Perceiving crowd attention: Gaze following in human crowds with conflicting cues

Zhongqiang Sun¹ · Wenjun Yu² · Jifan Zhou² · Mowei Shen²

Published online: 7 March 2017
© The Psychonomic Society, Inc. 2017

Abstract People automatically redirect their visual attention by following others' gaze orientation, a phenomenon called “gaze following.” This is an evolutionarily generated socio-cognitive process that provides people with information about their environments. Often, however, people in crowds can have rather different gaze orientations. This study investigated how gaze following occurs in situations with many conflicting gazes. In two experiments, we modified the gaze cueing paradigm to use a crowd rather than a single individual. Specifically, participants were presented with a group of human avatars with differing gaze orientations, and the target appeared randomly on the left or right side of a display. We found that (a) when a marked difference existed in the number of avatars with divergent gaze orientations, participants automatically followed the majority's gaze orientation, and (b) the strongest gaze cue effect occurred when all gazes shared the same orientation, with the response superiority of the majority's oriented location monotonically diminishing with the number of gazes with divergent orientations. These findings suggested that the majority rule plays a role in gaze following behavior when individuals are confronted with conflicting multigaze scenes, and that an increasing subgroup size appears to enlarge the strength of the gaze cueing effect.

Keywords Attention: selective · Attention and executive control · Visual perception

Zhongqiang Sun and Wenjun Yu contributed equally to this work.

✉ Mowei Shen
mwshen@zju.edu.cn

¹ Department of Psychology, Ningbo University, Ningbo, China

² Department of Psychology and Behavioral Sciences, Xixi Campus, Zhejiang University, Hangzhou, China 310028

As social animals, humans are experts in observing and understanding conspecifics' behaviors and signals in a crowd (Hare, Call, Agnetta, & Tomasello, 2000; Frith & Frith, 2012; Sumpter, 2006). Not only can this help with interaction but it also can provide information about the surrounding world, such as immediate predation risks or food sources (Galef & Giraldeau, 2001; Griffin, 2004). An effective means by which individuals receive information from others in a group is observing the attentional focus of these others through their gaze (or head) orientation. This orientation implies conspecifics' interests, which may provide important information about the surrounding physical and social environment (Shepherd, 2010).

This tendency to analyze others' gaze exists virtually from birth in humans (e.g., Beier & Spelke, 2012; Vecera & Johnson, 1995). Remarkably, humans exhibit rapid shifts in gaze and attention toward the locations on which others' gazes are fixated. Indeed, there is a great deal of research focusing on the properties of this “gaze following” response (e.g., Friesen & Kingstone, 1998; Langton & Bruce, 1999; Zhang, Tang, Zhang, & Zhang, 2015) and its underlying neural mechanisms (Hietanen, Nummenmaa, Nyman, Parkkola, & Hämäläinen, 2006; Joseph, Fricker, & Keehn, 2015), with a particular emphasis on its key role in human interaction. Furthermore, gaze following exists in a diverse range of animal species (Carpenter & Call, 2013; Emery, 2000; Schloegl, Kotrschal, & Bugnyar, 2007), suggesting that it has a broad adaptive significance.

Intriguingly, people exhibit a correspondingly rapid attentional shift when observing a group of people with a consistent gaze orientation. Compared to when using individual gaze cues, the proportion of individuals who engage in gaze following further increases when gaze information is available in crowds (e.g., Gallup et al., 2012; Milgram, Bickman, & Berkowitz, 1969). The earliest study of gaze following in

crowds dates back to 1969. On a crowded street in New York, Milgram et al. (1969) had a stimulus group of individuals stop and stare up together into a building window. The probability of passers-by following this behavior was then measured, and the results showed that the probability of a person following the stimulus group's action increased with the size of the stimulus group. In other words, the increasing number of consistently directed gazes induced a correspondingly greater cueing strength, which in turn led to an amplified gaze following effect (i.e., a higher probability of stopping and following). Recently, psychological researchers revived this field in several studies investigating the extent, influence, a contextual dependence of social visual attention transmission (Bayliss et al., 2013; Gallup, Chong, Kacelnik, Krebs, & Couzin, 2014; Gallup et al., 2012), gaze perception (Florey, Clifford, Dakin, & Mareschal, 2016; Sweeny & Whitney, 2014), and even virtual experience (Narang, Best, Randhavane, Shapiro, & Manocha, 2016).

In many situations, however, people's gaze orientations or preferences in crowds diverge (Conradt, 1998; Couzin, Krause, Franks, & Levin, 2005; Ruckstuhl & Neuhaus, 2000, 2002). Specifically, it is common to see multiple gazes of varying orientations at the same moment, especially during emergencies or frantic situations such as panic or rioting (Helbing, Farkas, & Vicsek, 2000; Shiwakoti & Sarvi, 2013). In such situations, how is the conflicting gaze information processed? In this case, the gaze following response must be reconsidered. To our knowledge, few studies have systematically examined the gaze following effect in conflicting gaze cue situations. Researchers are far from clear on how gaze cues guide individuals' attention when they are faced with a crowd with a variety of diverging gaze orientations. Therefore, this study aimed to clarify this point.

When comparing scenes involving a single gaze or consistently directed gazes, a more sophisticated mechanism in conflicting multigaze scenes is required. Take, for example, two subsets of individuals with totally opposing gaze orientations (left and right), which can be regarded as a relatively simple and common conflicting situation involving divergent gazing orientations. Based on previous studies (e.g., Gallup et al., 2012; Milgram et al., 1969), the group with more consistent gaze orientation (i.e., the majority) might have greater cueing strength than the other group (i.e., the minority), so we assumed that the majority would be prioritized and followed during processing. In an evolutionary sense, information from the majority has been considered rather reliable (Galef & Giraldeau, 2001; Griffin, 2004), thus following the majority could help a group to effectively resolve conflicts and make collective decisions, and finally maintain an adaptive advantage (Conradt & Roper, 2009; Sumpter & Pratt, 2009).

Moreover, in most situations involving conflicting gaze orientations, there is a size difference between subgroups (i.e., a majority and a minority). How is the distribution of

attention affected by the extent of this difference? This is our second concern. It is possible that attention follows an “all or none” principle, wherein all attentional resources would be distributed to the majority (and hence their gaze orientation). If this were the case, attentional shifts would be equally quick regardless of the size difference between the majority and minority. In contrast, the distribution of attention on the cued side might increase with the gap between the majority and the minority—although this allocation of attention may not follow a linear rule, it would almost certainly be monotonic. In other words, each subgroup would achieve a portion of an individual's attentional resources.

Considering the above background, we investigated the influence of social attention information derived from conflicting gazes in human crowds on individuals' attention allocation. We modified the gaze cueing paradigm (e.g., Friesen & Kingstone, 1998) so that a human crowd was employed. Specifically, a group of human avatars was presented at the center of a screen, and participants were asked to do a target identification task in which a target appeared randomly on the left or right side of this group. As the avatars' eyes would be too small to be recognized at the distance at which they were viewed, we used head orientation as the index of social attention. According to previous studies, head orientation, which is easier to recognize than gaze orientation (Perrett, Hietanen, Oram, Benson, & Rolls, 1992), has a similar capacity to trigger attentional shifts as eye gaze does (e.g., Hietanen, 1999; Nuku & Bekkering, 2008; Sato, Okada, & Toichi, 2007); hence, we adopted head-gaze orientation (wherein the head and gaze always share the same orientation) instead of merely gaze orientation. To investigate our second concern—namely, the relation between the degree of difference in subgroup size and the distribution of attention—we manipulated the sizes of the two subgroups of gaze orientation among the experimental conditions.

Experiment 1

Method

Participants

Sixteen graduate and undergraduate students (seven females; mean age 21.2 ± 2.1 years) were paid to participate in this experiment. None had a history of neurological problems, and all had normal or corrected-to-normal vision. The participants provided their written informed consent before the experiment, and all experimental procedures were approved by the Research Ethics Board of Zhejiang University.

The sample size of the study was determined via a power analysis using G*Power 3 (Faul, Erdfelder, Buchner, & Lang, 2009). Given a large effect size ($f = .40$; partial $\eta^2 = .14$), a

power of .80, and an alpha level of .05, the power analysis ultimately yielded an estimated sample size of 14. Furthermore, because some participants might be necessitate exclusion, we decided to stop collecting data at $N = 16$.

Stimuli

Ten human avatars (six males and four females) were created using the 3D animation software Poser 6® (E Frontier, Scotts Valley, California, USA). Three head orientations for each avatar were produced: one facing straight ahead (and therefore looking directly at participants when presented in the middle of a computer screen), one with the head turned 30° leftward, and one with the head turned 30° rightward. All faces showed a neutral expression. The gaze orientation was always the same as the head orientation.

All 10 avatars were positioned in three rows (with the whole picture measuring 12° high and 17° wide), with each row comprising three or four avatars. Each avatar in the back, middle, and front rows measured 3° × 6° with a 1° × 1.5° face, 3.5° × 7° with a 1.25° × 1.75° face, and 4° × 8° with a 1.5° × 2° face, respectively. The upper body and face of each avatar was clearly displayed, not being covered by any other avatar.

A capitalized letter *T*, subtending 0.4° of visual angle and centered 13° to the left or right of a central fixation cross, was set as the target. All stimuli were presented on a gray background (RGB, 80, 80, 80).

Procedure and design

Participants were seated in an electrically shielded and sound-attenuated recording chamber at a distance of 70 cm from a 17-inch CRT monitor (with a 100-Hz refresh rate). We used the Presentation® software to control stimulus presentation

and response acquisition. Participants were asked to keep their eyes centrally fixated and were given clear instructions on how to perform the experimental trials.

The experimental procedure is illustrated in Fig. 1. Each trial began with a fixation cross presented at the center of the screen for a random duration (500–1,000 ms), which was then replaced with an array of all 10 avatars looking straight forward. After remaining for 600 ms, a 300-ms cue array was displayed, wherein each avatar was equally likely to change his or her head orientation to the left or right. Then, the target appeared on the left or right of the display. The participant was required to indicate the position of the target by pressing *F* or *J* on the keyboard. If no response was made after 1,500 ms, the trial was coded as a missing response, and the next trial was presented. The interval between trials was randomly determined from 1,000 to 1,500 ms. Both the accuracy and reaction time (RT) were recorded.

The experimental design involved two within-subjects factors: consistency (with levels of all consistent, 9vs1, 8vs2, 7vs3, 6vs4, and 5vs5) and validity (with levels of majority-valid and majority invalid).

The *consistency* factor involved the manipulation of the avatars' head orientation in the cue array, and included six conditions: one wherein all 10 avatars had their heads turned leftward or rightward (i.e., the all-consistent condition), and five with varying numbers of heads (nine, eight, seven, six, or five) with the same orientation and the other heads (one, two, three, four, or five, respectively) with the opposite orientation (i.e., the 9vs1, 8vs2, 7vs3, 6vs4, and 5vs5 conditions, respectively). To exclude the processes of perceptual subgrouping and also to avoid displaying a same picture under each condition throughout the experiment as much as possible, the position of avatars shifting in one direction or the other was randomized across conditions. To be specific, we produced eight versions of pictures for each consistency condition in which the positions of the minority avatars were randomly selected.

Fig. 1 An example of a trial in a 6vs4 condition

The majority-valid and majority-invalid trials existed only in the all-consistent, 9vs1, 8vs2, 7vs3, and 6vs4 conditions. On majority-valid trials, the letter was presented on the side where the majority of heads were oriented, whereas in majority-invalid trials, the letter appeared on the opposite side. In the 5vs5 condition, wherein the numbers of heads orienting to the left and right were equal, the letter was presented on the left side in 50% of the trials and the right side in the remaining trials.

For each of the six consistency conditions, participant completed 48 trials, which were further divided evenly into the two validity conditions (i.e., 24 trials in both majority-valid and majority-invalid conditions), resulting in a total of 288 randomly presented trials. To ensure adequate rest, the whole experiment was divided into six sessions with a 2-minute break between sessions. Before the formal experiment, there were at least 20 practice trials to ensure that the participants understood the instructions.

After all trials were finished, participants were asked to complete a survey, including rating the task difficulty (from 1 = *extremely easy* to 7 = *extremely difficult*), reporting whether they indeed followed the instructions (e.g., keep eyes centrally fixed) and any strategy they adopted during the experiment as well as making a guess on the experimental objective.

Results

Because no trial could be defined as majority valid or majority invalid in the 5vs5 condition, we divided those trials into two halves based on the position of the target, namely the left visual field and right visual field, and conducted paired *t* tests to determine whether participants showed a prior attention bias to a certain visual field. For the remaining trials, a 5 (consistency) $\times$ 2 (validity) two-way analysis of variance (ANOVA) was conducted for RT. Significant main effects were always followed by Bonferroni post-hoc contrasts.

All participants' mean accuracies exceed 98%, which were fairly high and reached the level of ceiling effects, suggesting that the task was so simple that all participants completed it very well.

Concerning the RTs, all results with RTs either lower than 100 ms or exceeding three standard deviations from the mean were excluded from further analyses (Friesen & Kingstone, 1998; Hayward & Ristic, 2015), with 97.50% of trials remaining. We observed no significant difference between the left visual field and right visual field in the 5vs5 condition, $M_{\text{left-visual-field}} = 354$ ms, $M_{\text{right-visual-field}} = 357$ ms, $t(15) = 0.92$, $p = .37$, Cohen's $d = 0.08$, suggesting that participants showed evenly distributed attention due to the equal number of avatars in each orientation subgroup.

For the other conditions, the two-way ANOVA for RT revealed a significant main effect of validity, $M_{\text{majority-valid}} = 345$ ms, $M_{\text{majority-invalid}} = 361$ ms, $F(1, 15) = 76.32$, $p < .001$, $\eta_p^2 = 0.84$. The Consistency $\times$ Validity interaction was also significant, $F(4, 60) = 8.38$, $p < .001$, $\eta_p^2 = 0.36$.

Bonferroni post hoc tests revealed that the difference of RT between the majority-valid and majority-invalid trials was significant for either the all-consistent, 9vs1, 8vs2, or 7vs3 condition, $ps < .006$, while that for the 6vs4 condition was non-significant, $p = .96$ (see Fig. 2a). No significant main effect of consistency was found, $F(4, 60) = 2.10$, $p = .09$, $\eta_p^2 = 0.12$.

Taking the difference of RT between the majority-valid and majority-invalid trials in each consistency condition as the dependent variable, we then performed *t* tests for RT differences between two of each consistency condition (see Fig. 2b), for investigating the interaction effect in more detail. The results indicated that the RT difference in the all-consistent condition was larger than those in the 9vs1 condition, $t(15) = 2.37$, $p = .031$, Cohen's $d = 0.64$; 8vs2, $t(15) = 3.48$, $p = .003$, Cohen's $d = 1.18$; 7vs3, $t(15) = 4.46$, $p < .001$, Cohen's $d = 1.34$; and 6vs4, $t(15) = 4.02$, $p = .001$, Cohen's $d = 1.71$. Furthermore, there was a larger RT difference in the 9vs1 condition than in the 7vs3 condition, $t(15) = 2.14$, $p = .049$, Cohen's $d = 0.60$, and the 6vs4 condition, $t(15) = 2.45$, $p = .027$, Cohen's $d = 1.06$, as well as a larger RT difference in the 8vs2 condition than in the 6vs4 condition, $t(15) = 3.22$, $p = .006$, Cohen's $d = 0.97$. No significant difference was found between the remaining conditions, $ps > .11$.

Discussion

The results implied that when faced with two subgroups of diverging gaze cues, participants follow the majority's gaze orientation, and the superiority of the cueing strength of the majority group would monotonically increase as the size difference between two subgroups enlarges.

Despite that, there's one important reason to be skeptical about this interpretation. Namely, the head orientation of the avatars might prime a corresponding left versus right response rather than causing a shift of attention to the cued location, since the response property of the target (left vs. right) was identical to the cue. In other words, a corroborating measure might be helpful to clarify at what level of processing the observed biases are occurring.

Second concern for this interpretation is that, as participants were instructed to maintain central fixation, they might only select and process the central avatar in each trial, so that the different probability levels of this central face as a member of the majority or minority should be reflected in the different RT differences between majority valid and majority invalid. We improved the experimental design in Experiment 2 to solve this issue.

Experiment 2

Instead of the localization task in Experiment 1, we used a target identification task in Experiment 2 to confirm this

Fig. 2 Results of Experiment 1. **a** Reaction times (RTs) for the five consistency conditions. The asterisks represent a significant difference ($p < .05$) between the majority-valid and majority-invalid conditions. The numbers present RT values of the nearest data point. The error

bars present one SEM. **b** RT differences for the five consistency conditions. The asterisks represent a significant difference ($p < .05$) between two corresponding conditions

attentional effect, because head orientation would not prime the response.

To investigate whether the central faces dominated the processing of crowd gaze following in Experiment 1, we focused on the pair of conditions whose RT differences significantly differed, such as all-consistent and 9vs1 conditions. In Experiment 2, the central face in the 9vs1 condition was always arranged as a member of the majority, which is the same as the situation of central face in the all-consistent condition. If participants only encode the central face, the patterns of RT difference in the all-consistent and 9vs1 conditions should be the same. In addition, the condition where RT was faster in majority valid than in majority invalid, such as 8vs2, might provide further evidence after rearranging their central faces. Thus, we set the central faces in the 8vs2 condition to be equally possible as a member of the majority and minority. In this case, mean RT for majority-valid trials should be the same as that for majority-invalid trials if the central face is

solely processed; otherwise, RT difference in the 8vs2 condition should remain different, as in Experiment 1. To improve the reliability of this further analysis, more trials were adopted in Experiment 2.

Method

Participants

We followed the same data collection rule as in Experiment 1. Two suspicious participants were excluded from analyses (they reported to be aware of our experimental objective after the experiment), leaving 14 participants (seven females; mean age 21.0 ± 1.67 years). None had a history of neurological problems, and all had normal or corrected-to-normal vision. The participants provided their written informed consent before the experiment, and all experimental procedures were

approved by the Research Ethics Board of Zhejiang University.

Stimuli

Avatars were the same as those in Experiment 1. The arrangements of 10 avatars in the 9vs1, 8vs2, and 6vs4 conditions had minor changes according to corresponding experiment designs (see [Procedure and Design](#) section).

A capitalized letter *L* or *T*, subtending 0.4° of visual angle and centered 13° to the left or right of a central fixation cross, was set as the target.

Procedure and design

Most of the procedure was the same as in Experiment 1, except for the target. Participants were asked to indicate whether the target was *L* or *T* by pressing *F* or *J* on the keyboard.

Consistency (with levels of all-consistent, 9vs1, 8vs2, 7vs3, 6vs4, and 5vs5) and validity (with levels of majority valid and majority invalid) were adopted as two within-subjects factors in Experiment 2. For each of the six consistency conditions, participant completed 72 trials, which were further divided evenly into the two validity conditions (i.e., 36 trials in both majority-valid and majority-invalid conditions), resulting in a total of 432 randomly presented trials. In the 9vs1 condition, the central face (the middle face in the second row) was always a member of the majority, and in the 8vs2 condition, the central face had equal probability as a member of the majority and minority. In the other conditions, the probabilities of the central face being a member of the majority and minority were pseudorandom without prior manipulation. After all trials were finished, participants were asked to complete a survey, which was the same as that in Experiment 1.

Results

The mean accuracy for all participants was 96%, implying a ceiling effect in accuracy, as in Experiment 1.

Concerning the RTs, all results with RTs either lower than 100 ms or exceeding three standard deviations from the mean were excluded from further RT analyses, with 95.19% of trials remaining. A same result pattern as in Experiment 1 was found. First, there was no significant difference between the left and right visual field in the 5vs5 condition, $M_{\text{left-visual-field}} = 688$ ms, $M_{\text{right-visual-field}} = 690$ ms, $t(13) = 0.23$, $p = .82$, Cohen's $d = 0.02$, suggesting that participants showed evenly distributed attention due to the equal number of avatars in each orientation subgroup. Second, the two-way ANOVA for RT in other conditions revealed a significant main effect of validity, $M_{\text{majority-valid}} = 662$ ms, $M_{\text{majority-invalid}} = 713$ ms, $F(1, 13) = 49.60$, $p < .001$, $\eta_p^2 = 0.79$, and a significant Condition $\times$ Validity interaction, $F(4, 52) = 16.50$, $p < .001$, $\eta_p^2 = 0.56$.

Bonferroni post hoc tests revealed that the difference of RT between the majority-valid and majority-invalid trials was significant for either the all-consistent, 9vs1, 8vs2, or 7vs3 conditions, $ps < .04$, while that for the 6vs4 condition was non-significant, $p = .40$ (see Fig. 3a). No significant main effect of consistency was found, $F(4, 52) = 0.36$, $p = .84$, $\eta_p^2 = 0.03$.

Again, taken the difference of RT between the majority-valid and majority-invalid trials in each consistency condition as the dependent variable, we then performed *t* tests for RT differences between each of the consistency conditions (see Fig. 3b). The RT difference in the all-consistent condition was larger than those in the 9vs1 condition, $t(13) = 5.02$, $p < .001$, Cohen's $d = 1.27$; 8vs2, $t(13) = 6.22$, $p < .001$, Cohen's $d = 1.54$; 7vs3, $t(13) = 5.10$, $p < .001$, Cohen's $d = 1.73$; and 6vs4, $t(13) = 7.45$, $p < .001$, Cohen's $d = 2.75$. Furthermore, RT difference in the 6vs4 condition was significantly smaller than in the 9vs1 condition, $t(13) = 4.47$, $p = .001$, Cohen's $d = 1.61$; 8vs2, $t(13) = 3.27$, $p = .006$, Cohen's $d = 1.24$; and 7vs3, $t(13) = 2.27$, $p = .04$, Cohen's $d = 0.88$. No significant difference was found between the remaining conditions, $ps > .16$.

As predicted, the overall RT in the 8vs2 condition for majority-valid trials was faster than that for majority-invalid trials. Since the central gazes were equally possible as a member of the majority and minority in the 8vs2 condition, we further separately analyzed the trials where the central gaze was consistent or inconsistent to the majority. Two central-majority consistency subconditions were generated accordingly: central-as-majority condition (the central face orientation is consistent with the majority) and central-as-minority condition (the central face orientation is inconsistent with the majority). We conducted 2 (central-majority consistency) $\times$ 2 (validity) ANOVA for RT in the 8vs2 condition and found a main effect for validity, $F(1, 13) = 13.28$, $p = .003$, $\eta_p^2 = 0.51$, just as that we have observed in Fig. 3a, but either the main effect for central-majority consistency, $F(1, 13) = 0.14$, $p = .72$, $\eta_p^2 = 0.01$, or their interaction, $F(1, 13) = 1.71$, $p = .21$, $\eta_p^2 = 0.12$, was nonsignificant (see Fig. 4).

Discussion

Using an identification task, we replicated the results in Experiment 1, confirming that by shifts of attention, RTs were facilitated to probes appearing at the location to which the majority of faces were looking.

More importantly, RT difference between the majority-valid and majority-invalid trials was larger in the all-consistent than in the 9vs1 conditions, even though central face was always as a member of the majority in both conditions. RT difference in the 8vs2 condition, where central faces had equal probability as a member of the majority and minority, was also found apparent. Moreover, further detailed analyses to split trials in the 8vs2 condition demonstrated that participants would follow the gaze orientation of the majority,

Fig. 3 Results of Experiment 2. **a** Reaction times (RTs) for the five consistency conditions. The *asterisks* represent a significant difference ($p < .05$) between the majority-valid and majority-invalid conditions. The *numbers* represent RT values of the nearest data point. The *error*

bars represent one *SEM*. **b** RT differences for the five consistency conditions. The *asterisks* represent a significant difference ($p < .05$) between two corresponding conditions

even the central gaze oriented to the opposite direction, implying that central gaze was not relied upon exclusively. These

comparison results also indicated that the effects for the majority-valid condition in 8vs2 were not entirely pulled by

Fig. 4 Results in the 8vs2 condition of Experiment 2. The *asterisks* represent a significant difference ($p < .05$) between the majority-valid and majority-invalid conditions. The *error bars* represent one *SEM*

the 50% of trials where the central gaze was consistent with the majority. Hence, it's clear that participants did not process only the central avatar.

General discussion

In this study, we explored the gaze following effect in scenes with conflicting gazes and obtained a rich set of results. First, when all gazes had the same orientation, individuals tended to follow these gazes (and thereby shift their attention to the corresponding location), which in turn led them to respond faster to a test stimulus presented on the cued side than on the other side. Second, when a marked difference existed between the numbers of gazes with divergent orientations, individuals would automatically follow the majority's gaze orientation; in contrast, when the number of gazes with divergent orientations was the same or only slightly different, the gaze following effect vanished. Third, the strongest gaze cue effect occurred when all gazes shared the same orientation, and the response superiority of the majority's oriented location monotonically diminished with the number of gazes with divergent orientations.

The majority rule of gaze following in conflict gaze cues

Our results showed that people follow the majority's gaze orientation when faced with a crowd of diverging gaze cues. This phenomenon corresponds to the majority rule that applies to group decision making (Kerr & Tindale, 2004; Hastie & Kameda, 2005). The majority rule—wherein groups make decisions based on the agreement of the majority of members—has been widely used across the full spectrum of human groups, from the Stone Age (Boehm, 1996; Boyd & Richerson, 1985) to the modern era (Mueller, 1989; Sorkin, West, & Robinson, 1998). A similar rule has been demonstrated in laboratory studies with highly controlled environments and parameters. For instance, when participants were asked to make yes-or-no decisions in a visual detection task, groups with a simple majority performed better in terms of detection sensitivity (Sorkin, Hays, & West, 2001). Arguably, the greatest benefit of the majority rule is that it allows for good decision-making performance for comparatively little cognitive effort (Kameda & Hastie, 1999).

In the current study, the majority's gaze orientation did not predict the location of the target during the experiments, and the participants were instructed to ignore the avatars. Despite this, the gaze following effect still occurred, which implies that it was an automatic process. This suggests that the majority rule of gaze following may be an evolutionary adaptation, and beneficial from two perspectives. First, one of the instincts of social animals is striving for the unity and stability of the group. The majority rule is one such strategy by which self-

interested behaviors can be constrained in complex decision-making scenarios (Henrich & Boyd, 1998; Kameda, Takezawa, Tindale, & Smith, 2002), as this helps in maintaining an adaptive advantage and achieving optimal species fitness (Conradt & Roper, 2009; Sumpter & Pratt, 2009).

Second, in general, information from the majority is rather reliable. For instance, during foraging or group migration (Galef & Giraldeau, 2001; Griffin 2004), conspecifics' gazes can be considered an initial alert toward a target. Thus, when a conflict of gaze orientation arises within the group, the majority's gaze orientation has a rather high likelihood of indicating the correct location of a food source or safe destination. To ensure that a consensus is reached regarding selection of a target, natural selection may have forced species to evolve strategies of following the majority's gaze; otherwise, the group might be misdirected, thereby leading to a loss of adaptive advantage. Indeed, there is evidence for a similar majority-rule gaze following effect among nonhuman animal groups, such as birds or fish; it is believed that this following allows for resolution of conflicts and making collective decisions (Conradt & Roper, 2003; Couzin et al., 2005; Seeley & Bhurman, 1999).

Taken together, this majority rule in the visual system might be shaped by evolutionary contingencies.

The strength of gaze cues is modulated by the difference in orientation subgroup size

The results indicated that the size of the gap in the numbers of the majority and minority group members influenced the cueing strength of the majority group. One intriguing aspect concerns the monotonic nature of this relationship. In other words, we could reject the “all or none” hypothesis mentioned above. More specifically, as the gap in the number of members between the two subgroups widened, there was a concomitant increase in the difference in RTs, which corresponded to a greater superiority of the cueing strength of the majority group.

One may argue that participants might do not have time to process all 10 gazes, and instead, they use a sampling strategy in which a single or several avatars were selected in each trial. Because of the probabilistic distribution of those 10 gaze orientations, the same pattern of results, which roughly matches the gaze distribution of the avatars, would be produced after all trials were finished.

This alternative possibility could be largely ruled out based on two reasons. First, human's visual system has sufficient capability to receive and process all gazes in a group. Summary statistical perception studies found that a human is equally rapid and precise to process a single stimulus and multiple stimuli (e.g., Ariely, 2001; Haberman & Whitney, 2009). Findings in gaze perception from Sweeny and Whitney (2014) also confirmed the simultaneous processing

ability of multiple gazes. Second, participants were asked to maintain central fixation during the experiment, and all of them reported that they followed this instruction in the survey after the experiment.¹ As discussed in Experiment 1, if participants indeed adopted this sampling strategy, the gaze orientation of the central face was the most likely to be sampled and encoded, so that the different probability levels of this central face as a member of the majority or minority should be reflected in the different RT differences between majority-valid and majority-invalid trials. We further analyzed the data in Experiment 2 and found that, although the central face was always set as a member of the majority in both the all-consistent and 9vs1 conditions, a gap of RT difference was still apparent between them. Together with a significant cueing effect after rearranging the central face in the 8vs2 trials in Experiment 2, we concluded that participants did not adopt this sampling strategy toward the central face.

The current results can be supported by previous studies at the individual level. In previous gaze cue studies, cueing strength was found to be modulated by social status (Dalmaso, Galfano, Coricelli, & Castelli, 2014; Dalmaso, Pavan, Castelli, & Galfano, 2012), political temperament (Carraro, Dalmaso, Castelli, & Galfano, 2015; Liuzza et al., 2011), facial expression (Lassalle & Itier, 2013; Tipples, 2006), facial dominance (Jones et al., 2010), and even social interaction history (Capozzi, Becchio, Willemse, & Bayliss, 2016; Dalmaso, Edwards, & Bayliss, 2016). Take, for example, social status. Previous studies have shown that high-status individuals, who might be considered as more relevant or reliable sources of information, had a stronger cueing strength and further resulted in a greater gaze cueing effect than did low-status individuals (Dalmaso et al., 2012). Analogously, in our study, it might have been adaptively appropriate for participants to orient themselves relatively quickly to objects that have captured the majority of gazes in a scene, because a growing number of gazes with identical gaze orientations may suggest increasing gaze cueing strength and group reliability to predict the existence of a target, and thus further raise likelihood of others' following behavior. This monotonic tendency is supported by the results of Milgram and his colleagues' field study (1969) described in the Introduction.

Nevertheless, choosing between the majority and minority may not be inevitable. In the 6vs4 condition, there was no selective gaze following response, despite the fact that these two subgroups had differing numbers of members. This suggests that the majority rule of gaze following only applies if the size difference between the two subgroups reaches a certain level. In

this sense, there may be a quorum-like relationship in the proportion of attention distribution and the size difference between subgroups (Franks, Dornhaus, Fitzsimmons, & Stevens, 2003; Ward, Sumpter, Couzin, Hart, & Krause, 2008). Animals typically change their own behavior only when the number of individuals performing that same behavior reaches a threshold (i.e., the quorum); otherwise, others' behavior is disregarded (Pratt, 2005). From an adaptive perspective, a quorum response could lower the possibility of nonadaptive behavior spreading, and may therefore maintain the group's stability (Sumpter & Pratt, 2009). Nevertheless, the current results merely provide preliminary supportive evidence for the possibility of this quorum; further study is needed to verify it.

Conclusion

By manipulating the size difference between majority and minority groups in an array, we explored how the quantity of conflicting gazes in a crowd influences the gaze-following effect. Our findings suggest that individuals automatically follow the majority's gaze orientation, and that the response superiority of the majority's oriented location monotonically diminishes with the size difference between the two subgroups.

Because conflicting multigaze scenes do not always contain two divergent subgroups in the real life, we must consider other possible conflicting scenes to fully understand this issue. Follow-up studies that aim to narrow the gap between picture stimuli and real human groups would also be needed and could be realized using certain technologies, such as virtual reality.

Acknowledgements This research is supported by the National Natural Science Foundation of China (No. 31571119; 31600881).

References

- Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, 12(2), 157–162. doi:10.1111/1467-9280.00327
- Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2013). "Gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, 142(1), 76–92. doi:10.1037/a0029286
- Beier, J. S., & Spelke, E. S. (2012). Infants' developing understanding of social gaze. *Child Development*, 83(2), 486–496. doi:10.1111/j.1467-8624.2011.01702.x
- Boehm, C. (1996). Emergency decisions, cultural-selection mechanics, and group selection. *Current Anthropology*, 37(5), 763–793. doi:10.1086/204561
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Capozzi, F., Becchio, C., Willemse, C., & Bayliss, A. P. (2016). Followers are not followed: Observed group interactions modulate

¹ Because it is important to provide direct evidence that participants followed the instructions, we conducted an additional eye-tracking experiment in which most settings and parameters were the same as those in Experiment 2. Results from a new set of 14 participants showed that eye movements >2.5° from the center occurred in 9.83% of the trials. For the remaining valid trials, the results patterns were totally the same as those in Experiment 2.

- subsequent social attention. *Journal of Experimental Psychology: General*, 145(5), 531–535. doi:10.1037/xge0000167
- Carpenter, M., & Call, J. (2013). How joint is the joint attention of apes and human infants. In J. Metcalfe & H. S. Terrace (Eds.), *Agency and joint attention* (pp. 49–61). New York, NY: Oxford University Press.
- Carraro, L., Dalmaso, M., Castelli, L., & Galfano, G. (2015). The politics of attention contextualized: Gaze but not arrow cuing of attention is moderated by political temperament. *Cognitive Processing*, 16, 309–314. doi:10.1007/s10339-015-0661-5
- Conradt, L. (1998). Could asynchrony in activity between the sexes cause intersexual social segregation in ruminants? *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1403), 1359–1368. doi:10.1098/rspb.1998.0442
- Conradt, L., & Roper, T. J. (2003). Group decision-making in animals. *Nature*, 421(6919), 515–518. doi:10.1038/nature01294
- Conradt, L., & Roper, T. J. (2009). Conflicts of interest and the evolution of decision sharing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1518), 807–819. doi:10.1098/rstb.2008.0257
- Couzin, I. D., Krause, J., Franks, N. R., & Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025), 513–516. doi:10.1038/nature03236
- Dalmaso, M., Edwards, S. G., & Bayliss, A. P. (2016). Re-encountering individuals who previously engaged in joint gaze modulates subsequent gaze cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(2), 271–284. doi:10.1037/xlm0000159
- Dalmaso, M., Galfano, G., Coricelli, C., & Castelli, L. (2014). Temporal dynamics underlying the modulation of social status on social attention. *PLoS ONE*, 9(3), e93139. doi:10.1371/journal.pone.0093139
- Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biological Letters*, 8, 450–452. doi:10.1098/rsbl.2011.0881
- Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6), 581–604. doi:10.1016/S0149-7634(00)00025-7
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. doi:10.3758/BRM.41.4.1149
- Florey, J., Clifford, C. W., Dakin, S., & Mareschal, I. (2016). Spatial limitations in averaging social cues. *Scientific Reports*, 6, 32210. doi:10.1038/srep32210
- Franks, N. R., Domhaus, A., Fitzsimmons, J. P., & Stevens, M. (2003). Speed versus accuracy in collective decision making. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1532), 2457–2463. doi:10.1098/rspb.2003.2527
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3), 490–495. doi:10.3758/BF03208827
- Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, 63, 287–313. doi:10.1146/annurev-psych-120710-100449
- Galef, B. G., & Giraldeau, L. A. (2001). Social influences on foraging in vertebrates: Causal mechanisms and adaptive functions. *Animal Behaviour*, 61(1), 3–15. doi:10.1006/anbe.2000.1557
- Gallup, A. C., Chong, A., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2014). The influence of emotional facial expressions on gaze-following in grouped and solitary pedestrians. *Scientific Reports*, 4, 5794. doi:10.1038/srep05794
- Gallup, A. C., Hale, J. J., Sumpter, D. J., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences*, 109(19), 7245–7250. doi:10.1073/pnas.1116141109
- Griffin, A. S. (2004). Social learning about predators: A review and prospectus. *Animal Learning & Behavior*, 32(1), 131–140. doi:10.3758/BF03196014
- Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 718–734. doi:10.1037/a0013899
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59(4), 771–785. doi:10.1006/anbe.1999.1377
- Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, 112(2), 494–508. doi:10.1037/0033-295X.112.2.494
- Hayward, D. A., & Ristic, J. (2015). Exposing the cuing task: The case of gaze and arrow cues. *Attention, Perception, & Psychophysics*, 77(4), 1088–1104. doi:10.3758/s13414-015-0877-6
- Helbing, D., Farkas, I., & Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature*, 407(6803), 487–490. doi:10.1038/35035023
- Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19(4), 215–241. doi:10.1016/S1090-5138(98)00018-X
- Hietanen, J. K. (1999). Does your gaze direction and head orientation shift my visual attention? *Neuroreport*, 10(16), 3443–3447. doi:10.1097/00001756-199911080-00033
- Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage*, 33(1), 406–413. doi:10.1016/j.neuroimage.2006.06.048
- Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L., Feinberg, D. R., & Tiddeman, B. P. (2010). Facial cues of dominance modulate the short-term gaze-cueing effect in human observers. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1681), 617–624. doi:10.1098/rspb.2009.1575
- Joseph, R. M., Fricker, Z., & Keehn, B. (2015). Activation of frontoparietal attention networks by non-predictive gaze and arrow cues. *Social Cognitive and Affective Neuroscience*, 10(2), 294–301. doi:10.1093/scan/nsu054
- Kameda, T., & Hastie, R. (1999). *Social sharedness and adaptation: Adaptive group decision heuristics*. Paper presented at 17th Subjective Probability, Utility, and Decision Making Conference, Mannheim, Germany.
- Kameda, T., Takezawa, M., Tindale, R. S., & Smith, C. M. (2002). Social sharing and risk reduction: Exploring a computational algorithm for the psychology of windfall gains. *Evolution and Human Behavior*, 23(1), 11–33. doi:10.1016/S1090-5138(01)00086-1
- Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology*, 55, 623–655. doi:10.1146/annurev.psych.55.090902.142009
- Langton, S. R., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, 6(5), 541–567. doi:10.1080/135062899394939
- Lassalle, A., & Itier, R. J. (2013). Fearful, surprised, happy, and angry facial expressions modulate gaze-oriented attention: Behavioral and ERP evidence. *Social Neuroscience*, 8(6), 583–600. doi:10.1080/17470919.2013.835750
- Liuzza, M. T., Cazzato, V., Vecchione, M., Crostella, F., Caprara, G. V., & Aglioti, S. M. (2011). Follow my eyes: The gaze of politicians reflexively captures the gaze of ingroup voters. *PLoS One*, 6(9), e25117. doi:10.1371/journal.pone.0025117
- Milgram, S., Bickman, L., & Berkowitz, L. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology*, 13(2), 79–82. doi:10.1037/h0028070

- Mueller, D. C. (1989). *Public choice II: A revised edition of public choice*. New York, NY: Cambridge University Press.
- Narang, S., Best, A., Randhavana, T., Shapiro, A., & Manocha, D. (2016). *PedVR: Simulating gaze-based interactions between a real user and virtual crowds*. Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, (pp. 91–100). doi:10.1145/2993369.2993378
- Nuku, P., & Bekkering, H. (2008). Joint attention: Inferring what others perceive (and don't perceive). *Consciousness and Cognition*, 17(1), 339–349. doi:10.1016/j.concog.2007.06.014
- Perrett, D. I., Hietanen, J. K., Oram, M. W., Benson, P. J., & Rolls, E. T. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 335(1273), 23–30. doi:10.1098/rstb.1992.0003
- Pratt, S. C. (2005). Quorum sensing by encounter rates in the ant *Temnothorax albipennis*. *Behavioral Ecology*, 16(2), 488–496. doi:10.1093/beheco/ari020
- Ruckstuhl, K. E., & Neuhaus, P. (2000). Sexual segregation in ungulates: A new approach. *Behaviour*, 137(3), 361–377. doi:10.1163/156853900502123
- Ruckstuhl, K. E., & Neuhaus, P. (2002). Sexual segregation in ungulates: A comparative test of three hypotheses. *Biological Reviews of the Cambridge Philosophical Society*, 77(1), 77–96. doi:10.1017/S1464793101005814
- Sato, W., Okada, T., & Toichi, M. (2007). Attentional shift by gaze is triggered without awareness. *Experimental Brain Research*, 183(1), 87–94. doi:10.1007/s00221-007-1025-x
- Schloegl, C., Kotrschal, K., & Bugnyar, T. (2007). Gaze following in common ravens, *Corvus corax*: Ontogeny and habituation. *Animal Behaviour*, 74(4), 769–778. doi:10.1016/j.anbehav.2006.08.017
- Seeley, T. D., & Buhrman, S. C. (1999). Group decision making in a swarm of honeybees. *Behavioral Ecology and Sociobiology*, 45(1), 19–31. doi:10.1007/s002650050536
- Shepherd, S. V. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, 4(5), 1–13. doi:10.3389/fnint.2010.00005
- Shiwakoti, N., & Sarvi, M. (2013). Understanding pedestrian crowd panic: A review on model organisms approach. *Journal of Transport Geography*, 26, 12–17. doi:10.1016/j.jtrangeo.2012.08.002
- Sorkin, R. D., Hays, C. J., & West, R. (2001). Signal-detection analysis of group decision making. *Psychological Review*, 108(1), 183–203. doi:10.1037/0033-295X.108.1.183
- Sorkin, R. D., West, R., & Robinson, D. E. (1998). Group performance depends on the majority rule. *Psychological Science*, 9(6), 456–463. doi:10.1111/1467-9280.00085
- Sumpter, D. J. (2006). The principles of collective animal behaviour. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 361(1465), 5–22. doi:10.1098/rstb.2005.1733
- Sumpter, D. J., & Pratt, S. C. (2009). Quorum responses and consensus decision making. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1518), 743–753. doi:10.1098/rstb.2008.0204
- Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention ensemble perception of a crowd's gaze. *Psychological Science*, 25(10), 1903–1913. doi:10.1177/0956797614544510
- Tipples, J. (2006). Fear and fearfulness potentiate automatic orienting to eye gaze. *Cognition & Emotion*, 20(2), 309–320. doi:10.1080/02699930500405550
- Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition*, 2(1), 59–87. doi:10.1080/13506289508401722
- Ward, A. J., Sumpter, D. J., Couzin, I. D., Hart, P. J., & Krause, J. (2008). Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences*, 105(19), 6948–6953. doi:10.1073/pnas.0710344105
- Zhang, Z., Tang, Y., Zhang, X., & Zhang, Z. (2015). The influence of dynamic and static gaze-cueing upon attention shift: Role of movement cues. *Chinese Journal of Applied Psychology*, 21(3), 195–202. doi:10.3969/j.issn.1006-6020.2015.03.001