 (Wannig et al. 2011; Stoll et al. 2020), V2 (Merigan et al. 1993), the lateral occipital complex (Fang et al. 2008; Murray et al. 2004) of the visual cortex, the middle temporal cortex, the inferior parietal cortex, and the prefrontal cortex (Seymour et al. 2008; Carther-Krone et al. 2020).

To date, the underlying cognitive neural mechanism is yet unclear when human beings simultaneously process 2 or more principles of perceptual grouping information in a natural environment. Existing behavioral studies illustrated that grouping effects were enhanced by the cooperation of 2 grouping principles, whereas decreased for competing case (Quinlan and Wilton 1998; Luna and Montoro 2011). However, those behavioral findings were mostly based on the descriptions of conscious experience regarding the units that people naturally perceive rather than manipulating these grouping principles in fine-grained psychophysical settings. Therefore, it is very hard to quantify the effects of 2 or more grouping principles or detect the specific principle in operation by using behavior data.

Interestingly, to reveal the neural correlates of differing grouping principles, EEG studies based on univariate analysis have been performed to differentiate the 2 grouping principles generated by independent stimuli. For example, previous Electroencephalography (EEG) studies (Han 2004; Han, Jiang, Mao, Humphreys and Gu 2005) demonstrated that grouping by proximity, relative to shape similarity, was consistently associated with enhanced positive event-related potential (ERP) component peaking around 100 ms after stimuli onset over occipital electrodes (P100) and around 300 ms upon central and parietal electrodes (P300). In addition, it

was discovered that grouping by similarity relative to proximity elicited a higher temporo-occipital negative component N200. In contrast, EEG findings (Luna et al. 2016; Villalba-García et al. 2018) also exhibited a null P100 effect to show the difference between proximity and similarity grouping. This discrepancy in various ERP studies might be attributed to the limitation of the used phenomenological paradigm and univariate analysis approach. In particular, only ERP components showing significant difference between proximity and similarity grouping were examined for these work, whereas ERP signals that might be able to distinguish different perceptual grouping states yet failed to reach a significance were discarded (Cichy and Pantazis 2017).

To bridge the theoretical and methodological gap, the present study aims at directly quantifying the relationship between the 2 basic grouping principles (proximity and similarity) by using a dimotif lattice paradigm. Therefore, we hypothesize that the grouping effects of two principles might be precisely quantified when manipulating the strength of one grouping effect and controlling the other. In addition, it is also assumed that the dissociative processing between the 2 grouping principles would be realized in various time windows, implicating differing stages in both temporal and spatial domains. To test the hypothesis, both psychophysical and EEG data were recorded to inspect the precise values of different perception between the proximity and similarity grouping. More specifically, to resolve the limitation of EEG univariate analysis, novel time-resolved multivariate pattern analysis (MVPA) was carried out to dissociate well the 2 grouping principles. This analysis method is able to generate the interactions between multiple channels/trials, so as to detect the subtle changes in brain activation patterns associated with proximity and similarity groupings. Additionally, source estimation was also conducted to inspect how the neural underpinnings underlying 2 grouping principles were integrated in spatial domain. It is expected that this study will be able to pave a new avenue for fully understanding the mechanisms associated with human perceptual space of grouping decision-making and neural space of these brain response patterns.

Materials and methods
Participants

Thirty college students (19 males, mean age: 21.4 ± 2.8 years) from the University of Macau participated in this experiment. All participants reported no histories of neurological illness or mental disorders and were right-handed with normal or corrected-to-normal vision. Four participants (3 males and 1 female) were excluded for further analysis due to not following the task instructions (pressing the same button throughout the whole test) or the large measurement noise (over 10% EEG artifacts). Informed written consent was obtained from each participant prior to the experiment. The

protocol for the present study was approved by the Institutional Review Board of the University of Macau.
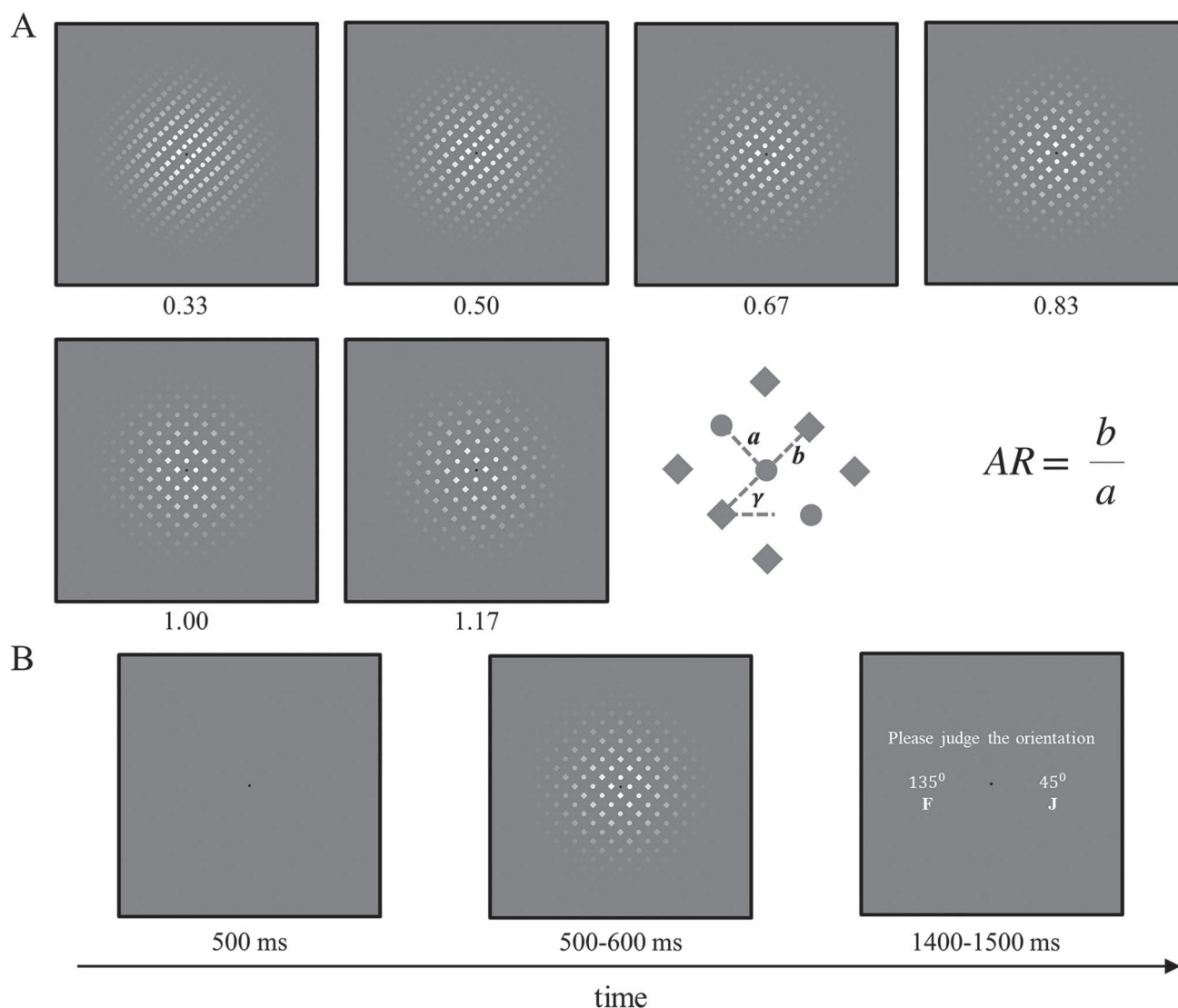
Dimotif lattice stimuli

Two categories of motifs (i.e. squares and discs, Kubovy and Van Den Berg 2008) were used as visual stimuli during the dimotif lattice task (Grünbaum and Shephard 1987), which were able to quantify the interaction between proximity and similarity grouping effects. The arrangement of all motifs was constrained by 3 parameters (a, b, and γ), as demonstrated in Fig. 1A. Specifically, a and b denote the distances of neighboring motifs along the 2 main axes, which are perpendicular to each other. All elements in either string of direction a consist of the identical motifs, while adjacent strings include heterogeneous motifs. Yet, each element string in direction b consists of alternating occurrences of the 2 motifs, whose pattern remains the same across all strings. The last parameter γ represents the angle between b axis and the horizontal line (measured counterclockwise).

The orientation of the dimotif lattice patterns was defined by the aspect ratio ($AR = \frac{b}{a}$) and γ, given a fixed a value. For the present design, a and γ were set to 1.5° (visual angle) and 45° (stimuli orientation), respectively, whereas AR was respectively chosen from 0.33, 0.50, 0.67, 0.83, 1.00, and 1.17, thus forming 6 dimotif lattice patterns. The AR adjustment and local element difference (squares and discs) were able to result in minor changes in dot density and local surface difference, which further caused the global luminance difference among each AR. As a consequence, the maximum difference of average luminance between AR = 1.17 and AR = 0.33 was by a factor of 2.25. Interestingly, it has been shown that when the luminance was up to a factor of 40, increased luminance of the stimulus can affect the amplitudes of ERP components (Johannes et al. 1995). Additionally, previous study using dot lattice paradigm explored the neural mechanisms of proximity grouping, which also found that no significant effect was driven by the small changes in lattice luminance (Nikolaev et al. 2008). Therefore, these evidence could justify that the significant results in current study corresponding to the manipulation of aspect ratio were not caused by the minor luminance difference in demotif lattice patterns.

To access participants' grouping preference and eliminate participants' bias to see the orientation of stimuli, half trials in one condition of perceived orientation of 45° might be seen as a proximity, while 135° as the similarity grouping preference. And the other half trials in one condition of perceived orientation of 135° might be seen as a proximity, while 45° as the similarity grouping preference.

Procedures

Participants were seated 55 cm from the display with a maximum visual angle of 25 × 25 degrees. Visual stimuli were generated by a Dell 64 bit-based machine (12G RAM) with an AMD Radeon HD graphics card

Fig. 1. Experimental design and procedure. A) Visual stimuli consist of 6 types of dimotif lattice pattern with a coherent (45° or 135°) or ambiguous global orientation, as a result of altered AR values (0.33, 0.50, 0.67, 0.83, 1.00, and 1.17). The calculation formula and diagram of AR were presented in middle right. B) Timing and sequence of stimuli in an experimental trial. Participants were instructed to press "F" key with the left index finger for a 135° global perceived orientation and to press "J" key with the right index finger for a 45° perception.

running Psychtoolbox-3 software (www.psychtoolbox. org) on Windows Professional 7. All motifs (squares and discs) were presented with a diameter of 1° against the gray background on a Dell P2312H monitor with a resolution of 1024×768. The element-background contrast was 150% (Weber contrast, $c = (I - I_b)/I_b$, in which I and I_b were the luminance of the motifs and the background, respectively). The whole pattern was masked with a Gaussian mask (20° diameter), whose center was a black point as fixation (0.5° diameter). The point fixation was kept on the screen for the whole test and participants were asked to fixate at it. For each trial, a blank with the point fixation was presented for 500 ms, followed by a dimotif lattice pattern lasting for 500 ms with a jitter of 100 ms to decrease expectancy effects. After the dimotif lattice disappeared, participants would encounter a probe asking them to judge the orientation of the previous pattern within 1400–1500 ms

(Fig. 1B). Participants were instructed to perform a two-alternative forced choice (2AFC) for two orientations (45° or 135°) by pressing the corresponding buttons labeled in the keyboard. Each of the 6 dimotif lattice patterns was presented for 80 times, thus making 480 trials in a random order. Each participant completed three blocks with 160 trials for each block.

Psychophysical analysis

Psychophysical performance was quantified as the percentage of dimotif lattice pattern with a response of 135° (i.e. similarity grouping preference). Psychometric functions were generated by fitting a cumulative Gaussian sigmoid curve by using the Psignifit toolbox (version 4, https://github.com/wichmann-lab/psignifit/wiki), which implements the maximum-likelihood method for parameter estimation and finds confidence intervals by bias-corrected and accelerated bootstrap method

(Schütt et al. 2016). The point of subjective equality (PSE) for each participant was extracted from the horizontal axis value corresponding to 0.5 on the vertical axis of psychometric curve. The statistical visualization of PSEs across participants was conducted by using the boxplot function in MATLAB (Mathworks, Natick, MA, USA).

EEG recordings and preprocessing

Participants were seated in a sound-attenuated and electrostatically shielded room during the experiment. Continuous EEG data were recorded from a 64-channel Biosemi Active Two EEG amplifier system (Biosemi, Amsterdam, Netherlands) with Ag/AgCl scalp electrodes placed according to the international 10–20 system on an elastic cap. During the online acquisition, EEG data were sampled at 2,048 Hz with a bandpass filter of 0.01–200 Hz. The input impedance of all channels was kept below 5 kΩ.

The EEG data were processed offline by using custom-made scripts in MATLAB, the EEGLAB toolbox (Delorme and Makeig 2004), and sLORETA software (Pascual-Marqui 2002) for source analyses. After down-sampling the data to 500 Hz, a built-in fourth-order Butterworth band-pass filter was applied with cutoff frequencies between 0.15 and 40 Hz. Then, epochs lasting from 100 ms before the stimuli onset to 500 ms afterward were extracted, among which those with unique, non-stereotypic artifacts were discarded. Independent component analysis was subsequently performed, and components representing common ocular or cardiac artifacts were visually identified and removed for further analysis. Overall, less than 10% of all trials were rejected. Finally, data were re-referenced to the grand average of whole head. It should be pointed out there that the orientation effect generated by perceiving for 45°/135° proximity grouping and 135°/45° similarity grouping would affect further EEG analysis. Therefore, instead of separating into the 2 halves, the whole dataset of each condition was pooled for further analysis.

Global electric field analysis

Modulations in the strength of the electric field at the scalp were assessed by the global field power (GFP; Murray et al. 2008). GFP is calculated as the square root of the averaged squared voltage value recorded at each electrode, which can index the spatial standard deviation of the electric field at the scalp. Larger GFP value denotes stronger electric field. Differences in GFP waveform data were analyzed as a function of time when the data post stimulus onset were significantly different from the baseline among 6 dimotif lattice conditions. GFP amplitude was considered significant if it exceeded a 95% confidence interval for at least 20 ms consecutively relative to the baseline of 100-ms prestimulus. Subsequently, GFP peaks were determined from the averaged GFP waveform across participants.

Topographic modulations were identified using randomization statistics applied to global map dissimilarity

measures (GMD; Murray et al. 2008). GMD is calculated as the root mean square of the difference between strength-normalized vectors. Differences in GMD values were also analyzed as a function of time, using stimulus type as a within-subject factor. The statistics of GMD were validated by a topographic ANOVA with 5,000 permutations ($P < 0.05$). Notably, GMD is independent of field strength and a significant GMD is indicative for different neural generators across 6 dimotif lattice conditions. Significant results of GMD were corrected by a duration threshold (20 ms).

Multidimensional scaling and source estimation

Global electric field analysis demonstrated that the dissociative information of 2 perceptual grouping was available in 3 time windows, while the GFP peaks at 108, 236, and 310 ms corresponded to the latency of the P1/N1, P2, and N3 in the visual stimuli-evoked potential, respectively. To assess mean scalp field differences among 6 conditions and construct the representational distances associated with proximity and similarity grouping preferences in brain, multidimensional scaling (MDS) analysis was performed for each time window among all conditions. The similarities of mean scalp fields among conditions were firstly assessed by the covariance among these maps. Then, 2D space that optimally represented the entire matrix of covariances was spanned between the first 2 eigenvectors. Furthermore, mean scalp fields of each condition were projected as 2D coordinates and then tested by calculating Euclidean distance among 6 AR conditions.

The neural sources of P1/N1, P2, and N3 were reconstructed for the significant period of GMD around each latency by using the sLORETA software, which provides current density values of 6,239 voxels (5 × 5 × 5 mm resolution), modeled in Montreal Neurologic Institute average MRI brain (MNI152) (Mazziotta et al. 2001). This method used a Laplacian-weighted minimum norm algorithm with no priori assumption about a predefined number of activated brain regions, thus constituting a more open solution to the EEG inverse problem.

The group differences at each component were examined by voxel-by-voxel single *t*-test statistics. T-test thresholds were computed using the built-in program of the sLORETA software. Correction for multiple comparisons was conducted by using a randomization test of statistical nonparametric mapping (SnPM) with 5,000 randomizations.

Multivariate pattern analysis

MVPA reveals topographic weightings of EEG signals that maximally distinguish perceptual grouping states within a given time interval. Here, a linear classifier was used based on L2-regularized logistic regression (Fan et al. 2008) to detect the optimal projections of the sensor space for discriminating among 6 dimotif lattice patterns (one-vs.-all decoding) or between 2 dimotif lattice patterns (one-vs.-one decoding) at a specific time point

(Fig. 5a). This allowed the assessment of how and when the dissociation of perceptual information between proximity and similarity was available in the stimulus-locked EEG data. The timing of this availability was accessed by using time-resolved decoding, in light of previous visual MEG/EEG studies (Crouzet et al. 2015; Cauchoix et al. 2016).

The accuracy of the classifier (linear L2-regularized logistic regression) fed with the multi-electrode single-trial EEG signals was evaluated for each time point independently. To verify the brain areas obtained from source estimation and ensure the electrodes mostly contributing to perception grouping, all electrodes were enrolled for one-versus-all decoding. And then, based on the findings of one-versus-all decoding and source estimation, 11 electrodes (posterior electrodes including P1, Pz, P2, PO7, PO3, POz, PO4, PO8, O1, OZ, and O2) were chosen for one-versus-one decoding, which was able to increase the signal-to-noise ratio and improve decoding performance. For each time point, the performance of the classifier was determined by using a Monte-Carlo cross-validation (CV) procedure ($n = 100$), in which the entire data were randomly partitioned into 10 portions including a training set (90% of the trials) and a test set (the remaining 10%). Here, the cost parameter C was used the default value of 1 for all analysis. These time windows were centered on and shifted from −100 to 500 ms relative to stimulus onset on stimulus-locked data.

For each participant, decoding accuracy was approximated according to the averaged performance across CVs. Error bars in the analysis corresponded to the nonparametric 95% confidence intervals of the mean obtained via bootstrapping. For each time point and CV, a measure of chance performance was obtained by performing an identical classification analysis using randomly permuted labels. At the single-subject level, classification accuracy was considered above chance when it was higher than classification accuracy obtained from permuted labels (paired *t*-test, $\alpha = 0.05$). The group analysis was performed following the same procedure, except that the group averages were computed across single-subject averages. Correction for multiple comparisons was used for a time-cluster-based approach, in which a time point was considered significant only when it was a member of a cluster of at least 10 consecutively significant time points (i.e. 20 ms).

Representational similarity analysis

Representational similarity analysis (RSA; Kriegeskorte et al. 2008) was carried out based on the multiclass decoding results. For multiclass decoding, classifiers were trained to discriminate among 6 ARs of dimotif lattice patterns, whose results were similar to the participants' psychophysical performance. This analysis constructed a dynamic neural representational similarity matrix (RSM) for each subject and time point. For each time point, there was a 6×6 confusion matrix, where each cell represented the proportion of trials.

Given one cell from the matrix, dimotif lattice pattern X was presented and the classifier categorized the trial as dimotif lattice pattern Y. These matrices denoted the representation of neural space, in which dimotif lattice pattern categories were coded, since they were able to demonstrate which dimotif lattice pattern-evoked neural patterns were similar. To construct the behavioral RSM, the percentage results of dimotif lattice stimuli were compared to each other and converted to Euclidean distance metric. Then, the representational similarity between dimotif lattice patterns was able to be compared by calculating Pearson correlations. The behavioral RSM was symmetrical along its diagonal, and the off-diagonal areas indicated the representational similarity across 6 dimotif lattice patterns. In particular, this behavioral RSM could estimate the participants' perceptual space to proximity or similarity grouping preference. Furthermore, the Pearson correlation coefficients were computed for each participant and each time point separately across the 36 cells of the neural and behavioral RSM. This correlation would therefore provide an estimation of the representational similarity of perceptual grouping in perceptual space and neural space.

Results
Psychophysical results
Psychophysical performance was measured as the percentage of trials perceived as 135° (i.e. similarity preference). The averaged percentage of behavioral responses across all participants for each dimotif lattice pattern (ARs from 0.33 to 1.17) was $7.62 \pm 1.28\%$, $32.52 \pm 2.67\%$, $53.75 \pm 3.36\%$, $83.07 \pm 2.12\%$, $91.97 \pm 1.69\%$, and $95.88 \pm 1.01\%$, respectively. Fitting the behavioral data using psychometric functions, it was found that as the AR increases, the percentage of similarity preference showed an S-curve growth tendency (Fig. 2A). This finding illustrated that participants were inclined to group discrete motifs into parallel strings in the direction *a* when the distance between 2 motifs increased in the direction *b*. The experimentally obtained PSEs were averaged across all participants, yielding an AR of 0.601 ± 0.011 (Fig. 2B). This result suggested that subjectively equated proximity effect of 2 perceptual grouping stimuli does not guarantee similar effects in the visual system due to the modulation of similarity effect.

Global electric field results
The neural dissociative processing of grouping preference between proximity and similarity evoked a significant electrophysiological response starting from 84 ms (GFP) relative to the prestimulus period (Fig. 3A). Accordingly, the GFP showed that the earliest signal increase for each dimotif lattice pattern (AR from 0.33 to 1.17) was 84, 90, 84, 86, 90, and 84 ms, respectively. Three GFP peaks were identified among 6 dimotif lattice conditions at 108,
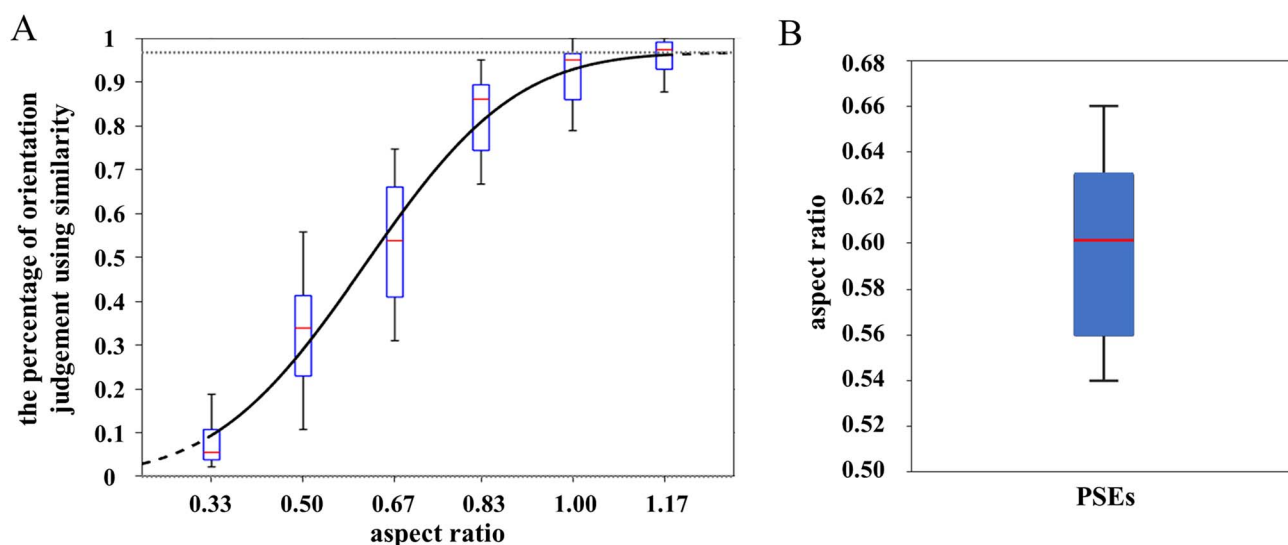
Fig. 2. The fitted psychometric functions and averaged PSEs of all participants. A) Illustrations of behavioral response between proximity and similarity grouping preferences among 6 dimotif lattice patterns. The cumulative Gaussian sigmoid curve was used to the fitted behavioral data with 95% confidence intervals of individual participant (dotted line). The S-curve was averaged across all participants. B) Boxplot with whiskers representing PSEs. The box extends from the 25th to 75th percentiles, while the whiskers extend from minimum to maximum values. The horizontal red line indicates the mean of PSEs ($n = 26$). The PSEs indicate that the AR for participants has the same global perceived orientation for proximity and similarity grouping.

235, and 310 ms, corresponding to the visual stimulus-evoked potentials P1/N1, P2, and N3, respectively. The significant differences in topographical distribution of the electric field independent of the electric field strength (GMD; $Ps < 0.05$ lasting for 20 ms) revealed that the underlying neural generators varied between dimotif lattice stimuli at the following 3 time windows. The first stage started from 86 to 130 ms after stimulus onset with a negative potential (N1) in medial occipital cortex and a positive potential (P1) over bilateral occipital cortex. The second time window implicated a positive distribution in medial occipital cortex (P2) from 188 and 238 ms. The late time window from 254 to 386 ms with a negative potential (N3) in frontal and parietal cortex corresponded to a higher order cognitive processing (Fig. 3B). Meanwhile, in light of GMD results, topographies between 6 dimotif lattice stimuli showed a similar pattern regarding the 3 stages of differentiating 6 AR conditions by proximity and similarity principles (86–130, 188–238, and 254–386 ms; Fig. 3C).

MDS and source estimation results

To identify the brain representational distances and activation patterns associated with proximity and similarity grouping preferences, we calculated the MDS and cortical generators of the stimulus-evoked responses for dimotif lattice patterns at 3 time windows (86–130, 188–238, and 254–386 ms). The MDS findings indicated that the initial representation of dimotif lattice patterns in participants' brain was arranged according to the AR values (86–130 ms). Then, based on the visual processing of elements' features and spatial locations, the representational distances of dimotif lattice patterns were rearranged. Especially, the ARs of 0.67 and 0.83 were transformed to those close to the representation

of similarity grouping preference (i.e. AR = 1.00/1.17) (188–238 ms). Finally, at the final stage of global perceived orientation decision-making (254–386 ms), the representational distances showed comparable patterns to behavioral responses (Fig. 4A and B). In addition, brain activation areas were detected that were associated with the dissociative processing of grouping preferences. First, the visual areas in cuneus and middle occipital gyrus [Brodmann areas (BA) 17, 18, and 19] were engaged in the early processing of spatial relationships between local elements with a time window of 86–130 ms. The attentional postperceptual operations employed the inferior occipital gyrus and fusiform gyrus (BA 17, 18, 19), indexing the discrimination and selection of stimuli features that were required for completing shape similarity identification from 188 to 238 ms. At the third stage (254–386 ms), the postcentral gyrus (BA 7) was linked to the discrimination in confidence for perceptual grouping decisions (Fig. 4C and Table 1).

Multiclass decoding and representational similarity results

While the global electric field analysis failed to sufficiently accommodate the discrimination in each dimotif lattice pattern, the dimotif lattice pattern might contain mixed proximity and similarity grouping preferences per se. Alternatively, time-resolved MVPA was employed to evaluate whether the single-trial, instantaneous topographical pattern of EEG activity carries the information about proximity and similarity grouping preferences. In the multiclass analysis, the grouping preference in proximity and similarity principle was also decoded at 3 time windows after stimulus onset at group level, which were significant above chance level: 98–146 ms ($t_{(25)} = 4.202$, $P < 0.001$, 95% CI [0.996, 2.912]), 164–190 ms ($t_{(25)} = 4.230$,

Table 1. Brain regions showing significant differences across 3 time windows between max proximity vs. max similarity grouping preference.

Time windows	Brain areas	BAs	MNI coordinates			Voxel peak values
			X	Y	Z	
86–130 ms	Cuneus	17	20	−90	5	10.17
	Middle occipital gyrus	18	25	−90	10	10.75
		19	30	−95	15	10.38
188–238 ms	Inferior occipital gyrus	17	−20	−95	−15	7.93
	Fusiform gyrus	18	−25	−95	−20	8.19
		19	−25	−85	−20	7.41
254–386 ms	Postcentral gyrus	7	15	−65	65	7.45

Note. All brain regions were manifested with a threshold of $P < 0.05$.

$P < 0.001$, 95% CI [0.763, 2.211]), and 248–500 ms ($t_{(25)} = 6.710$, $P < 0.001$, 95% CI [1.354, 2.553]). Specifically, those statistics were generated from averaged classification accuracies between normal label and random label during the respective time windows and following-up paired *t*-test comparisons. The results suggested a 3-stage process of early, middle, and late perceptual grouping in the human brain, which therefore provided sufficient and accurate time-resolved information for discriminating between grouping preferences in proximity and similarity (Fig. 5B).

To further quantify the resemblance between behavioral performance and brain responses, we explored the representational similarity between subjective and neural perceptual grouping representations. Since the multiclass decoding results were consistent with behavioral patterns, the dynamic association between participants' perceptual space and neural space could be obtained. Correlational analyses revealed that behavioral and neural integration were significantly correlated at 3-time windows: 96–146 ms ($t_{(25)} = 8.734$, $P < 0.001$, 95% CI [0.209, 0.338]), 150–198 ms ($t_{(25)} = 5.057$, $P < 0.001$, 95% CI [0.109, 0.259]), and 212–500 ms ($t_{(25)} = 6.458$, $P < 0.001$, 95% CI [0.159, 0.309]) (Fig. 5C). This pattern was comparable to the 3 stages at which the neuronal signal carried information about grouping preference as indicated by the multiclass analysis (Fig. 5B). This line of findings confirmed that the information used by the decoding classifiers formed the basis of perceptual grouping decision-making for each participant.

One-versus-one decoding results

To further examine how the global perceived orientation varies across different stimuli patterns, we performed one-versus-one classification for all paired dimotif lattice patterns. As shown in Fig. 6, pairwise classifications yielded significant accuracies above chance level at certain time points across the trials ($Ps < 0.05$ lasting for >20 ms). Interestingly, the paired patterns between proximity and similarity grouping indicated that as the AR values increased, more significant components were discriminated within the 3 time windows (Fig. 6 bule curves for 0.33 vs. 0.67/0.83/1.00/1.17, and pink curves for 0.50 vs. 0.67/0.83/1.00/1.17), indicating that different brain

processing between proximity and similarity grouping might influence perceptual grouping decision. However, significant decoding performance was only found at the early time window (108–140 ms for decoding 0.67 vs. 1.00, 108–142 ms for decoding 0.67 vs. 1.17, 86–124 ms for decoding 0.83 vs. 1.17; Fig. 6 green and yellow curves) between each discontinued similarity grouping condition (Fig. 6 green and yellow curves). Yet, no significant effect was detected between the adjacent AR values ($Ps > 0.05$; Fig. 6).

Discussion

A whole picture of the dissociative processing between 2 perceptual grouping principles in humans was depicted in both perceptual and neural domains in the current study. First, psychophysical results revealed that as AR values increased, more global perceived judgments for similarity orientation were made. This pattern was in line with the findings from Kubovy and Van Den Berg (2008) and Wei et al. (2018), which showed that similarity grouping effect was enhanced with decreased proximity grouping effect when the 2 unstable grouping cues were presented simultaneously.

In addition, regarding the global electric field differences among 6-stimulus patterns, the time point at which perceptual grouping information was available coincided with the latency of the first visual-evoked response at around 90 ms. It suggested that perceptual grouping information might initially be represented by one of the most elementary features. In visual system, the initial event-related response is associated with the representation of visual features and a fast category-selective activity, when the visual system is highly optimized for the processing of natural scenes (Wang et al. 2012; Kaneshiro et al. 2015). Existing EEG studies have identified that the early component (P1) was only associated with proximity-defined grouping (Han 2004; Han, Jiang, Mao, Humphreys and Qin 2005). In light of the comparable patterns generated, our GFP results therefore extended this line of evidence by demonstrating that the human brain might also initially manifest a proximity preference with the presence of both grouping cues. Importantly, GMDs confirmed the
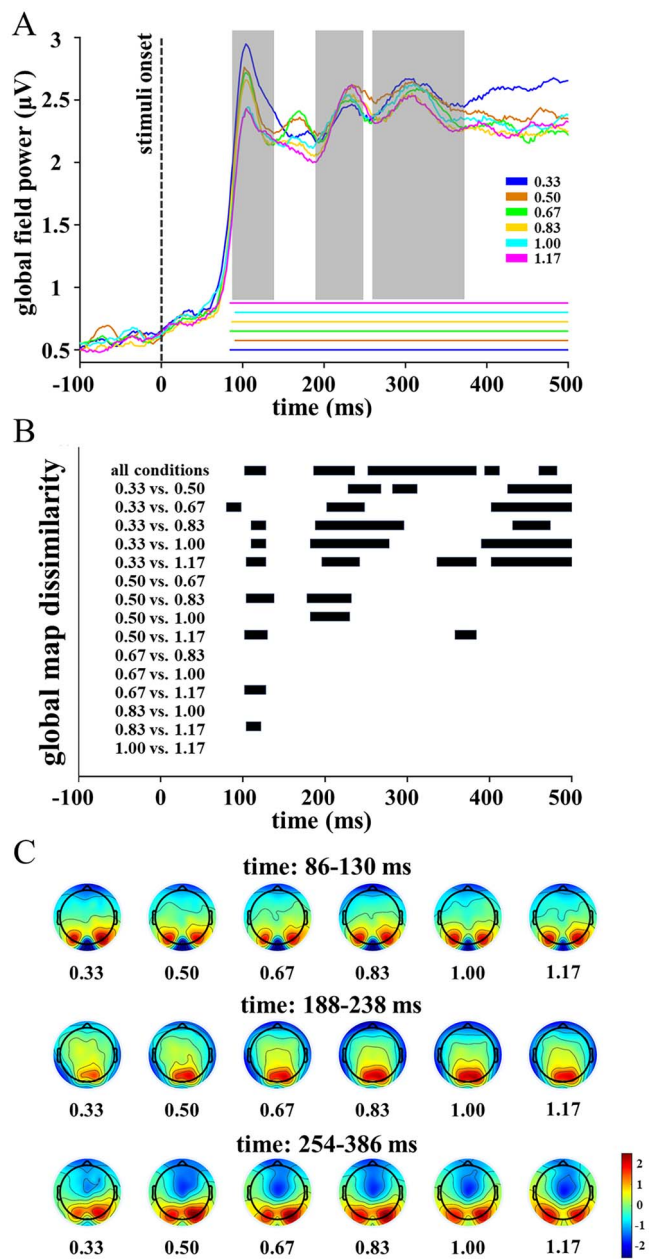
Fig. 3. Event-related responses to different ARs exhibited differences in electric field strength and topographic distribution. A) GFP for each AR pattern (averaged across participants). The time segments whose GFP significantly differed from pre-stimulus baseline were marked with colored straight lines for each AR pattern. Colored shaded areas were denoted for 95% confidence interval. B) The time segments of significant GMD across 6 AR patterns were indicated in black bars. The top row depicts the significant main effect of ARs, while the following rows represent all significant pairwise comparison results across 6 ARs. C) Topography among 6 AR patterns in 3 time windows (86–130, 188–238, and 254–386 ms). For each topography, EEG signals were averaged across time points within a specific time window. Color bar denotes the voltage value (μV).

hypothesis for 3 time windows (86–130, 188–238, and 254–386 ms), indexing the processing of distinguishing 2 perceptual grouping principles. The current findings further indicated that different subregions within the parieto-occipital and fronto-parietal networks are activated by proximity and similarity preference to

different extents across 3 time windows. It is noted that these findings were compatible with different proposals for the extensive network, which argued the frontal, parietal, and occipital cortex were commonly engaged in the processing of proximity and similarity grouping (Han and Humphreys 2007; Seymour et al. 2008).

Furthermore, MDS and source estimation provided a visualization of likely representational distance and predominant sources in the human brain for the 3-stage processing of grouping preference dissociation. The discrimination primarily involves the middle occipital cortex at an early stage (86–130 ms), the inferior occipital cortex and fusiform cortex at middle-stage (188–238 ms), and the parietal cortex at late stage (254–386 ms). These areas have been widely recognized to be associated with perceptual grouping processing in previous functional neuroimaging studies. The middle occipital cortex in cuneus, in particular, was linked to the proximity-defined larger scale organization of discrete visual elements in the visual field (Han, Jiang, Mao, Humphreys, Gu 2005a). Our results validated the previous findings by using MDS analysis, which showed that the pattern distances are arranged consistently with the strength of proximity effect in the early stage. Then, in the middle stage, a stronger grouping-related activation was observed in the inferior occipital and fusiform gyrus (BA 17, 18, and 19). This pattern was consistent with fMRI studies which specified the functional contribution of object recognition (Joseph 2001; Sim et al. 2015). Furthermore, MDS showed that the aspect ratios (AR = 0.67 and 0.83) were closer to the similarity set, which were initially closer to the proximity set. This might reflect a feedback mechanism from the lateral occipital cortex, the inferior occipital, and fusiform gyrus, to V1/V2 selecting the object features and processing the object recognition (Chen et al. 2021). Finally, the parietal cortex is considered as the area for spatial attention (Kim et al. 2017; Grassi et al. 2018) and more engagements were found during a feature integration task (Shafritz et al. 2002; Freedman and Ibos 2018). The current findings in late component therefore implicated a further distinction between 2 grouping information and final decision-making based on spatial attention and feature integration.

To identify the subtle temporal representation of the dissociative processing across 3 time windows, we trained classifiers on the data collected from different grouping preference conditions, respectively. The classifiers were then tested by looking at the decoding performance and neural template across all electrodes for each time point. Interestingly, the neural response patterns obtained from multiclass MVPA of EEG recordings corroborated results of global electric field. With an early onset and relatively short peak, the classification performance suggested that earlier sources of dissociative information came from the dynamic visual representations (Guo et al. 2019). For each grouping preference pattern, significant decoding performance between differing preferences could cover
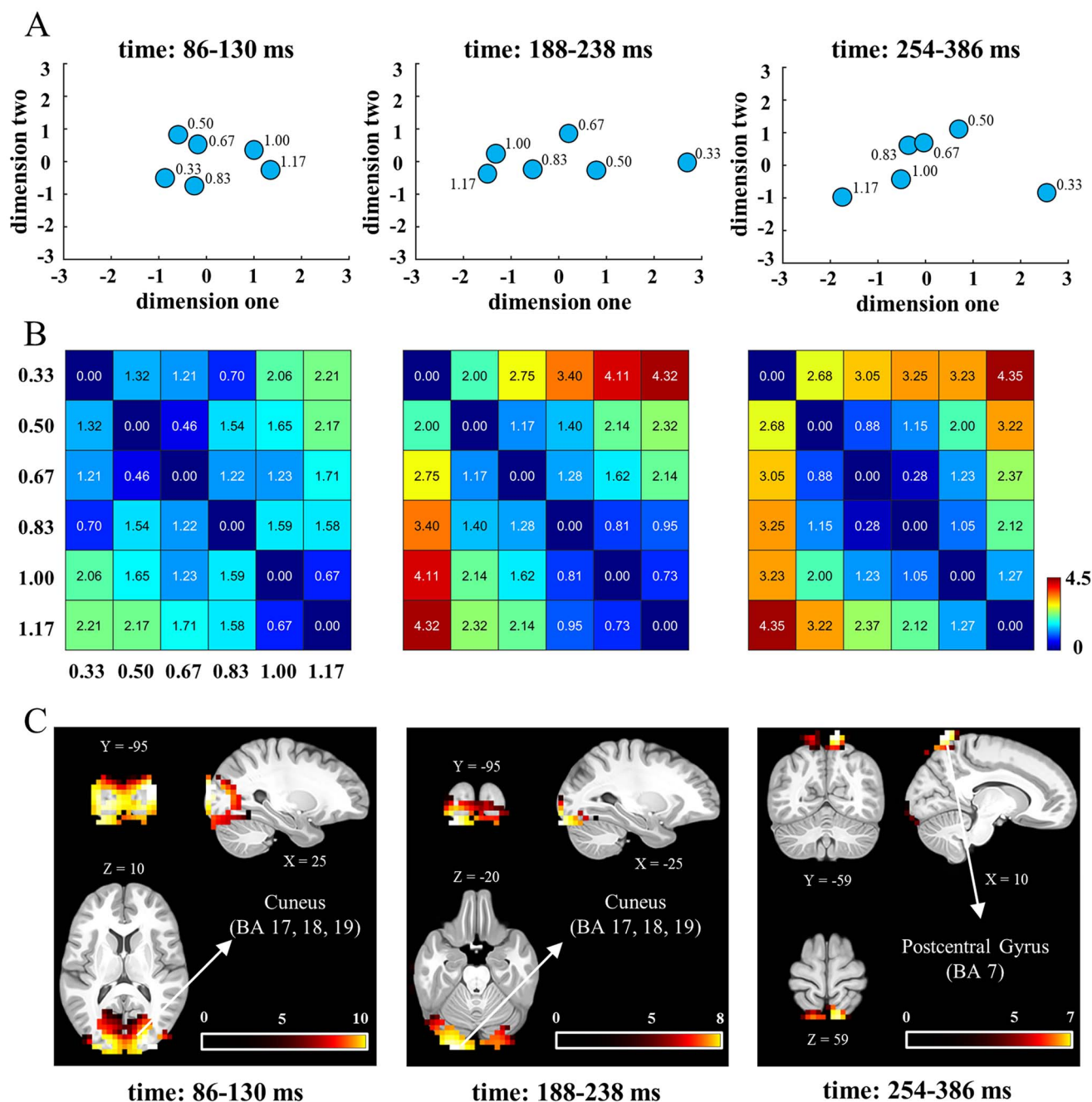
Fig. 4. The spatial layout of 6 ARs and EEG source estimation of proximity and similarity across 3 time windows. A) MDS provides a visual representation that projects the distances among 6 AR conditions to examine the representational distances. B) Calculated Euclidean distance among 6 AR conditions to offer statistical evidence of representational distances. C) Estimations of the neural sources in the dissociative processing of 2 grouping principles underlying the stimulus-evoked responses, including the cuneus and middle occipital gyrus (BA 17, 18, and 19; 86–130 ms), the inferior occipital gyrus and fusiform gyrus (BA 17, 18, and 19; 188–238 ms), and the postcentral gyrus (BA 7; 254–386 ms) ($P < 0.05$, single *t*-test).

more than one time segment. Interestingly, paired comparisons were pronounced in the early time window for the pairs of similarity grouping preference, while with increased AR values, discriminated components within the 3 time windows increased between the proximity and similarity pairs. For the present study, increased AR values were able to decrease the effect of proximity grouping. Therefore, for all similarity pairs, significant differences in the early time window indicated the difference effect of proximity. The pattern manifested in

the early time window could corroborate the notion that the distinct processing between 2 grouping principles initially employed proximity to group the ordered and discrete visual elements. Meanwhile, this line of finding reflected the proximity-defined large-scale organization (Han, Jiang, Mao, Humphreys and Gu 2005; Han, Jiang, Mao, Humphreys and Qin 2005), even in the paired conditions of similarity grouping preference. Importantly, the late time window (>260 ms) reported in previous EEG studies also elaborated the higher order cognitive
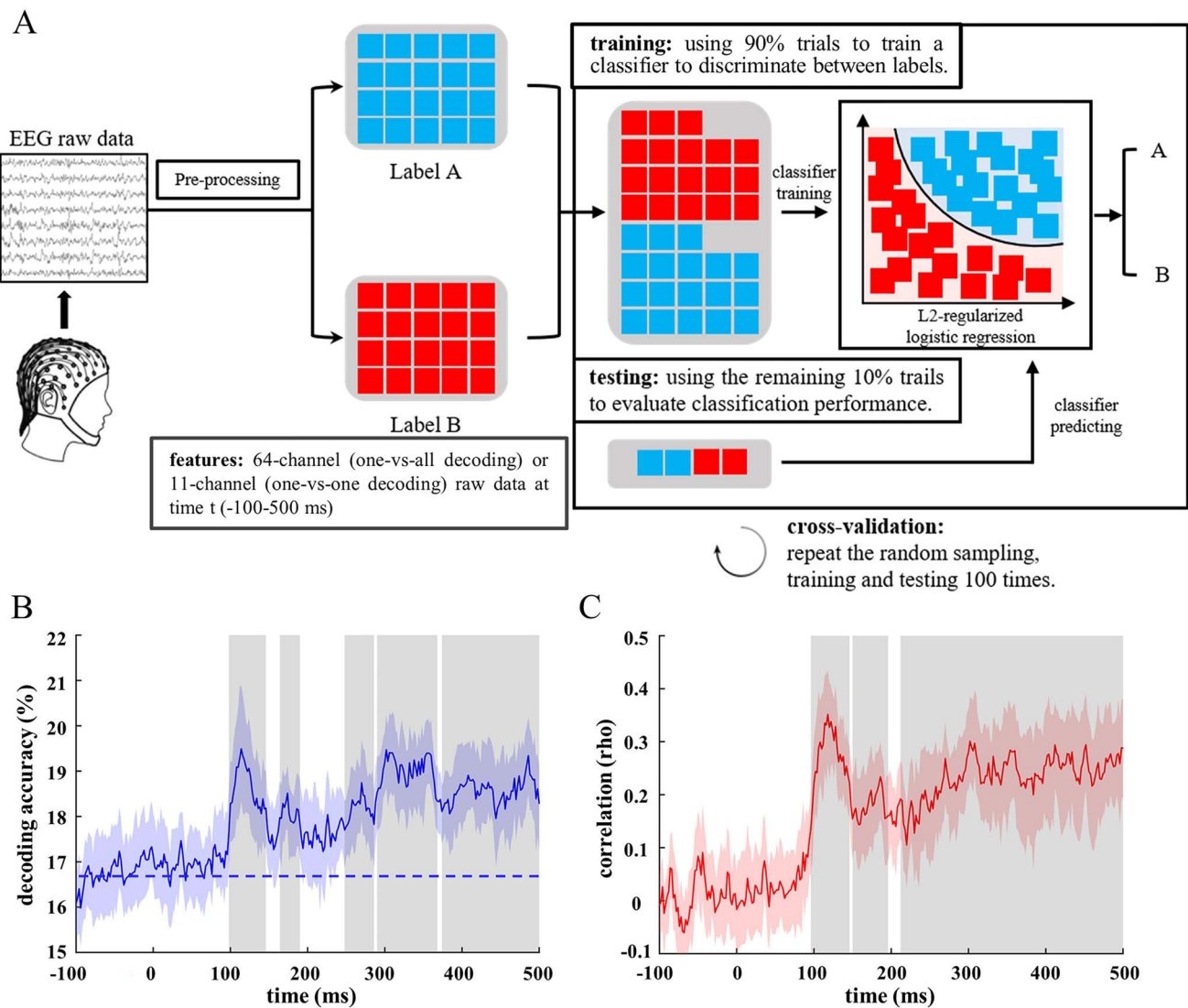
Fig. 5. The procedure of MVPA and the visualization of multi-class decoding and RSA results. A) For a given participant and time point, a random sample of 90% of the trials associated with 2 (one-vs-one classification) or 6 (multi-class decoding) conditions was used to train a classifier to discriminate between brain responses (EEG scalp topography) ARs. Classification performance was then evaluated by using the remaining 10% of the trials. The entire procedure was repeated for 100 CVs. B) Multi-class decoding: classifiers were trained at each time point among 6 ARs to reveal when ARs could be dissociated from each other. C) RSA: behavioral representational similarity matrices were correlated with neural representational similarity matrices at each time point to estimate the similarity between AR representations in perceptual space and neural space. Colored curves denote decoding accuracy across participants [colored shaded area: bootstrapped 95% confidence interval; light gray shaded area: the period of significant decoding at the group level (Ps < 0.05 lasting for >20 ms)].

function integrating perception to action (Deslandes et al. 2005; Takacs et al. 2020). The current findings further revealed that more cognitive resources were engaged to discriminate differing stimuli and determine global perceived orientation when distant AR values were compared.

Finally, in light of the correlational results, perceptual and neural space converged after first stimulus-evoked response (96–144 ms), demonstrating that the distinguished perceptual grouping information encoded in MVPA laid the basis for perceptual grouping decision-making. The more similarly any 2 dimotif lattice patterns were represented in MVPA, the more confused the participants might feel in the process of perceptual grouping decision-making. To our knowledge, our study is the first to show that the subjective discriminability of perceptual

grouping principles is related to their neural dissimilarity, thus manifesting a robust neural-perceptual mapping.

There are some caveats needing to be noted regarding the present study. For visual stimuli, comparable paradigms (demotif dot lattice) used in this study indicated that both proximity and similarity grouping preference simultaneously modulated participants' responses and their brain representation. Therefore, statistical analysis on one-versus-one perceptual patterns, instead of directly on different perceptual patterns for each grouping condition, was performed for quantifying the effects of 2 perceptual groupings and detecting the neural dissociation between proximity and similarity from each paired condition. Additionally, performing analysis on different perceptual patterns for each grouping condition might also lead to dataset
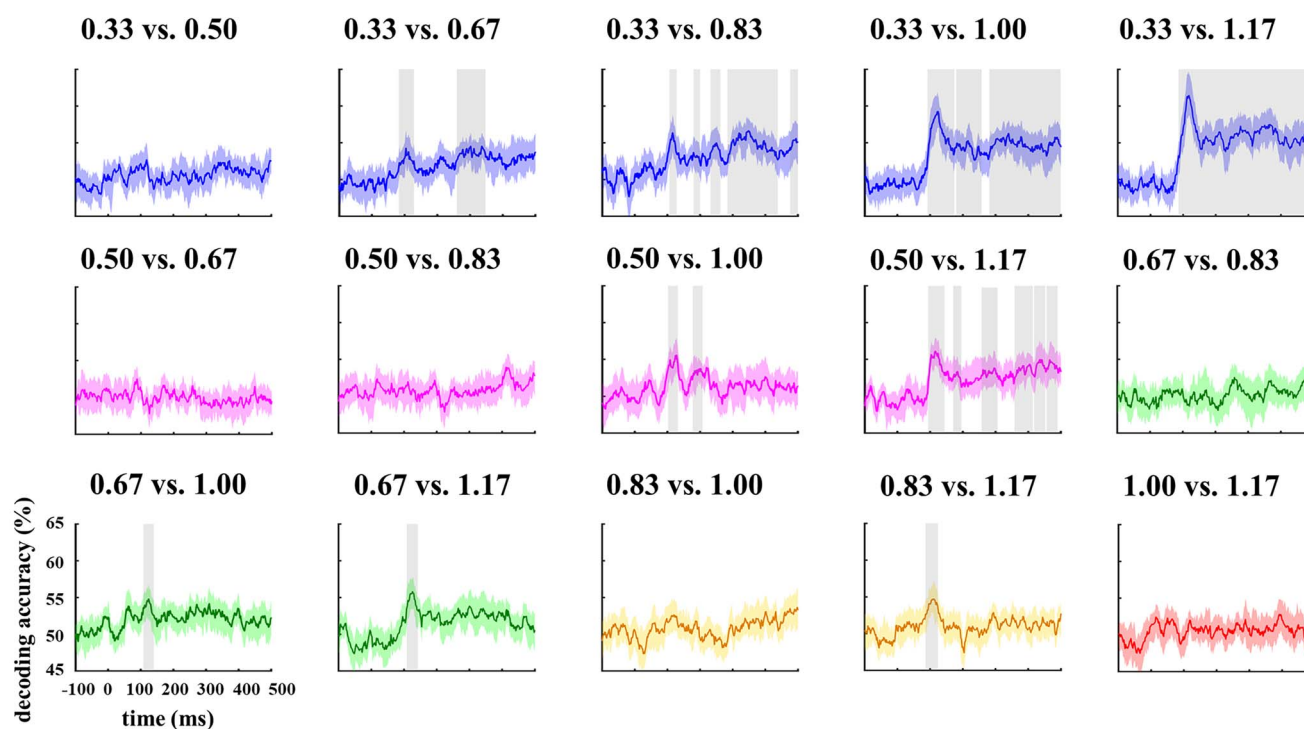
Fig. 6. One-versus-one decoding reveals that different ARs were represented by distinctive neural patterns. For one-versus-one decoding, classifiers were trained at each time point to provide the time points which were different between paired ARs. Colored curves indicate decoding accuracy averaged across participants [colored shaded area: bootstrapped 95% confidence interval; light gray shaded area: the period of significant decoding at the group level ($Ps < 0.05$ lasting for >20 ms)]. Various colors were employed to denote differing comparisons.

imbalance. For example, the dataset of proximity grouping preference includes more EEG data of small AR values, which is much larger than those in the data of similarity grouping, and vice reverse for EEG data of large AR values. Further study can use the paradigm such as Luna and Montoro (2011) to directly compare proximity and similarity grouping. Although these stimuli cannot manipulate grouping principles in fine-grained psychophysical settings like the current study, it may provide further evidence addressing when and where the perceptual grouping occurs. Furthermore, participants were instructed to focus on the fixation point during experiments; however, eye movements may still be a possible concern in current decoding study. Further study can focus on these caveats to examine whether they will perform as factors in perceptual grouping research.

In summary, the present study was among the first to apply MVPA to decode the processing of 2 perceptual grouping information, particularly with an emphasis on the representational similarity of perceptual space and neural space. Notably, the psychophysical responses reported here were in strong alignment with previous behavioral studies and we further quantified the perception gap between proximity and similarity principles. Our decoding results showed that the human visual system dissociating 2 perceptual grouping information is encoded at 3 stages of perceptual grouping. The grouping starts with the proximity-defined arrangement of local discrete visual elements in the middle occipital cortex, followed by the feature selection and spatial attention modulating 2 perceptual grouping cues in the inferior occipital cortex and fusiform cortex. At last, it ended with the higher cognitive integration for spatial information and feature conjunction to determine the decision-making in the parietal cortex. Thus, our findings offer fundamental insights into the mental separation of 2 perceptual grouping processing and its relation to ultimate perceptual grouping-related decisions. We anticipate that our methods and results will ignite further investigations using time-resolved whole-brain responses to understand other perceptual grouping principles, which could also contribute to other sensory perceptual grouping processing.

Funding

Conflict of interest statement: The authors declare no competing financial interests.

Data and materials availability

Data and code are available from the authors on request.

References

Carther-Krone TA, Lawrence-Dewar JM, Shomstein S, Nah JC, Collegio AJ, Marotta JJ. Neural correlates of perceptual grouping under conditions of inattention and divided attention. *Perception*. 2020:49(5):495–514.

Cauchoix M, Crouzet SM, Fize D, Serre T. Fast ventral stream neural activity enables rapid visual categorization. *NeuroImage*. 2016:125:280–290.

Chen S, Weidner R, Zeng H, Fink GR, Müller HJ, Conci M. Feedback from lateral occipital cortex to V1/V2 triggers object completion: Evidence from functional magnetic resonance imaging and dynamic causal modeling. *Hum Brain Mapp*. 2021:42(17): 5581–5594.

Cichy RM, Pantazis D. Multivariate pattern analysis of MEG and EEG: a comparison of representational structure in time and space. *NeuroImage*. 2017:158:441–454.

Crouzet SM, Busch NA, Ohla K. Taste quality decoding parallels taste sensations. *Curr Biol*. 2015:25(7):890–896.

Delorme A, Makeig S. EEGLAB-open source Matlab toolbox for electrophysiological research. *J Neurosci Methods*. 2004:134(1): 9–21.

Deslandes A, Veiga H, Cagy M, Piedade R, Pompeu F, Ribeiro P. Effects of caffeine on the electrophysiological, cognitive and motor responses of the central nervous system. *Braz J Med Biol Res*. 2005:38:1077–1086.

Fan R-E, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J. LIBLINEAR: a library for large linear classification. *J Mach Learn Res*. 2008:9:1871–1874.

Fang F, Kersten D, Murray SO. Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *J Vis*. 2008:8(7): 2–2.

Freedman DJ, Ibos G. An integrative framework for sensory, motor, and cognitive functions of the posterior parietal cortex. *Neuron*. 2018:97(6):1219–1234.

Grassi PR, Zaretskaya N, Bartels A. A generic mechanism for perceptual organization in the parietal cortex. *J Neurosci*. 2018:38(32):7158–7169.

Grünbaum B, Shephard GC. *Tilings and patterns*. New York: Courier Dover Publications; 1987

Guo LL, Nestor A, Nemrodov D, Frost A, Niemeier M. Multivariate analysis of electrophysiological signals reveals the temporal properties of visuomotor computations for precision grips. *J Neurosci*. 2019:39(48):9585–9597.

Han S. Interactions between proximity and similarity grouping: an event-related brain potential study in humans. *Neurosci Lett*. 2004:367(1):40–43.

Han S, Humphreys GW. The fronto-parietal network and top-down modulation of perceptual grouping. *Neurocase*. 2007:13(4):278–289.

Han S, Jiang Y, Mao L, Humphreys GW, Gu H. Attentional modulation of perceptual grouping in human visual cortex: functional MRI studies. *Hum Brain Mapp*. 2005:25(4):424–432.

Han S, Jiang Y, Mao L, Humphreys GW, Qin J. Attentional modulation of perceptual grouping in human visual cortex: ERP studies. *Hum Brain Mapp*. 2005:26(3):199–209.

Johannes S, Münte T, Heinze H, Mangun GR. Luminance and spatial attention effects on early visual processing. *Cogn Brain Res*. 1995:2(3):189–205.

Joseph JE. Functional neuroimaging studies of category specificity in object recognition: a critical review and meta-analysis. *Cogn Affect Behav Neurosci*. 2001:1(2):119–136.

Kaneshiro B, Perreau Guimaraes M, Kim H-S, Norcia AM, Suppes P. A representational similarity analysis of the dynamics of object processing using single-trial EEG classification. *PLoS One*. 2015:10(8):e0135697.

Kim Y-J, Tsai JJ, Ojemann J, Verghese P. Attention to multiple objects facilitates their integration in prefrontal and parietal cortex. *J Neurosci*. 2017:37(19):4942–4953.

Kriegeskorte N, Mur M, Bandettini PA. Representational similarity analysis-connecting the branches of systems neuroscience. *Front Syst Neurosci*. 2008:2:4.

Kubovy M, Van Den Berg M. The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns. *Psychol Rev*. 2008:115(1):131.

Luna D, Montoro PR. Interactions between intrinsic principles of similarity and proximity and extrinsic principle of common region in visual perception. *Perception*. 2011:40(12): 1467–1477.

Luna D, Villalba-García C, Montoro PR, Hinojosa JA. Dominance dynamics of competition between intrinsic and extrinsic grouping cues. *Acta Psychol*. 2016:170:146–154.

Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Woods R, Paus T, Simpson G, Pike B. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos Trans R Soc Lond Ser B Biol Sci*. 2001:356(1412):1293–1322.

Merigan WH, Nealey TA, Maunsell J. Visual effects of lesions of cortical area V2 in macaques. *J Neurosci*. 1993:13(7): 3180–3191.

Murray SO, Schrater P, Kersten D. Perceptual grouping and the interactions between visual cortical areas. *Neural Netw*. 2004: 17(5-6):695–705.

Murray MM, Brunet D, Michel CM. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr*. 2008:20(4):249–264.

Nikolaev AR, Gepshtein S, Kubovy M, Van Leeuwen C. Dissociation of early evoked cortical activity in perceptual grouping. *Exp Brain Res*. 2008:186(1):107–122.

Pascual-Marqui RD. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol*. 2002:24(Suppl D):5–12.

Quinlan PT, Wilton RN. Grouping by proximity or similarity? Competition between the Gestalt principles in vision. *Perception*. 1998:27(4):417–430.

Schütt HH, Harmeling S, Macke JH, Wichmann FA. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vis Res*. 2016:122:105–123.

Seymour K, Karnath H-O, Himmelbach M. Perceptual grouping in the human brain: common processing of different cues. *Neuroreport*. 2008:19(18):1769–1772.

Shafritz KM, Gore JC, Marois R. The role of the parietal cortex in visual feature binding. *Proc Natl Acad Sci*. 2002:99(16):10917–10922.

Sim E-J, Helbig HB, Graf M, Kiefer M. When action observation facilitates visual perception: activation in visuo-motor areas contributes to object recognition. *Cereb Cortex*. 2015:25(9):2907–2918.

Stoll S, Finlayson NJ, Schwarzkopf DS. Topographic signatures of global object perception in human visual cortex. *NeuroImage*. 2020:220:116926.

Takacs A, Zink N, Wolff N, Münchau A, Mückschel M, Beste C. Connecting EEG signal decomposition and response selection processes using the theory of event coding framework. *Hum Brain Mapp*. 2020:41(10):2862–2877.

Villalba-García C, Santaniello G, Luna D, Montoro P, Hinojosa J. Temporal brain dynamics of the competition between proximity and shape similarity grouping cues in vision. *Neuropsychologia.* 2018:121:88–97.

Wagemans J, Elder JH, Kubovy M, Palmer SE, Peterson MA, Singh M, von der Heydt R. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychol Bull.* 2012:138(6):1172.

Wang C, Xiong S, Hu X, Yao L, Zhang J. Combining features from ERP components in single-trial EEG for discriminating four-category visual objects. *J Neural Eng.* 2012:9(5):056013.

Wannig A, Stanisor L, Roelfsema PR. Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex. *Nat Neurosci.* 2011:14(10):1243–1244.

Wei N, Zhou T, Chen L. Objective measurement of gestalts: Quantifying grouping effect by tilt aftereffect. *Behav Res Methods.* 2018:50(3):963–971.

Wertheimer M. Untersuchungen zur Lehre von der Gestalt. *Psychol Forsch.* 1922:1(1):47–58.

Wertheimer M. Untersuchungen zur Lehre von der Gestalt II [Investigations on the Gestalt Theory II]. *Psychol Forsch.* 1923:4(1):301–350.