ELSEVIER

Interactions between proximity and similarity grouping: an event-related brain potential study in humans

Shihui Han [a,b,*]

[a] *Department of Psychology, Center for Brain and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China*
[b] *Learning & Cognition Lab, Capital Normal University, 105 Xisanhuan Bei Road, Beijing 100037, PR China*

Abstract

The current work examined neural substrates underlying interactions between two types of Gestalt grouping by recording event-related brain potentials (ERPs). Subjects were presented with stimulus arrays in which local circles and squares were grouped into columns or rows. The cues dominating grouping (proximity and similarity) were either congruent or incongruent. Subjects responded faster to the identification of orientations of perceptual groups based on proximity than similarity cues. Responses were slowed by incongruent cues and this effect was larger in the similarity than proximity conditions. Proximity grouping generated enhanced positivity over the posterior occipital cortex between 100 and 140 ms relative to similarity grouping. Relative to congruent grouping cues, incongruent cues elicited enlarged positivity over the temporal–parietal areas between 180 and 220 ms only in the similarity condition, but generated smaller P3 amplitudes with longer latencies in both proximity and similarity conditions. The results provide ERP indices of the dominance of proximity over shape similarity in guiding perceptual grouping when two grouping principles are present in stimulus displays.
© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Event-related potential; Perceptual grouping; Proximity; Similarity

Perceptual grouping takes place at an early stage of visual processing to form perceptual units for object recognition [9]. Gestalt principles play an important role in guiding perceptual grouping [11] but are not equally efficient. There is a stronger tendency to group local elements by proximity and common color than by similarity of shape [15], and response are faster to global structures formed by proximity grouping than by grouping by similarity of shapes [1,4–6].

Recent event-related brain potential (ERP) studies [3,8] found that proximity grouping is indexed by a positive activity at 100–120 ms after sensory stimulation over the medial occipital cortex and an occipito-parietal negativity with an onset of 180 ms. Grouping by shape similarity is reflected only in a long-latency occipito-temporal negativity with an onset of 260 ms. The ERP results suggest distinct neural substrates underlying grouping processes defined by different Gestalt laws. However, as only one Gestalt law determined the grouping operations in each stimulus display used in the prior work, it is unknown how ERPs are modulated by grouping operations when more than one grouping principle is at work and how the grouping processes interact with each other.

The current work further examined neural substrates underlying the interaction between two grouping operations defined by proximity and similarity. Given that proximity cues may be stronger or function earlier than similarity cues [1,4–6,15], it may be hypothesized that, when proximity and similarity cues are conflicting in determining local element grouping, proximity grouping may produce stronger or earlier interference over similarity grouping than vice versa. We designed stimulus arrays in which local circles and squares were grouped into columns or rows based on either proximity or similarity of shape (Fig. 1). In two of the stimulus arrays proximity and similarity cues are congruent in grouping the local elements into columns (Fig. 1a) or rows (Fig. 1b). In the other two stimulus arrays proximity and similarity cues are incongruent (Fig. 1c and d). For instance, proximity cues in Fig. 1c grouped local items into columns whereas similarity cues grouped them into rows. Subjects were instructed to identify columns versus rows in the stimulus arrays formed

* Tel.: +86 10 6275 9138; fax: +86 10 6276 1081.
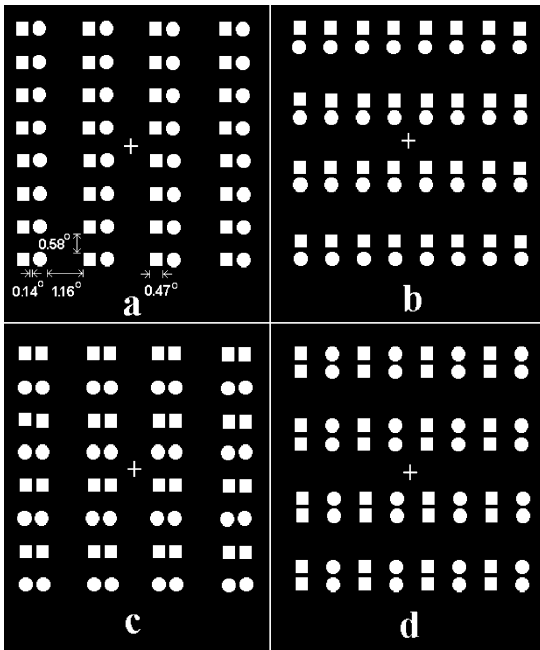*E-mail address:* shan@pku.edu.cn (S. Han).

Fig. 1. Stimulus arrays used in the present study. (a and b) Congruent stimuli; (c and d) incongruent stimuli.

by either proximity or similarity in separate blocks of trials while high density ERPs were recorded to the identification of orientations of perceptual groups. The difference between ERPs in the congruent and incongruent conditions revealed the time course and brain areas that mediate interactions between proximity- and similarity-grouping operations.

Fourteen adults (12 men, 2 women, aged between 18 and 35 years) were employed. All participants were right-handed, had normal or corrected-to-normal vision, and gave informed consent. The stimuli consisted of a square lattice of white elements (either filled circles or squares) in an 8×8 array on a black background (Fig. 1). The circles and squares were arranged into rows or columns by adjusting the distances between two adjacent rows or columns of local elements so that the distances between two near or remote rows (or columns) were $0.14°$ and $1.1°$, respectively. The local elements were also grouped into columns or rows by shape similarity cues. The proximity and similarity cues were congruent for half of the stimuli whereas incongruent for the others. Each local shape and stimulus array subtended an angle of $0.47° \times 0.47°$ and of $7.8° \times 7.8°$ at a viewing distance of 57 cm, respectively. The background and a local shape had a luminance of 0.02 and $3.46 \, cd/m^2$, respectively. The stimulus duration was 200 ms and interstimulus intervals varied between 800 and 1200 ms randomly. A white fixation cross of $0.3° \times 0.2°$ was continuously visible at the center of the screen.

Subjects identified orientations of perceptual groups formed by proximity or similarity in separate blocks of trials by pressing one of two keys with either the left or the right thumb. In the proximity-grouping condition, subjects responded vertical in displays of Fig. 1a and c but horizontal in displays of Fig. 1b and d. In the similarity-grouping condition, subjects responded vertical in displays of Fig. 1a and d but horizontal in displays of Fig. 1b and c. There were six blocks of 100 trials after 100 trials for practice in each grouping conditions. The electroencephalogram (EEG) was recorded from 120 scalp electrodes. The position of each electrode was measured with a 3D probe relative to fiducial marks on the skull. The average of the recordings from electrodes at the left and right earlobes was used as reference. Eye blinks and vertical eye movement were monitored with electrodes located below the left and right eyes. The horizontal electro-oculogram was recorded from electrodes placed 1.5 cm lateral to the left and right external canthi. The EEG was digitized at a sampling rate of 250 Hz and filtered with band pass 0.1–40 Hz. The ERPs with correct responses were averaged separately off-line with averaging epochs beginning 200 ms before stimulus onset and continuing for 1000 ms. Trials contaminated by eye blinks, eye movements, or muscle potentials exceeding $75 \, \mu V$ at any electrode were excluded from the average. ERP amplitudes were measured with respect to the mean voltage during the 200-ms pre-stimulus interval and peak latencies were measured relative to stimulus onset. Measurements and analyses of mean amplitudes and peak latencies of ERP components were conducted at selected electrodes over the frontal (16–20; 53–56), parietal (71–75; 90–94), temporal (88–96; 102–108;), and occipital (103–107; 118–120) regions, respectively, where specific ERP components or difference waves showed maximum amplitudes. Voltage topographies of ERPs and difference waves were plotted on a realistic head model of a randomly selected subject.

Reaction times (RTs) and error rates were subjected to repeated measure analyses of variance (ANOVAs) with Grouping (proximity versus similarity of shape) and Congruency (proximity and similarity cues were congruent or incongruent) as independent variables. The mean amplitudes of ERPs were subjected to ANOVAs with Grouping, Congruency, and Hemisphere (electrodes over the left versus right hemisphere) as independent variables.

RTs were shorter in the proximity- than similarity-grouping conditions (461 versus 520 ms, $F(1, 13) = 7.76$, $P < 0.015$), but were longer in the incongruent than congruent conditions (506 versus 476 ms, $F(1, 13) = 18.72$, $P < 0.001$). The congruency effect was larger in the similarity- than proximity-grouping conditions ($F(1, 13) = 4.73$, $P < 0.05$). Post hoc analyses confirmed that the congruence effect was significant in both proximity- and similarity-grouping conditions ($F(1, 13) = 12.00$ and 11.52, respectively, $P < 0.005$). Error rates did not differ between proximity- and similarity-grouping conditions (5.7% versus 7.0%, $F(1, 13) = 3.31$, $P > 0.08$) but were higher in the incongruent than congruent conditions (7.7% versus 5.1%, $F(1, 13) = 20.09$, $P < 0.001$). The interaction of Grouping $\times$ Congruency was not significant.
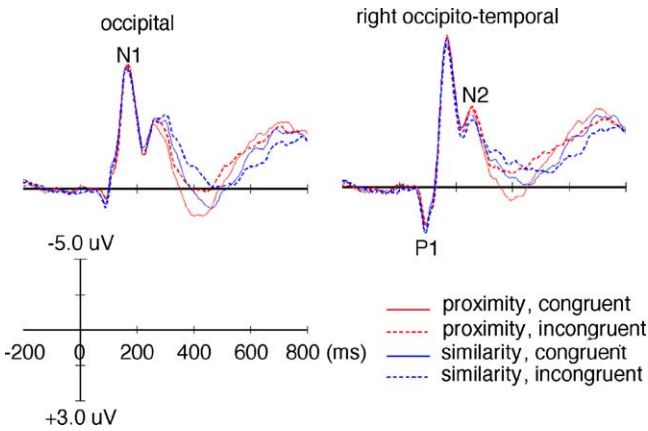
Fig. 2. ERPs in different stimulus conditions at electrodes over the posterior sites.

Fig. 2 shows grand averaged ERPs in different conditions. To illustrate the difference between proximity and similarity grouping, difference waves were obtained by subtracting ERPs in the similarity-grouping condition from ERPs in the proximity-grouping condition. The difference waves at electrodes over the posterior occipital areas were first characterized with a positive wave between 100 and 140 ms (Pd120) (Fig. 3), which was confirmed by ANOVAs showing a significant main effect of Grouping between 120 and 140 ms ($F(1, 13) = 5.17$, $P < 0.04$). The Pd120 was slightly larger over the left than the right hemisphere ($F(1, 13) = 10.17$, $P < 0.007$) but was not different between congruent and incongruent conditions ($F < 1$). Proximity grouping also elicited larger P3 amplitudes between 300 and 440 ms at electrodes over parietal, temporal, and occipital areas ($F(1, 13) = 8.00–13.65$, $P < 0.02$). Grouping modulation of the early phase of the P3 (300–340 ms) was stronger over the right then left parietal cortex ($F(1, 13) = 9.65$, $P < 0.008$). The P3 peaked earlier in the proximity-

than similarity-grouping conditions ($F(1, 13) = 10.14$, $P < 0.007$).

Congruent grouping cues elicited larger amplitude at the descending phase of the N1 component (180–220 ms) relative to incongruent grouping cues over the temporal–parietal areas ($F(1, 13) = 5.05$, $P < 0.04$). The congruency effect was stronger in the similarity- than proximity-grouping conditions ($F(1, 13) = 5.06$, $P < 0.04$). Post hoc analyses confirmed the congruency effect in the similarity-grouping condition ($F(1, 13) = 6.36$, $P < 0.02$) but not in the proximity-grouping condition ($F < 1$). Congruent grouping cues also generated larger amplitudes in the rising phase of the P3 (300–440 ms) ($F(1, 13) = 7.57$, $P < 0.02$) and shorter peak latencies ($F(1, 13) = 20.95$, $P < 0.001$) relative to incongruent cues. However, the congruency effect on the P3 amplitudes and latencies did not differ between proximity- and similarity-grouping conditions ($P > 0.2$). In order to visualize the congruency effect, difference waves were obtained by subtracting ERPs in the congruent condition from ERPs in the incongruent condition (Fig. 4). Incongruent grouping cues elicited a positivity peaking at about 220 ms with maximum over the right temporal–parietal areas (Pd220). The late negative wave at 300–440 ms (Nd370) resulted from the congruency effect on the P3 amplitudes over the parietal sites.

The behavioral data are consistent with the proposal that proximity grouping occurs earlier than and dominates similarity grouping in visual perception [1,4–6]. Proximity grouping elicited an early stronger positivity (Pd120) than similarity grouping over the posterior occipital cortex, which was observed in the previous work when only one Gestalt law was at work in each stimulus display [3,8] and even when the stimulus arrays were not required to respond [7]. The results suggest that the early proximity-grouping related activity is evident regardless of whether top-down instructions emphasized the proximity cues in displays and thus may mainly reflect a bottom-up grouping process. As
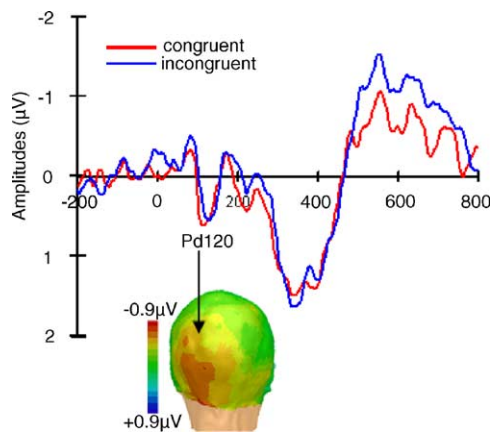

Fig. 3. The difference waves indexing early proximity grouping at the medial occipital electrode. Stronger occipital activities peaking at about 120 ms are shown in the congruent and incongruent conditions, respectively. The voltage topography shows the distribution of the Pd120 over the occipital area at 108–128 ms.
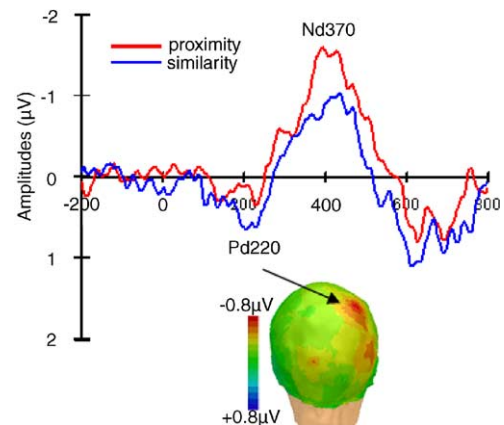

Fig. 4. The difference wave indexing the congruency effect at an electrode over the right parieto-temporal region. The voltage topography at 208–236 ms shows the distribution of the Pd220 over the right parieto-temporal areas.

the stimulus arrays were identical in the proximity and similarity conditions, the Pd120 could not be simply attributed to the difference in the power of low spatial frequencies between the two conditions [2], according with the finding that this proximity-grouping-related activity was evident even when proximity grouping could not be conducted based on low spatial frequencies [8]. We showed further that the Pd120 was not influenced by the incongruent similarity cues, providing ERP evidence that proximity dominates similarity at an earlier stage of grouping operations.

The first sign of interactions between proximity and similarity grouping was evident at 180–220 ms over the posterior temporal–parietal areas. The N1 amplitudes in the similarity-grouping condition were enlarged by congruent relative to incongruent proximity cues, whereas the N1 amplitudes in the proximity-grouping condition were not influenced by congruency with similarity cues. It appears that, at an early stage, proximity cues produced interference over the similarity-grouping process whereas the reverse did not occur. Moreover, the congruency effect indexed by the Pd220 was salient over the right parieto-temporal areas, possibly reflecting the dominance of the right hemisphere in processing low spatial frequencies [10] that, at least partially, support the proximity grouping [2]. The congruent grouping cues also elicited larger P3 amplitudes and shorter latencies relative to the incongruent cues. The effect did not differ between proximity and similarity conditions and might arise from either perceptual factors such as delayed stimulus evaluation and categorization [13,14] or motor factors such as slowed response selection and execution.

Kubovy et al. showed that the strength of grouping into strips of local items of a particular orientation can be measured quantitatively [12]. According to their analysis, the stimulus parameters used here generated a hierarchical grouping process. The two elements closest together are grouped into pairs, which are then grouped at a higher level of the hierarchy based on the shortest distance between the pairs. In the proximity-grouping condition, the second-level grouping operation determined the perception of columns versus rows in stimulus displays. In the similarity-grouping condition, the first-level proximity grouping formed pairs of elements with different shapes in the congruent condition (impairing similarity grouping) but with identical shape in the incongruent condition (facilitating similarity grouping). The second-level proximity grouping, however, was conflicting with similarity cues in the incongruent condition but consistent with similarity cues in the congruent condition. Given the behavioral and ERP results, it may be proposed that the second-level proximity grouping plays a critical role in the interactions between proximity- and similarity-grouping processes in the stimulus displays, which occurred at different stages of processing indexed by the N1 and P3 modulations.

The current work did not measure the relative weight of similarity and proximity by comparing the shape difference with the spatial distances. However, our findings suggest that, with the parameters determining local element grouping used in the current experiment, representation of spatial relations between local elements based on proximity cues provides the initial cue for segmenting the visual field into perceptual units. Similarity cues are engaged at a later stage in the grouping process and can be weaken by incongruent proximity cues at a later stage of processing.

Acknowledgements

References

[1] M.B. Ben-Av, D. Sagi, Perceptual grouping by similarity and proximity: experimental results can be predicted by intensity autocorrelations, Vis. Res. 35 (1995) 853–866.

[2] A.P. Ginsburg, Spatial filtering and visual form perception, in: K.R. Boff, L. Kaufman, J.P. Thomas (Eds.), Handbook of Perception and Human Performance, John Wiley & Sons, Inc., 1986, Chapter 34, pp. 1–71.

[3] S. Han, Y. Ding, Y. Song, Neural mechanisms of perceptual grouping in humans as revealed by high density event related potentials, Neurosci. Lett. 319 (2002) 29–32.

[4] S. Han, G.W. Humphreys, Interactions between perceptual organization based on Gestalt laws and those based on hierarchical processing, Percept. Psychophys. 61 (1999) 1287–1298.

[5] S. Han, G.W. Humphreys, L. Chen, Uniform connectedness and classical Gestalt principles of perceptual grouping, Percept. Psychophys. 61 (1999) 661–674.

[6] S. Han, G.W. Humphreys, L. Chen, Parallel and competitive processes in hierarchical analysis: perceptual grouping and encoding of closure, J. Exp. Psychol. Hum. Percept. Perform. 25 (1999) 1411–1432.

[7] S. Han, Y. Jiang, Attentional modulations of perceptual grouping in the visual cortex, Perception 32 (Suppl.) (2003) 59.

[8] S. Han, Y. Song, Y. Ding, E.W. Yund, D.L. Woods, Neural substrates for visual perceptual grouping in humans, Psychophysiology 38 (2001) 926–935.

[9] D. Kahneman, A. Henik, Perceptual organization and attention, in: M. Kubovy, J. Pomerantz (Eds.), Perceptual Organization, Erlbaum, Hillsdale, NJ, 1981, pp. 181–211.

[10] F. Kitterle, S. Christman, J. Hellige, Hemispheric differences are found in the identification, but not the detection, of low versus high spatial frequencies, Percept. Psychophys. 48 (1990) 297–306.

[11] K. Koffka, Principle of Gestalt Psychology, Harcourt Brace Jovanovich, New York, 1935.

[12] M. Kubovy, A.O. Holcombe, J. Wagemans, On the lawfulness of grouping by proximity, Cogn. Psychol. 35 (1998) 71–98.

[13] G. McCarthy, E. Donchin, A metric for thought: a comparison of P300 latency and reaction time, Science 211 (1981) 77–80.

[14] A. Mecklinger, P. Ullsperger, P3 varies with stimulus catergorization rather than probability, Electroencephalogr. Clin. Neurophysiol. 86 (1993) 95–407.

[15] P.T. Quinlan, R.N. Wilton, Grouping by proximity or similarity? Competition between the Gestalt principles in vision, Perception 27 (1998) 417–430.