AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Human Perception and Performance

ISSN: 0096-1523

2020, Vol. 46, No. 12, 1434–1442
http://dx.doi.org/10.1037/xhp0000866

Selective Attention Operates on the Group Level for Interactive Biological Motion

Huichao Ji
Sun Yat-Sen University

Jun Yin
Ningbo University

Yushang Huang and Xiaowei Ding
Sun Yat-Sen University

How do we distribute attention to interactive biological motion (BM)? There are 2 main hypotheses: (a) distribution-by-individual hypothesis, suggesting that interactive BM will not be taken as one unit in attention distribution, and an individual BM is independently selected; and (b) distribution-by-group hypothesis, indicating that interactions between BM can integrate them as one attention unit. We examined these hypotheses using a modified cueing paradigm. Participants observed 4 upright BM interacting in pairs (paired condition) or not interacting (unpaired condition), and after a cue for 1 agent, they determined whether the probe was the letters T or L (Experiment 1, sample size = 20). The results demonstrated better performance for probes presented on BM in the same interaction compared to BM equidistant but in different interactions. These differences were not present in the unpaired condition or for inverted BM (Experiment 2, sample size = 20), excluding a possible influence of low-level features. The findings indicate that interactive BM can serve as the elementary unit of attention in accordance with the distribution-by-group hypothesis.

Public Significance Statement

Interactive biological motion can serve as the elementary unit of attention in accordance with the distribution-by-group hypothesis, but attention is not selectively distributed to agents in interactions depicted by unpaired or inverted biological motion.

Keywords: selective attention, social interaction, biological motion, distribution-by-individual, distribution-by-group

Humans operate in a social world crowded with single and interactive biological motion (BM). The manner in which people selectively attend to this information is critical for survival and social adaptation (Blakemore & Decety, 2001; Blake & Shiffrar, 2007; Hirai & Senju, 2020; Yovel & O'Toole, 2016). For example, perception of BM facilitates remote identification (Yovel & O'Toole, 2016), as well as action imitation and learning (Giese & Poggio, 2003). Most previous studies on BM focused on selective attention allocated to a single BM (for reviews, see Blake & Shiffrar, 2007; Giese & Poggio, 2003); however, these studies have neglected to investigate BM in social situations. The study of interactive BM is important, because the social interactions involved with multiple animate entities are important to people's survival (Hafri, Papafragou, & Trueswell, 2013). Then faced with

This article was published Online First September 24, 2020.

Huichao Ji, Department of Psychology, Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Sun Yat-Sen University; Jun Yin, Department of Psychology, Ningbo University; Yushang Huang and Xiaowei Ding, Department of Psychology, Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Sun Yat-Sen University.

Huichao Ji and Jun Yin contributed equally to this work.

This work was supported by grants from Humanities and Social Sciences Foundation of the Ministry of Education of China [19YJC190004] and Fundamental Research Funds for Colleges and Universities–Key Training Program for Young Teachers [19wkzd23] to Xiaowei Ding. We have no conflicts of interest to disclose. The materials used in this project were obtained from http://mocap.cs.cmu.edu. The database was created with funding from NSF EIA-0196217. All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/uvqgh/?view_only=112d5ebd070f4350b72fe7ba000746ae.

Huichao Ji, Xiaowei Ding, Jun Yin, and Yushang Huang conceived and designed the experiments. Huichao Ji and Yushang Huang performed the experiments under the supervision of Xiaowei Ding. Huichao Ji analyzed the data. Huichao Ji, Yushang Huang, and Xiaowei Ding wrote the manuscript.

Correspondence concerning this article should be addressed to Xiaowei Ding or Yushang Huang, Department of Psychology, Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Sun Yat-Sen University, 132 Waihuan East Road, Guangzhou 510006, China. E-mail: dingxw3@mail.sysu.edu.cn or huangysh5@mail.sysu.edu.cn

numerous instances of interactive BM in daily life (e.g., two people quarreling), the question emerges as to how selective attention operates on interactive BM. Two possible hypotheses exist in answering this question: the distribution-by-individual hypothesis and the distribution-by-group hypothesis.

The Distribution-by-Individual Hypothesis: From Single Object to Single BM

Object-based theory suggests a distribution-by-individual hypothesis. That is, although involved in social interactions, interactive BM will not be taken as one unit in attention distribution, showing that individuals are independently selected. Object theory takes single object as a unit in attention system (for a review, see Scholl, 2001), as supported by both behavioral (e.g., Vickery & Chun, 2010) and neurophysiological evidence (e.g., Müller & Kleinschmidt, 2003). Researchers have developed various paradigms to study attention distribution, such as the Eriksen flanker task (Eriksen & Eriksen, 1974), choice reaction time (RT) task to examine the Simon effect (Simon & Rudell, 1967), Posner cueing task (Posner, 1980), and visual search task (Yantis & Jonides, 1984). One of the most widely used methods to study object-based attention is the spatial cueing paradigm proposed by Egly, Driver, and Rafal (1994). In their experiment, participants were asked to detect a target after a cue appeared at any of the four ends of two distinct rectangles. They found that participants responded faster to the targets appearing on the same rectangle as the cue but at a different end, compared with targets appearing on another rectangle. They thus concluded that attention can be distributed automatically to objects. The same theory has been presented for subjective contours (e.g., Moore, Yantis, & Vaughan, 1998), grid-like patterns (i.e., tiles on a floor; Ongchoco & Scholl, 2019), and even human faces (Baldauf & Desimone, 2014). Recently, object-based theory was further examined in real-world settings (Malcolm & Shomstein, 2015). The results indicated that this theory is still supported in the real world, full of both lower-level (such as edges) and higher-level (semantic) stimuli.

Previous studies have provided evidence consistent with the distribution-by-individual hypothesis. For example, researchers found that participants first and mainly attended to people in real-world settings compared to other, nonliving stimuli (Birmingham, Bischof, & Kingstone, 2008). Furthermore, eye movement results have suggested that participants distribute attention almost equally to individuals, whether they are involved in social interactions or not. Subsequent studies then only used a single person to study attention distribution, acquiescently taking a single person as an attention unit (e.g., Bindemann, Scheepers, Ferguson, & Burton, 2010; Fletcher-Watson, Findlay, Leekam, & Benson, 2008). Meanwhile, studies on working memory used a single actor to study the capacity of working memory for retaining action information, taking individual BM as the elementary unit (Gao, Ye, Shen, & Perry, 2016; Wood, 2007). Based on the similarities between working memory and attention (deBettencourt, Keene, Awh, & Vogel, 2019), it is reasonable to assume that a similar selection mechanism exists in attention allocation. Taken together, these findings suggest that even when involved in social interactions, only a single BM may be given priority during processing.

The Distribution-by-Group Hypothesis: Interactions Between BM Integrate Them as a Unit

The distribution-by-group hypothesis represents an alternative to the distribution-by-individual hypothesis and claims that the interactions between BM integrate them as an attention unit. In daily life, we often need to perceive and process human interactions from a third-party perspective (for a review, see Quadflieg & Penton-Voak, 2017). In this context, beyond a physical entity, it is necessary to detect social information between different agents, make social judgments, and modify our behavior in accordance with social interactions at any given moment (Quadflieg & Penton-Voak, 2017; Schwarzkopf, Büchner, Hölscher, & Konieczny, 2017), such as avoiding walking between two interacting individuals (Knowles, 2015). Researchers found that infants already can perceive collaborative relationships between two actors acting sequentially (Fawcett & Gredebäck, 2013) and can detect the nonmatching relation in a dyadic interaction (Gredebäck & Melinder, 2010). An fMRI study indicated that, compared to teasing situations, the right amygdala, fusiform gyrus, extrastriate body area-human motion complex, and superior temporal sulcus exhibited stronger activation when participants were presented with a threatening interaction, even when the interacting dyad was task-irrelevant (Sinke, Sorger, Goebel, & de Gelder, 2010), indicating that people have an excellent ability to detect social interaction information. These studies of social perception thus make it reasonable to assume that in these complex social situations, selective attention will be more effective when socially interactive BM is taken as an attention unit.

In accordance with the distribution-by-group hypothesis, previous studies have found that attention is automatically distributed to cooperative objects in a dynamic chase (Yin et al., 2013) and two static hands performing a handshake (Shen, Yin, Ding, Shui, & Zhou, 2016; Yin, Xu, Duan, & Shen, 2018). Using a visual search paradigm, researchers found that a static-facing dyad sketch was more likely to capture attention and was more easily recognized than a nonfacing dyad (Papeo & Abassi, 2019; Papeo, Goupil, & Soto-Faraco, 2019; Papeo, Stein, & Soto-Faraco, 2017). A similar processing priority was found in short-term and longer-term visual memory (for a review, see Kaiser, Quek, Cichy, & Peelen, 2019). That is, compared to noninteracting individuals, automatically binding people into interacting groups induced faster processing and enhanced working memory and longer-term memory (Vestner, Tipper, Hartley, Over, & Rueschemeyer, 2019). However, hardly any of the published research has examined the distribution of attention on dynamically interactive BM, in which irrelevant physical features, such as hair and body shape, are controlled. The absence of such research has resulted in a lack of direct evidence concerning how selective attention operates on interactive BM. In this study, we conducted two experiments to examine the two possible hypotheses on this question: the distribution-by-individual hypothesis and the distribution-by-group hypothesis.

Current Study

We used point-light displays (PLDs) attached to the major joints of human agents (Johansson, 1973) to depict interactive BM, which is effective for studying the perception of human action (e.g., Abernethy, Gill, Parks, & Packer, 2001). To study responses to interactive BM, this research adopted a modified two-rectangle cueing paradigm (Egly et al., 1994). In the classical cueing paradigm, participants were presented with a cue for one location and then asked to complete a letter discrimination task. The cue could be valid (V; the same location as the target) or one of two types of invalid (different locations from the target) cues. The invalid cues could appear on the same object (IS; invalid, same object) or on a different object (ID; invalid, different object; Figure 1a). Comparing performance in the three cue conditions allows for a comparison of spatial and object-based allocation of attention.

In the present research, the classical cueing paradigm was modified by two objects being replaced by two interactive BM pairs (Figure 1b, Experiment 1), and participants were told that there were four PLDs. Thus, if the distribution-by-individual hypothesis is correct, the four BM would be selected individually, whether they are involved in the same interactions or not. Then, only the cueing effect would be observed, as response times (RTs) to targets appearing in cued locations would be faster than to those in uncued locations ($RT_V < RT_{IS}$; $RT_V < RT_{ID}$). However, if the distribution-by-group hypothesis is correct, then both the cueing effect ($RT_V < RT_{IS}$; $RT_V < RT_{ID}$) and an automatic allocation of attention to BM involved in the same interaction ($RT_{IS} < RT_{ID}$) would be observed. Additionally, this would only occur in the paired condition (meaningful interaction). To exclude the possible influence of low-level features, we conducted Experiment 2, in which the BM stimuli were inverted.

Experiment 1

In Experiment 1, we compared interactive BM pairs against noninteractive ones under different cue settings. According to the

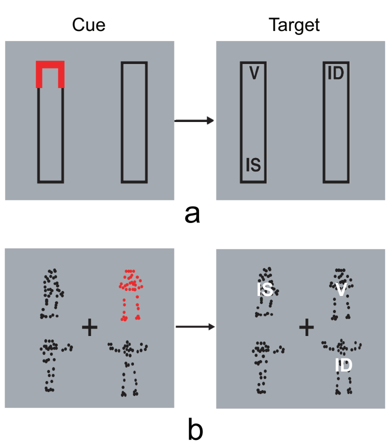

Figure 1. The comparison between the classical paradigm (a) in attention studies (Egly et al., 1994) and our research (b). In our research (b), the four PLDs were presented on the four corners of a square, with the geometric center of each being equidistant from the others. The upper two PLDs are presented as interacting with each other, as are the lower two PLDs. Participants were presented with a cue to one location and then asked to complete a letter (T or L) discrimination task. The cues are represented here by V for valid trials, IS for invalid, same-object trials in a and, correspondingly, invalid, same-event trials in b, and ID for invalid different-object trials in a and, correspondingly, invalid different-event trials in b. See the online article for the color version of this figure.

distribution-by-group hypothesis, if the task target appeared in an uncued BM, and this BM formed an interactive pair with the cued BM, task performance would be better. If no such a difference were found, then the distribution-by-individual hypothesis would be correct.

Method

Participants. Experiment 1's participants comprised 20 students (5 men and 15 women) from Sun Yat-Sen University, aged between 18 and 23 years ($M = 21$ years). All the participants were right-handed and had normal or corrected-to-normal vision. The sample size was decided based on previous studies that provided the background knowledge for the current research (Yin et al., 2018). We used BUCSS (Anderson, Kelley, & Maxwell, 2017) to perform the power analysis for the within-subject analysis of variance (ANOVA) in our study (https://designingexperiments.shinyapps.io/BUCSS_ss_power_wa/). Specifically, the *observed F value* was 15.65. The total sample size of the previous study was 20. The main effect we focused on was the 2 × 3 interaction between display type (paired/unpaired) and validity (valid/invalid same-event/invalid different-event). The target alpha level for the previous study was 0.05, as well as the planned study. The desired statistical power was set to 0.8, with a level of assurance of 0.8, which means we had an 80% chance to obtain the desired statistical power or higher with the estimated sample size. With the above settings, the calculated value was 12. The sample size of 20 in the present study should be sufficient for obtaining a moderate statistical power. Written informed consent was obtained before participation. Participants received financial rewards after completing the experiment. The experimental protocol was approved by the institutional review board of the Department of Psychology, Sun Yat-Sen University.

Stimuli and apparatus. The PLDs of the agents' motion used in Experiment 1 were selected and adapted from the Motion Capture Database built by Carnegie Mellon University Graphics Laboratory (http://mocap.cs.cmu.edu). The following selection rules were applied: (a) clear interactive behaviors and (b) little direct physical touch between the two agents. Five types of interactions were selected, including quarreling, fighting, conversing, dashing, and drinking. Each agent in the PLD consisted of 41 points, and the display size of a single agent's PLD was approximately 0.5° (width) × 1.8° (height), with variations depending on body gestures. Paired PLDs were drawn from the matching motion data displaying two agents' motion showing one interaction type (e.g., two agents quarreling with each other). In unpaired PLDs, scrambled combinations were drawn from different motion data from the five interactions (e.g., one agent is in conversation while the other is fighting), while each agent was not changed or flipped. That is, the two agents in the unpaired condition were selected from different sides (left or right) of different interactive pairs. Since there were only five types of motion interactions available, we conducted a preliminary Likert evaluation of all possible randomized combinations and then selected the five combinations that were reported to be the most improbable matching pairs. This selection process was conducted to ensure that the number of stimuli in each set was equal in the paired and unpaired conditions. All PLDs were clipped from raw data with the key portions of the interaction preserved (e.g., an outburst of body movements in a threatening interaction; Figure 2).

The target to be reported was a white letter T or L, and the distractors were combinations of T and L at 0°, 90°, 180°, and 270° orientations (Figure 3a). Both the target and the distractors subtended a 0.6° visual angle. The program was prepared in MATLAB 2016 with PsychToolbox extensions and implemented on a Windows 7 (64-bit) desktop computer with a 27-in. LCD (100Hz refresh rate) screen and a standard QWERTY keyboard, in a dark laboratory room with acoustic shielding.

Procedure. At the beginning of each trial, a black cross was presented at the screen's center as a fixation point. Participants were asked to focus on the cross and told that four agents' PLDs would appear 500ms later in four positions around the fixation cross (top-left, top-right, bottom-left, and bottom-right). The PLDs were constantly pinned to their positions for all motion types (similar to walking on a treadmill). Pinning the PLDs ensured that the distances from the body centers of each of the four agents to the screen center remained equal (3.9° approximate visual angle). The PLD array was presented for 3000ms, and each motion was continuously played as repetitions during the presentation. After this period, one of the PLDs was highlighted (color change) as a visual cue, lasting for 400ms before the cue disappeared (color was restored to the original). The target and distractors appeared 500ms after the cue disappeared. The participants were asked to detect the target and report whether it was a T or an L. The selection was indicated by pressing corresponding keys as quickly as possible without losing accuracy. Participants were not reminded of the interactions, apart from the presence of the four PLDs. RTs and accuracy levels were recorded (RTs longer than 1500ms were recorded as misses).

A total of 320 trials were conducted, with four blocks each for the paired and unpaired conditions. This total was the accumulation of each participant in each experiment. Each block contained 40 trials. In each block, 60% of the trials were valid (V), 20% of the trials were invalid same-event (IS), and 20% of trials were invalid different-event (ID). As a result, we set a within-subject 2 (display type: paired/unpaired) × 3 (validity: valid/invalid same-event/invalid different-event) design (Figure 3b). In the paired blocks, both the top and bottom halves of the PLD array were paired PLDs, randomly drawn from the five motion types. Thus, the top half of the display contained one interaction, while the bottom half contained a different interaction. In the unpaired blocks, the two halves comprised PLD combinations randomly drawn from the five unpaired combinations, and the diagonal PLDs were never drawn from matching interactions.

The validity of a trial was defined by the relationship between the location of the cue and the location of the target. Specifically, in valid trials, the target (a discrete T or L) was positioned in the center of the highlighted PLD (cued position). In invalid same-event trials, the target was positioned horizontally opposite (left vs. right, within the same interaction) the highlighted PLD (cued position), and in invalid different-event trials, the target was presented vertically opposite (top vs. bottom, within a different interaction) to the cued position.

Data analysis. Only trials with correct performances on the discrimination task were considered for further analysis. We examined participants' RTs using repeated measures ANOVA with display type and validity as the within-subject factors. Then, a linear mixed model (LMM) analysis was conducted to verify the effects' uniformity across the five interaction types. Finally, we compared participants' accuracy rates in the same way to detect any trade-off between response speed and response accuracy.

Results

The mean RTs for the correct responses are shown in Figure 4. The RT data were analyzed in a 2 (display type: paired/unpaired) × 3 (validity: valid/invalid same-event/invalid different-event) repeated measures ANOVA. The main effect of display type was not statistically significant, $F(1, 19) = 0.78$, $p = .389$, $\eta_p^2 = 0.04$. There was a significant main effect for validity, $F(1.31, 24.81) = 49.73$, $p < .001$, $\eta_p^2 = 0.72$, and a significant interaction between the two factors, $F(2, 38) = 6.51$, $p = .004$, $\eta_p^2 = 0.26$. Post hoc analyses with Bonferroni correction confirmed that RTs in valid trials were faster than those in invalid same-event trials, $p < .001$, 95% CI [−123.95, −62.88], and invalid different-event trials, $p < .001$, 95% CI [−143.17, −66.13]. These findings reflected facilitation of RTs induced by the attention cue.

To examine whether there were any differences among the five interaction types, we took interaction type as the fixed factor in an additional LMM analysis on the RTs, while adding a random intercept for each participant. The results suggested that in the paired condition for invalid same-event trials, interaction type made no significant contribution to the model depicting participants' RTs, AIC = 7723.13, $F(4, 599) = 0.69$, $p = .602$. Further pairwise comparisons showed that there was no significant differ-

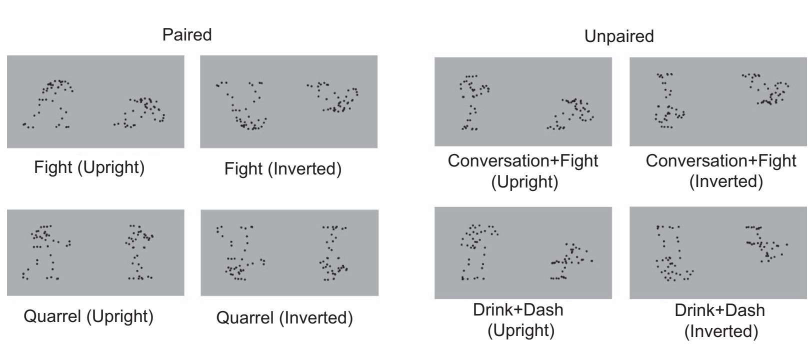

Figure 2. Examples of the PLD motion stimuli used in Experiments 1 (upright) and 2 (inverted).

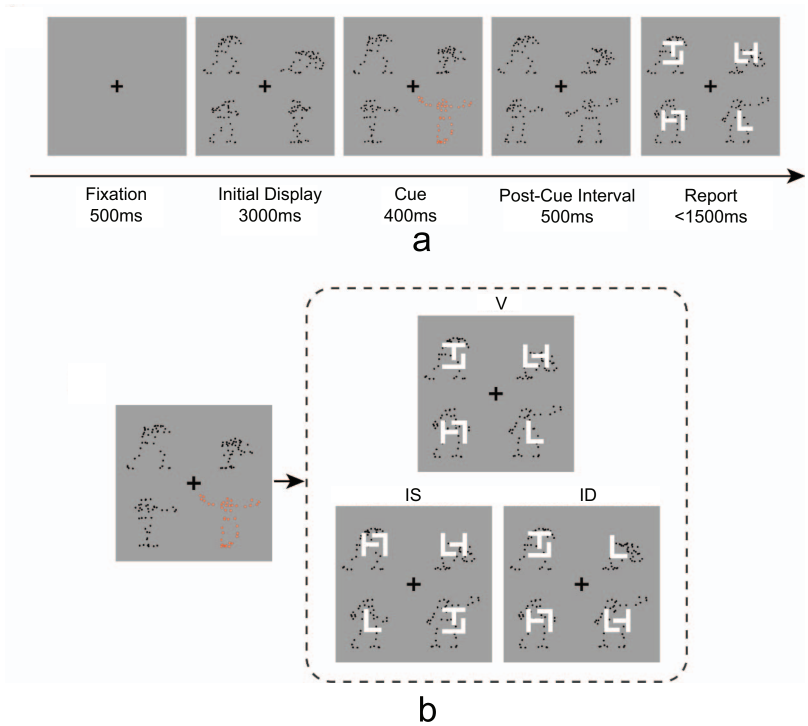

Figure 3. The design of Experiment 1. Participants were told there would be four PLDs after the fixation point, then they would be presented with a cue and asked to complete a letter (T or L) discrimination task (a). Three types of trials (b) were used: the cue might be valid (V, in the same location as the target), invalid same-event (IS, in a different location from the target, but within the same event), or invalid different-event (ID, in a different location from the target and in a different event). See the online article for the color version of this figure.

ence between any two of the five types, *p*s $>$ .05. Similar results were found in the unpaired condition. The interaction type made no significant contribution to the model depicting participants' RTs, AIC = 7721.76, $F(4, 599) = 1.04$, $p = .388$. Further pairwise comparisons showed that there was no significant difference between any two of the five types, *p*s $>$ .05.

To further analyze the interaction effect, we calculated the cue effect by subtracting the mean RT in the valid trials from the mean RTs in the invalid same-event trials and invalid different-event trials. This was done for both display types (Figure 4b). For the paired condition, the cue effect of the invalid same-event trials ($M = 75.76$ms) was significantly smaller than that for the invalid different-event trials ($M = 111.07$ms), $t(19) = -2.28$, $p = .034$, Cohen's $d = 0.51$, 95% CI [−67.71, −2.91]. No significant differences were observed in the unpaired condition (invalid same-event: $M = 108.48$ms; invalid different-event: $M = 100.83$ms), $t(19) = 0.63$, $p = .537$, Cohen's $d = 0.14$, 95% CI [−17.82, 33.12]. These findings indicated that when one of the two interactive BM was selected for attentional allocation, the other could also benefit. This supports the distribution-by-group hypothesis' statement that interactive BM is the elementary unit of selective attention, even when observers were not aware of the interactions between the four BM.

In addition to the RT analysis, the accuracy data were also analyzed. The main effect of display type was not statistically significant (paired: $M = 95.10\%$; unpaired: $M = 95.00\%$), $F(1, 19) = 0.00$, $p = .951$, $\eta_p^2 = 0$. The interaction effect was not statistically significant, $F(2, 38) = 0.64$, $p = .533$, $\eta_p^2 = 0.03$, although the main effect of validity was statistically significant (valid: $M = 96.41\%$; invalid same-event: $M = 94.22\%$; invalid different-event: $M = 94.45\%$), $F(2, 38) = 3.74$, $p = .033$, $\eta_p^2 = 0.17$. Post hoc analyses with Bonferroni correction did not reveal any accuracy differences among different trials ($p > .05$ in all cases). Thus, the results demonstrated that the RTs could not be attributed to a trade-off between response speed and response accuracy.

Experiment 2

Previous studies indicated that BM perception was significantly vulnerable to vertical inversion. That is, inversion could disrupt the semantic processing of single (Pavlova & Sokolov, 2000) as well as interactive BM (Ding, Gao, & Shen, 2017). Therefore, we conducted Experiment 2 in which all BM stimuli were inverted, to prevent the perception of social interaction. If the effect found in Experiment 1 disappeared in Experiment 2, we could confirm that the effect was caused by the social interaction information rather than the stimuli's low-level features.

Method

Participants. The sample for Experiment 2 comprised 20 students (8 men and 12 women) from Sun Yat-Sen University,

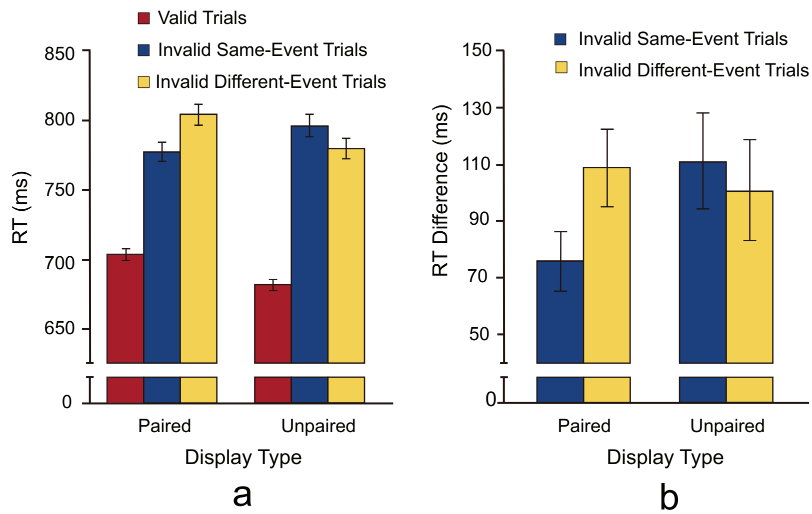

Figure 4. Results of Experiment 1. Mean response times (RTs; a) varied as a function of display condition and validity. The mean RT differences between valid trials to invalid same-event and invalid different-event trials (b). Error bars indicate 95% CI with between-subjects variance removed (Cousineau & O'Brien, 2014). See the online article for the color version of this figure.

aged between 18 and 25 years (M = 20 years). In accordance with the discussion in Experiment 1's Method section, this sample size should be sufficient for our target statistical power. All participants were right-handed and had normal or corrected-to-normal vision. The participants received financial rewards after completing the experiment. Participants gave written informed consent before their participation.

Stimuli and apparatus. The PLDs used in Experiment 2 were the inverted versions of those used in Experiment 1. The targets and distractors were directly adopted from Experiment 1 with no changes.

Procedure. The procedure and design of Experiment 2 were identical to those of Experiment 1.

Results

Figure 5 shows the mean RTs for trials with correct responses in all conditions. To examine the cue effects for the inverted BM, we conducted a 2 (display type: paired/unpaired) × 3 (validity: valid/invalid same-event/invalid different-event) repeated measures ANOVA with the RT data. A significant main effect of validity was found, $F(1.15, 21.91) = 28.75$, $p < .001$, $\eta_p^2 = 0.60$. Post hoc analyses with Bonferroni correction showed that the RTs in the valid trials were faster than those in the invalid same-event trials, $p < .001$, 95% CI [−103.85, −36.32], and the invalid different-event trials, $p < .001$, 95% CI [−106.38, −37.96]. These findings confirmed the existence of the cue effect. No significant differences were found between the invalid same-event and invalid different-event trials, $p = 1.000$, 95% CI [−12.85, 8.68]. The effect was not modulated by display type, as the interaction effect between the two factors was not statistically significant, $F(2, 38) = 0.63$, $p = .539$, $\eta_p^2 = 0.03$. The main effect of display type was not statistically significant, $F(1, 19) = 0.12$, $p = .738$, $\eta_p^2 = 0.01$. Unlike in Experiment 1, invalid same-event and invalid different-event presentations showed no statistically significant differences in Experiment 2. This finding suggested that the invalid same-event advantage observed in Experiment 1 was not due to the influence of low-level stimuli.

We again examined the fixed effect of pair type on mean RTs using LMM analysis, taking subjects as a random factor. Separate analyses were conducted for the paired and unpaired conditions. The results suggested that in the paired condition, interaction type made no significant contribution to the model depicting participants' RTs in invalid same-event trials, AIC = 7855.92, $F(4, 600) = 0.43$, $p = .788$. Further pairwise comparisons showed that there was no significant difference between any two of the five types, $ps > .05$. Similar results were found in the unpaired condition. Interaction type made no significant contribution to the

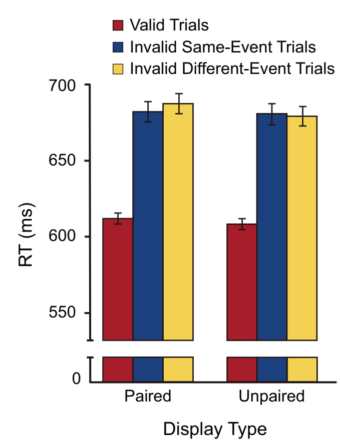

Figure 5. Results of Experiment 2. Mean response time (RT) varied as a function of display condition and validity. Error bars indicate 95% CI with between-subjects variance removed (Cousineau & O'Brien, 2014). See the online article for the color version of this figure.

model depicting participants' RTs, AIC = 8067.09, $F(4, 610) = 0.40$, $p = .813$. Further pairwise comparisons showed that there was no significant difference between any two of the five types, ps $> .05$.

The analyses regarding the accuracy data exhibited the same pattern of results. Specifically, only the main effect of validity was statistically significant (valid: $M = 98.02\%$; invalid same-event: $M = 95.31\%$; invalid different-event: $M = 95.47\%$), $F(2, 38) = 7.40$, $p = .002$, $\eta_p^2 = 0.28$. Neither the main effect of display type (paired: $M = 96.50\%$; unpaired: $M = 97.44\%$), $F(1, 19) = 1.37$, $p = .256$, $\eta_p^2 = 0.07$, nor the interaction effect, $F(2, 38) = 0.65$, $p = .526$, $\eta_p^2 = 0.03$, was statistically significant. Post hoc analyses with Bonferroni correction revealed that the responses in the valid trials were more accurate than those in the invalid same-event trials, $p = .003$, 95% CI [0.01, 0.05], and invalid different-event trials, $p = .023$, 95% CI [0, 0.05]. The difference between the two invalid conditions was not statistically significant, $p = 1.000$, 95% CI [−0.02, 0.02]. Thus, the RT results could not be attributed to a trade-off between response speed and response accuracy.

To further compare the performance of participants in Experiment 1 and Experiment 2, we conducted a 2 (orientation: Exp1_upright/Exp2_inverted) × 2 (display type: paired/unpaired) × 3 (validity: valid/invalid same-event/invalid different-event) mixed-design ANOVA with the RT data. The main effect of orientation was significant, M(Exp1_upright) = 760.15, M(Exp2_inverted) = 658.54, $F(1, 38) = 14.19$, $p = .001$, $\eta_p^2 = 0.27$. The main effect of validity was also significant, M(valid) = 652.621, M(invalid same-event) = 734.37, M(invalid different-event) = 741.03, $F(1.26, 47.84) = 77.53$, $p < .001$, $\eta_p^2 = 0.67$. Post hoc analyses with Bonferroni correction revealed that the RTs in the valid trials were faster than those in the invalid same-event trials, $p < .001$, 95% CI [−103.47, −60.03], and invalid different-event trials, $p < .001$, 95% CI [−112.99, −63.83]. There was no significant difference in invalid same-event and invalid different-event trials, $p = .316$, 95% CI [−3.397, 16.719]. The main effect of display type was not significant, M(paired) = 713.59, M(unpaired) = 705.09, $F(1, 38) = 0.715$, $p = .403$, $\eta_p^2 = 0.02$. The interaction between validity and display type was significant, $F(1.74, 65.96) = 6.66$, $p = .002$, $\eta_p^2 = 0.15$. However, neither the interaction between orientation and validity nor that between orientation and display type was significant, ps $> .05$. The three-way interaction was significant, $F(1.74, 65.96) = 3.39$, $p = .046$, $\eta_p^2 = 0.08$, indicating that there were different patterns among different conditions in Experiment 1 and Experiment 2, consistent with the aforementioned results.

Thus, the integrated analysis of Experiment 1 and Experiment 2 showed that inverted BM destroyed the perception of social interaction, and the effect in Experiment 1 was caused by social interaction information, rather than the stimuli's low-level features.

Discussion

To answer the question concerning how selective attention operates on interactive BM, the current study modified the two-rectangle cueing paradigm presented by Egly et al. (1994) to test two possible hypotheses: the distribution-by-individual hypothesis and the distribution-by-group hypothesis. PLDs were used as stimulus materials, and the interactive relationship (paired or unpaired) was manipulated between the two agents to examine the effect on a letter discrimination task. The findings indicated that when one of the two paired PLDs was cued, the partner interacting with it would also benefit from the attention guided by the cue. This suggests that the two interacting BM can be selected as a unit during attention allocation, even when we only told the participants that there were four PLDs in the display. These findings are congruent with the distribution-by-group hypothesis. Specifically, compared to the cue-invalid different-event trials, the cue-invalid same-event trials resulted in faster responses (Experiment 1). When the interaction was broken (unpaired condition in Experiment 1) or the motion stimuli were inverted (Experiment 2), the advantage attributed to the interaction vanished. These results suggest that the observed advantage may have stemmed from the presence of social interaction information when the PLDs could be perceived as meaningful human action (Ding et al., 2017; Neri, Luu, & Levi, 2006), rather than being due to the low-level features of the stimuli, such as the specific motion pattern of one dot. Additionally, the disappearance of the interaction-specific advantage in the unpaired condition excluded the possible influence of a horizontal advantage (Chen & Cave, 2019) or coupling.

This study elaborates and expands on previous object-based research. Contemporary behavioral and neuroimaging studies have yielded many meaningful findings concerning selective attention with objects as elementary units (e.g., Egly et al., 1994; Vickery & Chun, 2010), even when competing factors were present (Malcolm & Shomstein, 2015). For example, Erlikhman, Lytchenko, Heller, Maechler, and Caplovitz (2020) found object-based attention can be generalized to multisurface objects. In terms of social information processing, discrete entities and events defined under social context may also play similar roles to discrete objects in object-based attention. Previous studies on visual perception of social interactions laid a good foundation for this hypothesis. For example, researchers found people could detect dangerous social behavior between two interacting agents from third-party encounters (Hamlin, 2013; Mast & Hall, 2004), even with a short glance (Quadflieg & Penton-Voak, 2017). Furthermore, previous studies have found that events are key components of perception, attention, and memory (Zacks, Speer, Swallow, Braver, & Reynolds, 2007). This study, involving BM, depicted various social events. Thus, we can propose a new concept labeled event-based attention, reiterating the significance of events and expanding on the concept of object-based attention in cognition.

Our results echoed those studies on working memory. Previous researchers claimed that objects are the basic units for working memory storage (e.g., Luck & Vogel, 1997), and single BM has been used to study the capacity of working memory (Gao et al., 2016; Wood, 2007). However, recent studies suggested that two interactive BM can be stored in working memory as a unified chunk (Ding et al., 2017). Since attention shares common points with working memory (deBettencourt et al., 2019), and there is a mutual influence between them (for a review, see Soto, Hodsoll, Rotshtein, & Humphreys, 2008), the present study's findings provided new evidence on the similarity between working memory and attention. As taking interactive BM as a whole is common in working memory and selective attention, it may represent a universal mechanism that interactive BM can act as a unit in other cognitive processes, such as perception and long-term memory.

The present study also provides evidence and expansion regarding research on perceptual organization. The classical view proposes that sensory organization obeys principles such as proximity, closure, and common fate (e.g., Köhler, 1967); however, increasing numbers of studies have found that mutual gaze (e.g., Nawroj, Toneva, Admoni, & Scassellati, 2014), joint attention (e.g., Leach, Baxter, Robertson, & Sparks, 2014), interpersonal distance, and angles (Zhou, Han, Liang, Hu, & Kuai, 2019) can also affect the judgment of social interactions. In the present experiments, there were no such clues (e.g., gaze) or spatial Gestalt principles, but a similar effect was observed. Therefore, we inferred that the social interactive information carried by BM can serve as a basis for social grouping and attentional selection, such as motor direction.

However, there are several follow-up questions. First, the present study did not address the time course of the mechanisms related to social interaction processing, an important question for low-level processing. Second, we used only two-agent motion in the present study; thus, whether social interactions involving more agents would play the same role in selective attention is currently unclear. These questions require exploration in future work.

Conclusion

The present study found a "same-event advantage" demonstrating that, in meaningful interactions, attention tended to be assigned, or allocated, to BM in the same interaction. This led to faster RTs in comparison to BM which was equidistant but in different interactions. These findings support the distribution-by-group hypothesis, which states that interactive BM can serve as the elementary unit of attention.

References

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception, 30,* 233–252. http://dx.doi.org/10.1068/p2872

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science, 28,* 1547–1562. http://dx.doi.org/10.1177/0956797617723724

Baldauf, D., & Desimone, R. (2014). Neural mechanisms of object-based attention. *Science, 344,* 424–427. http://dx.doi.org/10.1126/science.1247003

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1477–1485. http://dx.doi.org/10.1037/a0019057

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real-world scenes: The roles of action, competition and social content. *The Quarterly Journal of Experimental Psychology, 61,* 986–998. http://dx.doi.org/10.1080/17470210701410375

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology, 58,* 47–73. http://dx.doi.org/10.1146/annurev.psych.57.102904.190152

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2,* 561–567. http://dx.doi.org/10.1038/35086023

Chen, Z., & Cave, K. R. (2019). When is object-based attention not based on objects? *Journal of Experimental Psychology: Human Perception and Performance, 45,* 1062–1082. http://dx.doi.org/10.1037/xhp0000657

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods, 46,* 1149–1151. http://dx.doi.org/10.3758/s13428-013-0441-z

deBettencourt, M. T., Keene, P. A., Awh, E., & Vogel, E. K. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour, 3,* 808–816. http://dx.doi.org/10.1038/s41562-019-0606-6

Ding, X., Gao, Z., & Shen, M. (2017). Two equals one: Two human actions during social interaction are grouped as one unit in working memory. *Psychological Science, 28,* 1311–1320. http://dx.doi.org/10.1177/0956797617707318

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123,* 161–177. http://dx.doi.org/10.1037/0096-3445.123.2.161

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143–149. http://dx.doi.org/10.3758/BF03203267

Erlikhman, G., Lytchenko, T., Heller, N. H., Maechler, M. R., & Caplovitz, G. P. (2020). Object-based attention generalizes to multisurface objects. *Attention, Perception, & Psychophysics, 82,* 1599–1612. http://dx.doi.org/10.3758/s13414-019-01964-5

Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science, 16,* 841–849. http://dx.doi.org/10.1111/desc.12074

Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception, 37,* 571–583. http://dx.doi.org/10.1068/p5705

Gao, Z., Ye, T., Shen, M., & Perry, A. (2016). Working memory capacity of biological movements predicts empathy traits. *Psychonomic Bulletin & Review, 23,* 468–475. http://dx.doi.org/10.3758/s13423-015-0896-2

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience, 4,* 179–192. http://dx.doi.org/10.1038/nrn1057

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114,* 197–206. http://dx.doi.org/10.1016/j.cognition.2009.09.004

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General, 142,* 880–905. http://dx.doi.org/10.1037/a0030045

Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition, 128,* 451–474. http://dx.doi.org/10.1016/j.cognition.2013.04.004

Hirai, M., & Senju, A. (2020). The two-process theory of biological motion processing. *Neuroscience and Biobehavioral Reviews, 111,* 114–124. http://dx.doi.org/10.1016/j.neubiorev.2020.01.010

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14,* 201–211. http://dx.doi.org/10.3758/BF03212378

Kaiser, D., Quek, G. L., Cichy, R. M., & Peelen, M. V. (2019). Object vision in a structured world. *Trends in Cognitive Sciences, 23,* 672–685. http://dx.doi.org/10.1016/j.tics.2019.04.013

Knowles, E. S. (2015). Spatial behavior of individuals and groups. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 53–86). New York, NY: Psychology Press.

Köhler, W. (1967). Gestalt psychology. *Psychologische Forschung, 31,* 18–30. http://dx.doi.org/10.1007/BF00422382

Leach, M. J., Baxter, R., Robertson, N. M., & Sparks, E. P. (2014). Detecting social groups in crowded surveillance videos using visual attention. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 467–473). Columbus, UK: IEEE.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281. http://dx.doi.org/10.1038/36846

Malcolm, G. L., & Shomstein, S. (2015). Object-based attention in real-world scenes. *Journal of Experimental Psychology: General, 144,* 257–263. http://dx.doi.org/10.1037/xge0000060

Mast, M. S., & Hall, J. A. (2004). Who is the boss and who is not? Accuracy of judging status. *Journal of Nonverbal Behavior, 28,* 145–165. http://dx.doi.org/10.1023/B:JONB.0000039647.94190.21

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9,* 104–110. http://dx.doi.org/10.1111/1467-9280.00019

Müller, N. G., & Kleinschmidt, A. (2003). Dynamic interaction of object- and space-based attention in retinotopic visual areas. *The Journal of Neuroscience, 23,* 9812–9816. http://dx.doi.org/10.1523/JNEUROSCI.23-30-09812.2003

Nawroj, A., Toneva, M., Admoni, H., & Scassellati, B. (2014). An exploration of social grouping in robots: Effects of behavioral mimicry, appearance, and eye gaze. *Proceedings of the Annual Meeting of the Cognitive Science Society, 36,* 1060–1065.

Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience, 9,* 1186–1192. http://dx.doi.org/10.1038/nn1759

Ongchoco, J. D. K., & Scholl, B. J. (2019). How to create objects with your mind: From object-based attention to attention-based object. *Psychological Science, 30,* 1648–1655. http://dx.doi.org/10.1177/0956797619863072

Papeo, L., & Abassi, E. (2019). Seeing social events: The visual specialization for dyadic human-human interactions. *Journal of Experimental Psychology: Human Perception and Performance, 45,* 877–888. http://dx.doi.org/10.1037/xhp0000646

Papeo, L., Goupil, N., & Soto-Faraco, S. (2019). Visual search for people among people. *Psychological Science, 30,* 1483–1496. http://dx.doi.org/10.1177/0956797619867295

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. *Psychological Science, 28,* 369–379. http://dx.doi.org/10.1177/0956797616685769

Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics, 62,* 889–899. http://dx.doi.org/10.3758/BF03212075

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32,* 3–25. http://dx.doi.org/10.1080/00335558008248231

Quadflieg, S., & Penton-Voak, I. S. (2017). The emerging science of people-watching: Forming impressions from third-party encounters. *Current Directions in Psychological Science, 26,* 383–389. http://dx.doi.org/10.1177/0963721417694353

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80*(1–2), 1–46. http://dx.doi.org/10.1016/S0010-0277(00)00152-9

Schwarzkopf, S., Büchner, S. J., Hölscher, C., & Konieczny, L. (2017). Perspective tracking in the real world: Gaze angle analysis in a collaborative wayfinding task. *Spatial Cognition and Computation, 17*(1–2), 143–162. http://dx.doi.org/10.1080/13875868.2016.1226841

Shen, M., Yin, J., Ding, X., Shui, R., & Zhou, J. (2016). Deployment of attention on handshakes. *Frontiers in Psychology, 7,* 681. http://dx.doi.org/10.3389/fpsyg.2016.00681

Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51,* 300–304. http://dx.doi.org/10.1037/h0020586

Sinke, C. B., Sorger, B., Goebel, R., & de Gelder, B. (2010). Tease or threat? Judging social interactions from bodily expressions. *NeuroImage, 49,* 1717–1727. http://dx.doi.org/10.1016/j.neuroimage.2009.09.065

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12,* 342–348. http://dx.doi.org/10.1016/j.tics.2008.05.007

Vestner, T., Tipper, S. P., Hartley, T., Over, H., & Rueschemeyer, S. A. (2019). Bound together: Social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. *Journal of Experimental Psychology: General, 148,* 1251–1268. http://dx.doi.org/10.1037/xge0000545

Vickery, T. J., & Chun, M. M. (2010). Object-based warping: An illusory distortion of space within objects. *Psychological Science, 21,* 1759–1764. http://dx.doi.org/10.1177/0956797610388046

Wood, J. N. (2007). Visual working memory for observed actions. *Journal of Experimental Psychology: General, 136,* 639–652. http://dx.doi.org/10.1037/0096-3445.136.4.639

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601–621. http://dx.doi.org/10.1037/0096-1523.10.5.601

Yin, J., Ding, X., Zhou, J., Shui, R., Li, X., & Shen, M. (2013). Social grouping: Perceptual grouping of objects by cooperative but not competitive relationships in dynamic chase. *Cognition, 129,* 194–204. http://dx.doi.org/10.1016/j.cognition.2013.06.013

Yin, J., Xu, H., Duan, J., & Shen, M. (2018). Object-based attention on social units: Visual selection of hands performing a social interaction. *Psychological Science, 29,* 1040–1048. http://dx.doi.org/10.1177/0956797617749636

Yovel, G., & O'Toole, A. J. (2016). Recognizing people in motion. *Trends in Cognitive Sciences, 20,* 383–395. http://dx.doi.org/10.1016/j.tics.2016.02.005

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133,* 273–293. http://dx.doi.org/10.1037/0033-2909.133.2.273

Zhou, C., Han, M., Liang, Q., Hu, Y. F., & Kuai, S. G. (2019). A social interaction field model accurately identifies static and dynamic social groupings. *Nature Human Behaviour, 3,* 847–855. http://dx.doi.org/10.1038/s41562-019-0618-2

Received February 6, 2020
Revision received July 14, 2020
Accepted July 15, 2020 ■