RESEARCH ARTICLE

Emotional cues and social anxiety resolve ambiguous perception of biological motion

Hörmet Yiltiz

by mistaking an approaching biological motion as receding; this bias is higher than perceiving the reverse (perceiving receding motion as approaching) (Weech et al. 2014; Yiltiz and Chen 2015).

Social and emotional information extracted from (biological) motion may significantly influence the visual processing of dynamic information associated with the motion stimuli (Brooks et al. 2008; Heenan and Troje 2014, 2015; Pavlova 2011; Schouten et al. 2010; Van de Cruys et al. 2013). On the one hand, PLW contained social information such as age (Insch et al. 2012), gender (Cutting and Kozlowski 1977; Pollick et al. 2005), and emotional state (Clarke et al. 2005). On the other hand, the internal emotional states of the observers could bias the perception of point-light walkers. Emotional factors that can induce bias include empathy level (Yiltiz and Chen 2015), anxiety level (Van de Cruys et al. 2013), and inhibition capacity (Heenan and Troje 2015).

Investigations into facing-the-viewer bias are important. First, it is ecologically important to protect humans and animals in dangerous environments, by actively and effectively avoiding approaching threatening predators or objects. The perception of PLW (with bi-stability in direction) imposes an example of perceptual decision under uncertainty. Judgment made under uncertainty can result in over-or underestimations. To cope with the perceptual decision with uncertainty, humans favor an "error management" bias toward making the less costly error. For instance, the costs of false alarm of wasting time by estimating too early the arrival time of the approaching object are relatively low compared to the costs of misses (i.e., not being prepared for an approaching object.) (Haselton et al. 2009; Holbrook et al. 2014). Second, examining the size of the facing-the-viewer bias effect will quantitatively reveal how the underlying cognitive and individual traits modulate and relate to this perceptual bias, and how the bias is modulated in the framework of "error management", hence to help humans make decisions and secure the chance for survival and reproduction (Haselton et al. 2009; Holbrook et al. 2014).

Humans and some animals are endowed with the ability to rapidly take in complex sensory arrays or events by means of perceptual averaging, i.e., 'ensemble coding' (Ariely 2001; Hunt et al. 2008). Being able to extract statistical properties—such as the mean of numbers, sizes, spatial layout, or even emotions from a set of simultaneously viewed objects (Alvarez 2011; Ariely 2001; Chong and Treisman 2003; de Gardelle and Summerfield 2011; Haberman et al. 2009; Haberman and Whitney 2007; Walker and Vul 2014), or a series of auditory beeps (Miller et al. 2013; Piazza et al. 2013), can greatly expedite perceptual decisions, as well as social cognition, in everyday life PLW perception is a high-level cognitive construct, which might differ from the percept upon other forms of visual (apparent) motion. Therefore, the FTV effect could be modulated by the 'correspondence' of perceived high-level cognitive and emotional factors but outside emotional information conveyed by PLW stimuli themselves. We hypothesized that the acquired mean emotion stimuli—facial figures (given by either a single face or multiple faces), though task-irrelevant, could be used by an ensemble coding—like processing, to modulate the FTV effect on PLWs but not on the percept of low/middle level random dot motion (containing non-sociobiological meanings). Moreover, this modulation (if exists) is dependent on the ability of association the background emotional information onto the PLWs. Previous evidence demonstrated the modulating effect of anxiety on the perceptions about PLW. Individuals with high anxiety level are more sensitive to emotional cues with higher valence (Bar-Haim et al. 2007; Fox et al. 2002; Gray et al. 2009; MacLeod et al. 1986; Mathews and MacLeod 1985; Singer et al. 2012). We further hypothesized that by presenting the task-irrelevant emotional facial figures, observers with high anxiety would use the ensemble coding to extract the mean emotion, and they would be subject to a stronger facing-the-viewer bias than individuals with low anxiety would be.

Based on this background, we implemented two experiments to examine the perception of PLW under a single facial emotional cue (Experiment 1) and multiple facial emotional cues (Experiment 2). We investigated individual differences in FTV bias, due to different levels of social anxiety traits (low vs. high anxiety). To our best knowledge, this study showed the first empirical evidence that task-irrelevant but subjectively (face) emotional cues, mediated by individual cognitive abilities, could modulate FTV bias.

Experiment 1

We investigated how a single task-irrelevant facial emotion cue as background affected the perception of dominant direction (facing-the-viewer bias) of an ambiguous point-light walker, and how social anxiety level interacts with emotional perception to modulate the facing-the-viewer bias.

Method

Participants

Thirty-two undergraduate students (18 female; 14 male) from Peking University participated in this experiment. Participants were aged from 18 to 25 (mean: 21.9, SE = 2.0), with either normal or corrected to normal vision, none of whom reported any neurological symptoms. Observers were not informed about the purposes of the study, and were paid for their participation. The experiment was performed in compliance with all institutional guidelines established by the Academic Affairs Committee, Department of

Psychology at Peking University. All observers provided written informed consent, according to institutional guidelines and the Declaration of Helsinki.

Visual stimuli and equipment

Visual stimuli included target stimuli (foreground) and a task-irrelevant facial emotion cue (background). The target stimuli included either a point light walker or circular random dots (Fig. 1). For each trial, we randomly chose an azimuth rotation angle for the point-light walker as either $90°$ or $-90°$, and counterbalanced both types of PLW. We then displayed an animation of PLW, with a full walking cycle of 1300 ms. We played the video with 130 frames on Cathode Ray Tube (CRT) monitor at a vertical refresh rate of 100 Hz (10 ms per frame), with a resolution of 1024×768 (pixels^2). The random dots stimuli consisted of 1000 random dots in a circular area within an imaginary contour diameter equal to the height of the PLW. Approximately half of the random dots moved towards the center point of this circular (centripetal motion), while the other half moved away from the center (centrifugal motion). Each dot was 3 pixels wide (approximately 0.06 degrees of visual angle).

For the task-irrelevant facial emotional cue, we presented human face figures in the background, in the same area as the point-light walker. The face stimuli were categorized into happy, neutral, or angry faces using a standard appraisal procedure/scoring procedure (Gong et al. 2011). We selected the facial stimuli with high rating consistency (over 80%) across evaluators. The target stimulus was either a point light walker or a display of circular random dots (Fig. 1). The 3D position data to render the PLW were collected from the CMU Graphics Lab Motion Capture Database (http://mocap .cs.cmu.edu). The PLW consisted of 13 dots, representing key joints of the body including the head, the shoulders, the elbows, the hands, the hips, the knees, and the feet (Ahlström et al. 1997). The position for PLW was always upward and extended approximately 6 (high) $\times$ 4 (wide) degrees of visual
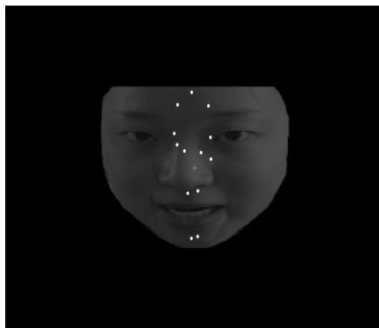

Fig. 1 A point-light walker is used as biological motion stimuli. A task-irrelevant facial emotional cue is presented as background

angle, viewed at 90 cm between the eye and the monitor. The PLW had a direction that was orthogonal to the monitor surface (along the azimuthal axis), inducing an ambiguous perception of the motion's direction (depth).

Dots for the PLW had an average luminance of 14.88 cd/m^2 on black screen background (0.17 cd/m^2) or on a background face (20 cd/m^2). The emotional face stimuli were selected from Chinese Facial Affective Picture System (Gong et al. 2011). The computer programs were developed with Matlab (Mathworks Inc.) and the Psychophysics Toolbox (Brainard 1997; Pelli 1997). The test booth was semi-anechoic and dimly lit throughout the experiment, with ambient luminance of 0.05 cd/m^2. Throughout the experiment, observers used a chin-rest to maintain a fixed viewing distance.

Design and procedure

We separated observers into a high social anxiety group and a low social anxiety group, based on the mean scores from the Chinese version of the Liebowitz Social Anxiety Scale (LSAS) (He and Zhang 2004; Pan et al. 2006). The Liebowitz Social Anxiety Scale (LSAS) was the first clinician-administered scale developed for the assessment of fear and avoidance associated with social phobia (Rytwinski et al. 2009; Safren et al. 1999). The LSAS consists of two subscales, measuring the level of social phobia and the frequency of avoiding a social activity or a typical social scene. Examples of questionnaire items include: "Making a phone call in a public area" and "Expressing di erent opinions to someone unfamiliar." We adopted a full factorial design with two anxiety levels (high anxiety vs. low anxiety), four background visual conditions (happy, neutral, anger, and baseline without facial stimulus), and two target stimuli type (a point light walker or random circular dots). The between-participants variable was social anxiety; the within-participants variables were valence of facial emotion cue (in each trial) and target stimuli type (in each block). We had four visual background conditions (happy, neutral, anger, and baseline without facial stimulus), and two target stimuli type (a point light walker or random circular dots) and four repetitions of each sub-condition. Therefore, the number of trials was 32. Each trial lasted for 70 s. Participants took about 50 min to finish the experiment.

Prior to the experiment, observers watched a short video demo to better understand the task. The task was to focus on upon a central cross and report the dominant walking/facing direction (either inward or outward) of the point light walker, and the dominant direction of the random dots motion, this was done by holding down a corresponding key ('up' arrow for inward motion in PLW and receding of random dots, 'down' arrow for outward motion in PLW and approaching of random dots). The duration for each trial was 70 s. The

first 10 s were dropped to prevent potential initial response bias and to allow the observer to establish bi-stability of the ambiguous motion. Data from the last 60 s were used for analysis.

During the experiment, PLW and random dots motion (RDM) were presented in blocks, while the background visual conditions were presented using randomized trial-by-trial. Each block consisted of 16 trials, with each of the four emotional valences repeated four times (Fig. 2). Observers rested for at least 30 s every five trials. To examine the validity of the subjective appraisal of emotional valences of facial cues (including the no-facial-image baseline condition), we asked observers to rate emotional valences for the faces on a Likert scale from 1 to 7 (i.e., 1-most angry, 7-most happy) after the experiment.

Results for Experiment 1

The dependent variable is the duration of perceived motion direction, which characterizes the stability of a dominant perceived direction over a non-dominant direction. Because the perceived direction is the result of a resolution of ambiguity (approaching vs. receding), the perception is bi-stable. The dominant perceived direction changes over time, switching between two perceptions. We excluded two observers' data for analysis. One of these experienced a ceiling effect, perceiving all the PLW as moving outward. That person did not perceive the bi-stability of the ambiguous stimuli. The second excluded observer did not complete the experiment due to an accidental failure of the response devices. Due to large individual variances, we normalized the duration within observers (the duration in each trial was divided by the mean duration from all trials). Further, the study's LSAS scores did not exactly follow normal distribution (confirmed by SPSS normality plots with tests). Because of this, due to the large individual variances, we divided subjects into two groups using split-half (median) method. We grouped participants with mean LSAS score greater than the median as the high anxiety group; the remaining participants were grouped as the low anxiety group.

We used a repeated measure analysis of variance (ANOVA) to analyze the data. Stimuli types (point-light walker vs. random dot motion), background visual information (negative, neutral, positive, and baseline black visual background), and perceived direction (approaching vs. receding) were used as within subject factors. Social anxiety group (high vs. low) was a between-subject factor and the dominant duration was a dependent variable. The main effect of the background information was statistically significant, $F(3,84) = 2.771$, $p < 0.05$. The mean normalized duration of dominant perceptions for negative, neutral, happy, and baseline conditions were 1.038 ± 0.027, 1.013 ± 0.019, 0.954 ± 0.012, and 1.042 ± 0.024, respectively. The duration under the happy condition (0.954) was significantly shorter than in the baseline condition (1.042), $p < 0.05$ (see Fig. 3). The main effects of stimuli types and dominant directions were non-significant, $F(1,28) = 0.089$, $p = 0.768$ and $F(1,28) = 3.944$, $p = 0.057$. The interaction between stimuli type (PLW vs. random dots) and perceived direction was statistically significant, $F(1,28) = 106.572$, $p < 0.001$. We also found a significant three-way interaction between stimuli type,7(s. r)/T1a7[(000S(nt).5(05 Tc.01f.005 Tof t)05. The m2(he(o g)

interactions by exploring the statistical results within each stimulus category (PLW vs. random dots).

To further investigate the stimulus/task effects, we applied two separate repeated measures ANOVA for PLW and random dot presentations, with normalized dominant duration as the dependent variable and the same set of predictor variables (emotion valences, perceived direction, and social anxiety group) as the independent variables.

For the PLW task (with the point-light walkers as the visual stimuli), the main effect of perceived direction was statistically significant, $F(1,28) = 107.243$, $p < 0.001$, where a Bonferroni corrected pairwise test showed the normalized duration for receding (facing away from the viewer) (0.787, SE = 0.0.023) was significantly lower than for approaching (1.235, SE = 0.023). This shows a typical FTV bias. The main effect of the background visual information was non-significant, $F(3,84) = 1.321$, $p = 0.273$. The interaction between visual information and anxiety groups was statistically significant, $F(3,84) = 4.332$, $p < 0.01$. Simple main effect ANOVA tests showed that, for neutral emotion cues, the normalized duration for high social anxiety group was significantly longer (1.125) than the duration in the low social anxiety group (0.914), $F(1,28) = 9.86$, $p < 0.01$. For appraisal of the neutral cues (Likert points 1–7), the mean scores for the high anxiety group was 3.53 (0.15); the mean for the low anxiety group was 4.06 (0.18), $F(1,29) = 4.923$, $p < 0.05$. In contrast, when the emotional cue has a positive valence, the normalized duration for the high social anxiety group was significantly shorter (0.919) than the duration in the low social anxiety group (1.010), $F(1,28) = 6.26$, $p < 0.05$.

The mean dominant durations for receding and approaching perceptions were 0.834 (0.038) and 1.183 (0.032), respectively, for the low social anxiety group; the durations for receding and approaching perceptions were 0.761(0.035) and 1.283 (0.030) for the high social anxiety group. The interaction between perceived direction and social anxiety group was statistically significant, $F(1,28) = 3.881$, $p < 0.05$. This result indicates a strong facing-the-viewer bias. The dominant duration of receding (facing away from the viewer) motion was 0.798 (0.026); the duration of approaching was 1.233 (0.022), $F(1,28) = 93.480$, $p < 0.001$. The difference in the normalized duration between perceiving receding and approaching motion was higher for high social anxiety group (0.52, SE = 0.03) than the difference in duration for low social anxiety group (0.35, SE = 0.03), $F(1,28) = 3.706$, $p < 0.05$.

For the non-sociobiologically relevant task (random dots as the visual stimuli), there was no two-way or three-way interaction between emotional cues and/or response type when the 'anxiety' factor was considered. The main effect of emotional valence and perceived direction was significant, with $F(3,84) = 4.802$, $p < 0.01$ and $F(1,28) = 40.11$,

$p < 0.001$. For emotional valences, a Bonferroni corrected pairwise test showed the normalized duration under a positive emotional cue (0.95, SE = 0.01) was significantly shorter than baseline (1.09, SE = 0.03), $p < 0.001$. For perceived direction, a Bonferroni corrected pairwise test showed the normalized duration for perceiving receding random dots (1.17, SE = 0.02) was significantly longer than for perceiving approaching random dots (0.86, SE = 0.03), $p < 0.001$. However, no significant interactions between the valence of emotional cues and perceived direction was found, $F(3,93) = 0.543$, $p = 0.654$. There were no group differences between the low and high anxiety groups, $F(1,28) = 2.896$, $p = 0.099$.

Experiment 2

Experiment 1 examined the influence of task-irrelevant single emotional cues (a single figure of human face at various emotional valences) on perception of ambiguous visual motion, using a point-light walker PLW for biological motion stimulus and random dots for non-biological motion. However, visual environments in the real world are much more complex. For example, when we see a friend walking far away from us among a crowd of people, we may be inclined to welcome the others approaching us if we feel safe. Here, we assessed whether the perceived averaged (mean) valence of facial cues, with an ensemble coding-like processing, could modulate the FTV bias for a group of PLWs.

Method

Experiment 2 included 30 undergraduate students (16 female; 14 males) from Peking University, aged from 19 to 24 (21.4, SD = 2.7). The experimental design was like Experiment 1, except we increased the number of Tc0 Tw(-)002(a

Fig. 4 Eight task-irrelevant facial emotion cues (located on an imaginary circle). First row shows three PLWs (left) with a crowd of faces; and random dots (right) with a crowd of faces. The second and third rows show an example of an average valences of facial emotion cues as: negative, with six angry faces and two happy faces (second row left); positive, with six happy faces and two angry faces (second row right); neutral, with eight neutral faces (third row left); and the baseline, without any facial emotion cues but a black background
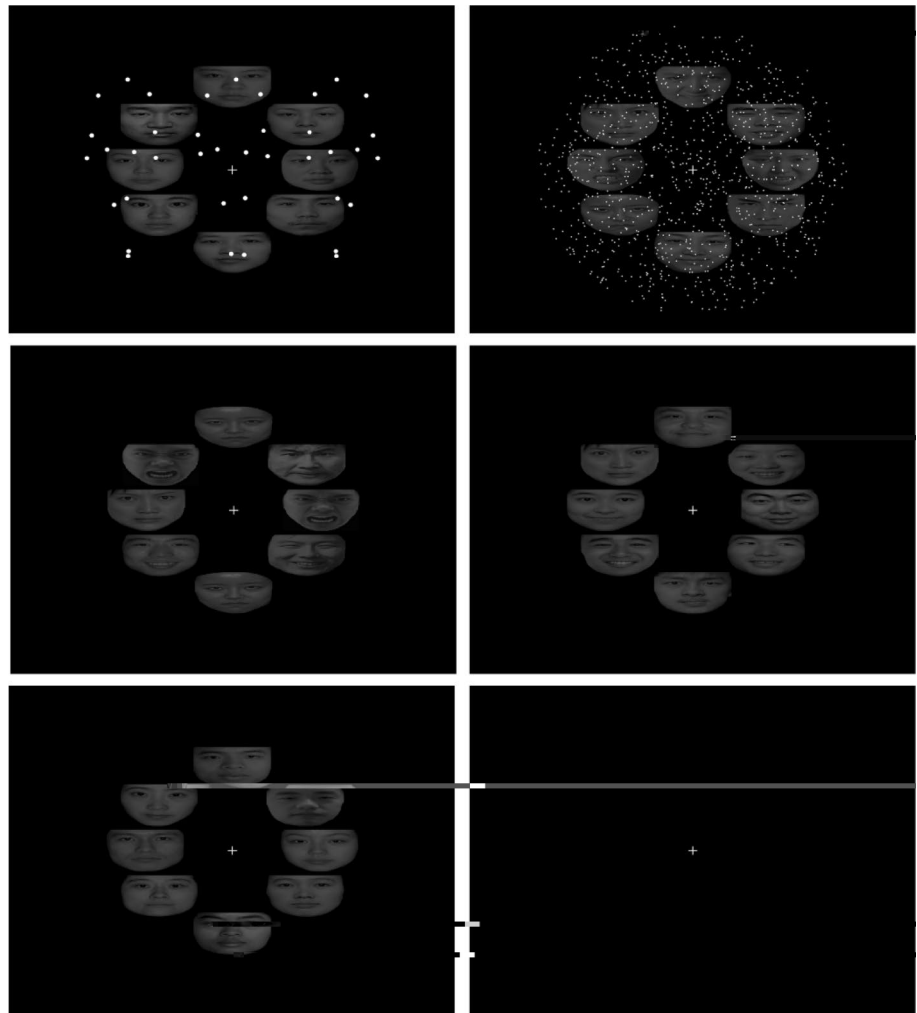
Three PLWs were presented side-by-side on the screen; the middle point-light walker was at the center of the screen, and the other two was set at 650 mm (roughly 36 degrees of visual angle) from the center. To prevent duplication of all the point-light walkers, at the very beginning of presentation (300 ms), the facing directions of the two surrounding PLWs (i.e., left and right PLW) were tilted randomly, with a 20° angle either to the right or left along the azimuth axis. All three point-light walkers had the same walking cycle, starting at a random time-point in their cycle (phases randomized). We used three PLWs to balance the visual display areas between background multiple faces and the target of biological motion stimuli, so that the ratio of focusing PLWs relative to the background faces in Experiment 2 was tantamount to the one in Experiment 1, and the relative saliency between the background faces and foreground PLWs was largely controlled. However, the perceptual identity and quality of each PLW was the same, which did not affect the bias of the perceived direction of PLWs. Participants reported the dominant walking/facing direction (either inward or outward) of the three PLWs or random dots.

To examine the validity of the subjective appraisal of emotional valences of facial cues (including the no-facial-image baseline condition), we asked observers to rate emotional valences for the faces on a Likert scale from 1 to 7 (i.e., 1-most angry, 7-most happy) after the experiment.

The remaining configurations were the same as in Experiment 1.

Results for Experiment 2

A similar set of analysis was conducted for Experiment 2 as for Experiment 1. We applied a repeated measure ANOVA with stimulus type (PLWs vs. random dots), average emotion valences of the task-irrelevant background visual cues (negative, neutral, positive, and baseline), and perceived direction (approaching vs. receding) as the within-subject factors. The social anxiety group (high social anxiety vs. low social anxiety) was the between-subject factor.

There were no significant main effects detected for stimuli type, $F(1,29) = 0.127$, $p = 0.725$; for visual cues, $F(3,87) = 0.682$, $p = 0.565$; and for perceived direction,

$F(1,29) = 3.128$, $p = 0.087$. The interaction between visual cues and social anxiety group was statistically significant, $F(3,87) = 2.91$, $p < 0.05$. The interaction between stimuli type and social anxiety group was borderline significant, $F(1,29) = 3.652$, $p = 0.06$. The interaction between perceived direction and social anxiety group was also statistically significant, $F(1,29) = 4.949$, $p < 0.05$. The interactions between target stimuli type and perceived direction was statistically significant, $F(1,29) = 57.16$, $p < 0.001$. The interaction between visual cues and perceived direction was borderline significant, $F(3,87) = 2.500$, $p = 0.065$. A three-way interaction between stimulus type, background visual cues, and social anxiety group was statistically significant, $F(3,87) = 3.217$, $p < 0.05$ (Fig. 5). Next, we explore these critical interactions, by investigating the statistical results within each type of stimuli (PLW vs. random dots).

We performed separate repeated measure ANOVA tests for the PLW and random dots stimuli, and examined the effects of the mean valences of facial emotion cues (negative, neutral, positive valences, and baseline)*, the perceived motion, and social anxiety group (low vs. high), on the normalized duration of dominant perception.
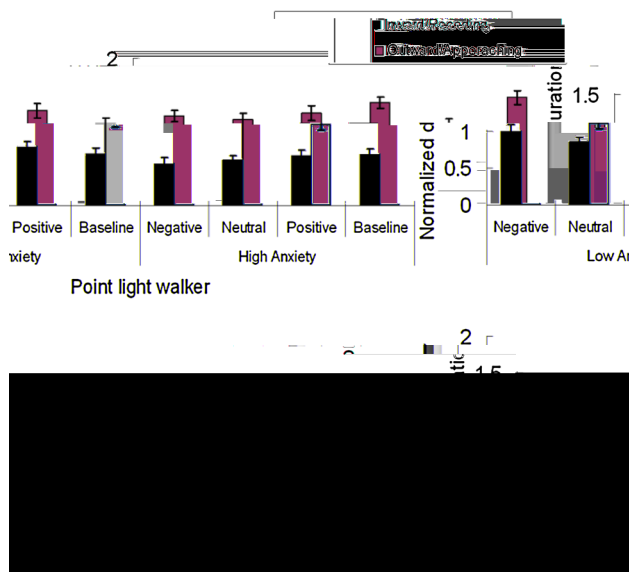

Fig. 5 Results for Experiment 2 (effect of mean emotion valence using a group of task-irrelevant facial emotion cues). The black bars indicate normalized durations for perceiving the dominant direction of a point-light walker as walking away from the viewer (inward/receding) or perceiving the dominant direction of the coherence motion (of random dots) as centripetal (moving inwards) when observing ambiguous visual stimuli. The gray bars indicate normalized durations of dominant perceptions for the approaching/outward direction. As the figure shows, participants reported a facing-the-viewer bias (perceiving the point-light-walker as approaching dominates); as with Experiment 1, this bias was magnified for individuals with higher anxiety

For the PLWs task, the three-way interaction between emotion valences, social anxiety groups, and perceived direction was significant, $F(3,87) = 2.88$, $p < 0.05$. A subsequent simple main effect analysis showed that the normalized duration of the dominant perception of 'approaching' is higher than the perception of 'receding' for all valences of emotional cues (negative, neutral, positive, and baseline), $p < 0.001$. The main effect of perceived direction is significant, $F(1,29) = 66.42$, $p < 0.001$. A Bonferroni corrected pairwise test showed that the normalized duration for the perception of approaching (1.25, SE = 0.029) is significantly longer than the duration of the perception of receding (0.741, SE = 0.026), $p < 0.001$, showing a typical facing-the-viewer bias as found in Experiment 1. The interaction between the task-irrelevant facial emotion cues and social anxiety was significant, $F(3,87) = 3.558$, $p < 0.05$.

A further simple main effects analysis found no significant group effect when emotion valence was neutral, positive, or when no facial emotion cues were presented (baseline), $p > 0.05$. However, when emotion valence was negative, the normalized duration for the high social anxiety group (0.912) was significantly shorter than the duration (1.144) for the low social anxiety group, $F(1,29) = 5.41$, $p < 0.05$. The interaction between perceived direction and social anxiety group was statistically significant, $F(1,29) = 6.20$, $p < 0.05$. A further analysis of simple main effect showed that, the low anxiety group reported a longer normalized duration of a dominant 'receding' perception (0.842) than the high anxiety group did (0.639). $F(1,29) = 3.72$, $p = 0.064$. However, for the dominant 'approaching' perception, this contrast between the two groups was non-significant, $F(1,29) = 2.73$, $p = 0.109$, with dominant durations of 1.197 for low anxiety and 1.307 for high anxiety. Also, a Bonferroni corrected pairwise test showed that the difference between normalized durations of perceived receding and approaching is higher for high social anxiety group (0.67, SE = 0.05) than the difference in low social anxiety group (0.36, SE = 0.04), $p < 0.05$.

The interaction between perceived direction and social anxiety group was significant in conditions where the average valence of emotion cues was either negative or neutral, $F(1,29) = 5.11$, $p < 0.005$ and $F(1,29) = 16.57$, $p < 0.001$, respectively. That is, facing-the-viewer bias was more readily observed when the average valence of facial emotion cues was negative and neutral. However, the interaction between perceived direction and anxiety group was not significant when the valence was positive or at a baseline level, $p > 0.05$.

For the random dot motion, we found no significant two-way or three-way interactions between any of the task factors with random dots motion. The main effect of perceived direction was significant, $F(1,29) = 22.34$, $p < 0.001$. The normalized duration for perceived receding random dot motion (1.18, SE = 0.036) was significantly longer than

13

the duration of perceived approaching random dots motion (0.816, SE = 0.042).

Likert scale rating showed that in Experiment 2, the main effect of the perceived emotion valence was significant, $F(3,81) = 6.934$, $p < 0.001$. The mean scores for negative, neutral, positive, and baseline were 3.395(0.252), 3.832(0.124), 4.797(0.279), and 3.720(0.166), respectively. A Bonferroni corrected comparison revealed significant differences in the appraisal of negative vs. positive cues, $p < 0.05$; neutral vs. positive cues, $p < 0.05$; and positive vs. baseline cues, $p < 0.05$.

We summarized and compared the key findings of the two experiments in Table 1.

Discussion

This study investigated how task-irrelevant emotional cues, presented as background visual information, affected the processing of ambiguous visual motion conveyed through point light walkers (biological motion) or random dots (non-biological motion). Emotional valences (happy, anger, or neutral) were presented as a single facial image (Experiment 1) or as an average of a group of faces (Experiment 2). The results indicated that both a single emotional cue and an average of a group of emotional cues affected how a dominant perception of PLW, but not random dots motion, was resolved. Specifically, negative and neutral valences of task-irrelevant facial emotion cues contributed to FTV bias. Moreover, this modulation effect was higher for the group with higher social anxiety than for the group with low social anxiety. This finding echoes a previous hypothesis that perceiving an approaching biological motion as receding is

Table 1 Key results with statistical significances (main effects or interaction effects) in single face and multiple face conditions

Factors	Single face	Multiple faces
Combined analysis		
Type (PLW vs.RDM) × Visual × anxiety	N	*
Type (PLW vs.RDM) × Resp × anxiety	0.063	**
Visual × anxiety	**	*
Type × anxiety	N	0.066
Resp × anxiety	N	*
Visual × Resp	0.060	0.065
Type × Resp	***	**
Visual × Type	*	*
PLW		
Resp	(Facing viewer bias)***	(Facing viewer bias)***
Visual	N	N
Visual × anxiety	(Neutral emotion)**	(Neutral, negative emotion)**
Resp × anxiety	*	*
RDM		
Resp	***	***
Visual	**	N
Visual × anxiety	N	N
Resp × anxiety	N	N

'Type' shows the stimuli categories (point light walker—PLW and random dots motion—RDM); 'visual' shows the conditions of background visual information (happy, angry, neutral, and baseline). 'Anxiety' indicates lower or higher social anxiety. 'Resp' represents the two dominant response directions: approaching and receding. The numbers indicate borderline significance. For RDM, there was no interaction effect among the given factors. In both single face and multiple face conditions, the facing-the-viewer bias was larger in individuals with higher anxiety compared to the group with lower anxiety. This bias was robust with 'neutral' and 'negative' background facial valences in the 'multiple faces' tests (Experiment 2), but was only seen with the 'neutral' face condition for in the 'single face' tests (Experiment 1)

N there was no statistical significance detected

*Shows that the p value is less than 0.05

**Indicate a p value less than 0.01

***Indicate a p value less than 0.001

riskier than the opposite (perceiving a receding figure as approaching), especially when this motion is accompanied by potentially menacing emotional information (angry faces) (Heenan and Troje 2015; Van de Cruys et al. 2013). Anxious individuals display an attentional bias towards more threatening stimuli (Bar-Haim et al. 2007; MacLeod et al. 1986; Mathews and MacLeod 1985).

We detected this facing-the-viewer bias even when the facial images had neutral valence. Previous studies found that observers with higher social anxiety perceived neutral emotional information in anxiety-inducing environments (such as crowds) as menacing/threatening (Gilboa-Schechtman et al. 1999). A post-hoc comparison found that individuals with high anxiety traits evaluated the neutral faces as more 'angry' than the low anxiety group did. As a result, even for PLWs with a neutral emotional background, observers that rated the faces as more threatening tended to perceive the dominant direction as facing the viewer (Heenan and Troje 2014, 2015; Van de Cruys et al. 2013).

Nevertheless, the above illusory facing-the-viewer bias when seeing 'negative' facial cues did not occur when we showed only a single facial cue (Experiment 1). The specificity of this modulation effect when comparing single and group facial cues shows that attentional focus and sensitivities to visual cues in different visual fields and across different levels (local vs. global features) might play an important modulating role. Previous studies identified qualitative changes in sensitivity for perceiving biological motion across the visual field. Sensitivities to the PLWs were enhanced for images in the central vision and were reduced for the peripheral vision (Ikeda et al. 2005). Indeed, this performance imbalance has been seen in several studies on visual perception (Beard et al. 1997; Daniel and Whitteridge 1961; Levi et al. 1985, 2000; Westheimer 1982).

Here, the primary function of perceiving biological motion is to carefully and quickly retrieve detailed information about the identity, intentions, and affective state of another individual (conveyed through PLWs). The visual system adopted selective attention, with flexible mechanisms for processing perceptions of biological motion and task-irrelevant figures (Battelli et al. 2003; Beintema and Lappe 2002; Cavanagh et al. 2001; Giese and Poggio 2003; Neri et al. 1998; Thornton et al. 2002). In Experiment 1, a single task-irrelevant facial image on the background is located at the screen's center. In Experiment 2, a group of faces were aligned on a virtual circle (at parafoveal or peripheral region), with the given spatial parameters, those facial figures crowded with each other (Bouma 1970; McNair et al. 2017; Pell et al. 2007; Shim et al. 2008; Whitney and Levi 2011). Participants might invest different attentional engagements for perceiving facial layouts. For the single (and only salient) face (Experiment 1), more attentional focus was directed to the one figure display; less attention was directed

to the target motion stimuli (PLWs). As such, observers had high probability to be less efficient in aligning the emotional valence on the background on the perceptual decision of visual motion. On the other hand, evidence has shown the sensory processing of facial information (facial expression) is hierarchical, with the lower cortical area (such as V1) having a spatially smaller receptive field. As such, it responds more to simple/local features of the stimuli compared to those in higher areas, such as inferotemporal cortex responding to the complex/global features (emotion expression) (Felleman and Van Essen 1991). Our visual system is usually especially attuned to local features in the bottom half of the point light walker during biological motion perception (Troje and Westhoff 2006). This local processing was more engaged in the single face display. However, during a longer presentation of facial stimuli, observers showed invariant (global) representation of the high-level facial expression with large scale of display (with a crowd of faces are shown on the background); however, attentional resources were also distributed among them, with less attention directed to individual faces(Ying and Xu 2017). Therefore, in our case, multifocal attention (spreading across the multi-face displays) was limited but focused attention may not be strictly necessary for ensemble-like perception to occur (Cavanagh and Alvarez 2005; Whitney and Levi 2011), observers could still focus on the global outline (mean emotion) of the faces by saving the benefits from ensemble coding. However, we should be cautious in above arguments, in viewing that the presumed local vs. global processing as shown in the literature, has not been implemented and strictly tested using the stimuli (configurations) in current experimental setting.

Attentional control theory has been used to interpret the link between anxiety, inhibition capacity, and the perception of ambiguous visual stimuli (Heenan and Troje 2015). Anxious individuals display an attentional bias towards more threatening stimuli (Bar-Haim et al. 2007; MacLeod et al. 1986; Mathews and MacLeod 1985). Individuals with higher anxiety were less able to suppress the salient emotional properties of a single (large) face (Experiment 1); however, they readily inhibited the processing of local features of individual faces in a multi-face display and still efficiently

bi-stability rather than the other direction. The perception of motion direction using point-light walkers was in the direction coming out from the center (facing-the-viewer). Human observers tend to process visual stimuli in the central visual field more efficiently than in the peripheral field. This resulted in a perception of dominant receding (centripetal) motion (Aaen-Stockdale et al. 2008). Unexpectedly, we found that the presence of the single facial image affected the perceived dominant direction in random dots motion; however, the group of facial images did not have this effect. This further confirms that the central visual cues (a single face) played a major role in modulating the lower-level perception of random dots motion. Nevertheless, we could not find a specific modulation effect based on social anxiety level in the random dots motion experiment. This results implies that individual cognitive abilities, including anxiety level, is specifically aligned with life-relevant sociobiological motion stimuli such PLWs (Heenan and Troje 2014, 2015), but is not aligned with non-biological motion stimuli.

Unlike other approaches in ensemble coding, which directly reveal that many objects were pooled into a summary, in our current approach, we did not pool directly the facial stimuli into a summary representation. This is probably due to the task (discriminating the PLW rather than the faces) and the stimuli sizes (larger than the stimuli in typical studies using ensemble coding) we adopted. In the current study, the participants could use the mean 'emotional' information from multiple faces with lower possibility of attending each individual, task-irrelevant faces. In this sense, we infer that ensemble coding-like process might take place. With that said, we cannot rule out completely there might be some potentially unsolved/un-identified on-going process.

The present study has some other limitations. First, the experiments did ask people whether the visual target is approaching or going away; this protocol may introduce a bias, because people are aware that the figure is bi-stable (Heenan and Troje 2014). However, with the procedure of two alternative forced choices, observers still exploited the task-irrelevant but differential emotional cues to modulate the FTV bias, so that the baseline/general response bias is controlled. On the other hand, in a given trail, we omitted the first 10 s and used the 60 s recording instead, to largely prevent the initial bias to establish the perception of domination direction. Second, we dichotomized the group based on high and low anxiety. Initially, in designing this study, we tried to recruit as many participants as possible so that the possibility of discriminating individuals with lower vs. higher anxiety is acceptable. However, with LSAS test, we found that it is not easy to obtain the participants with extreme scores corresponding to typical lower vs. higher anxiety traits. In current approach, we used the split-half method (by median LSAS score) instead to separate the group into two-halves. Lastly, In the future, relevant neuroimaging studies may be useful to explore the neural mechanism underlying the emotional modulation upon the FTV (Hindi Attar et al. 2010).

In sum, this study provides the first empirical evidence that task-irrelevant but subjectively emotional cues, conveyed overtly through a single face at a single emotional valence or a group of faces with different emotional valences, could modulate perceptions of the dominant direction of ambiguous motion (i.e., 'facing-the-viewer bias'). This modulation effect was present in life-relevant situations (PLWs), but was not seen in less life-related stimuli (random dots motion). This shows that the semantic correlation between the modulator (facial expressions) and target motion (PLWs) is important in allowing the emotional information to modulate the bi-stable perception. This modulation effect was greater in the group with higher social anxiety traits than in the group with lower-anxiety traits, and was greater with multiple PLWs than with a single PLW. Further neuroimaging studies will reveal the potential neural mechanism in resolving the differential emotional modulation effect using more or less life-relevant motion stimuli.

References

Aaen-Stockdale C, Thompson B, Hess RF, Troje NF (2008) Biological motion perception is cue-invariant. J Vis 8(8):6

Ahlström V, Blake R, Ahlström U (1997) Perception of biological motion. Perception 26(12):1539–1548

Alvarez GA (2011) Representing multiple objects as an ensemble enhances visual cognition. Trends Cogn Sci 15(3):122–131

Ariely D (2001) Seeing sets: representation by statistical properties. Psychol Sci 12(2):157–162

Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, van Ijzendoorn MH (2007) Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. Psychol Bull 133(1):1–24. https://doi.org/10.1037/0033-2909.133.1.1

Battelli L, Cavanagh P, Thornton IM (2003) Perception of biological motion in parietal patients. Neuropsychologia 41(13):1808–1816

Beintema JA, Lappe M (2002) Perception of biological motion without local image motion. Proc Nat Acad Sci 99(8):5661–5663

Bertenthal BI, Pinto J (1994) Global processing of biological motions. Psychol Sci 5:221–225

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10(4):433–436

Bouma H (1970) Interaction effects in parafoveal letter recognition. Nature 226(5241):177–178

Brooks A, Schouten B, Troje NF, Verfaillie K, Blanke O, van der Zwan R (2008) Correlated changes in perceptions of the gender and orientation of ambiguous biological motion figures. Curr Biol 18(17):R728–R729. https://doi.org/10.1016/j.cub.2008.06.054

Cavanagh P, Alvarez GA (2005) Tracking multiple targets with multifocal attention. Trends Cogn Sci 9(7):349–354. https://doi.org/10.1016/j.tics.2005.05.009

Cavanagh P, Labianca AT, Thornton IM (2001) Attention-based visual routines: sprites. Cognition 80(1–2):47–60

Chong SC, Treisman A (2003) Representation of statistical properties. Vision Res 43(4):393–404

Clarke TJ, Bradshaw MF, Field DT, Hampson SE, Rose D (2005) The perception of emotion from body movement in point-light displays of interpersonal dialogue. Perception 34(10):1171–1180. https://doi.org/10.1068/p5203

Cutting JE, Kozlowski LT (1977) Recognition of friends by their walk: gait perception without familiarity cues. Bull Psychon Soc 9:353–356

Daniel PM, Whitteridge D (1961) The representation of the visual field on the cerebral cortex in monkeys. J Physiol 159(2):203–221

de Lussanet MH, Lappe M (2012). Depth perception from point-light biological motion displays. J Vis 12(11). doi: 2.11.14[pii]10.1167/12.11.14

de Gardelle V, Summerfield C (2011) Robust averaging during perceptual judgment. Proc Natl Acad Sci USA 108(32):13341–13346. https://doi.org/10.1073/pnas.1104517108

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1(1):1–47

Fox E, Russo R, Dutton K (2002) Attentional bias for threat: evidence for delayed disengagement from emotional faces. Cogn Emot 16(3):355–379. https://doi.org/10.1080/02699930143000527

Giese MA, Poggio T (2003) Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4(3):179–192

Gilboa-Schechtman E, Foa EB, Amir N (1999) Attentional biases for facial expressions in social phobia: the face-in-the-crowd paradigm. Cogn Emot 13(3):305–318

Gong X, Huang YX, Wang Y, Luo YJ (2011) The revision of chinese facial a ective system. Chin Mental Health J 25(1):40–46

Gray KL, Adams WJ, Garner M (2009) The influence of anxiety on the initial selection of emotional faces presented in binocular rivalry. Cognition 113(1):105–110. https://doi.org/10.1016/j.cognition.2009.06.009

Haberman J, Whitney D (2007) Rapid extraction of mean emotion and gender from sets of faces. Curr Biol 17(17):R751–753

Haberman J, Harp T, Whitney D (2009) Averaging facial expression over time. J Vis 9(11):1 1–13

Haselton MG, Bryant GA, Wilke A, Frederick DA, Galperin A, Frankenhuis WE, Moore T (2009) Adaptive rationality: an evolutionary perspective on cognitive bias. Soc Cogn 27(5):733–763

He Y, Zhang M (2004) Psychometric investigation of liebowitz social anxiety scale. J Diagn Concepts Pract 3:89–93

Heenan A, Troje NF (2014) Both physical exercise and progressive muscle relaxation reduce the facing-the-viewer bias in biological motion perception. PLoS One 9(7):e99902. https://doi.org/10.1371/journal.pone.0099902

Heenan A, Troje NF (2015) The relationship between social anxiety and the perception of depth-ambiguous biological motion stimuli is mediated by inhibitory ability. Acta Psychol (Amst) 157:93–100. https://doi.org/10.1016/j.actpsy.2015.02.012

Hindi Attar C, Muller MM, Andersen SK, Buchel C, Rose M (2010) Emotional processing in a salient motion context: integration of motion and emotion in both V5/hMT + and the amygdala. J Neurosci 30(15):5204–5210. doi:30/15/5204 [pii]10.1523/JNEUROSCI.5029–09.2010

Holbrook C, Galperin A, Fessler DM, Johnson KL, Bryant GA, Haselton MG (2014) If looks could kill: anger attributions are intensified by ordances for doing harm. Emotion 14(3):455–461. https://doi.org/10.1037/a0035826

Hunt S, Low J, Burns KC (2008) Adaptive numerical competency in a food-hoarding songbird. Proc Biol Sci 275(1649):2373–2379. https://doi.org/10.1098/rspb.2008.0702

Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the intentions of others with one's own mirror neuron system. PLoS Biol 3(3):e79. https://doi.org/10.1371/journal.pbio.0030079

Ikeda H, Blake R, Watanabe K (2005) Eccentric perception of biological motion is unscalably poor. Vision Res 45(15):1935–1943

Insch PM, Bull R, Phillips LH, Allen R, Slessor G (2012) Adult aging, processing style, and the perception of biological motion. Exp Aging Res 38(2):169–185. https://doi.org/10.1080/0361073X.2012.660030

Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:195–204

Levi DM, Klein SA, Aitsebaomo AP (1985) Vernier acuity, crowding and cortical magnification. Vision Res 25(7):963–977

Levi DM, McGraw PV, Klein SA (2000) Vernier and contrast discrimination in central and peripheral vision. Vision Res 40(8):973–988

MacLeod C, Mathews A, Tata P (1986) Attentional bias in emotional disorders. J Abnorm Psychol 95(1):15–20

Manera V, Schouten B, Becchio C, Bara BG, Verfaillie K (2010) Inferring intentions from biological motion: a stimulus set of point-light communicative interactions. Behav Res Methods 42(1):168–178. https://doi.org/10.3758/BRM.42.1.168

Mather G, Radford K, West S (1999). Low-level processing of biological motion. Proc Royal Soc London Ser B 249:149–155

Mathews A, MacLeod C (1985) Selective processing of threat cues in anxiety states. Behav Res Ther 23(5):563–569

McNair NA, Goodbourn PT, Shone LT, Harris IM (2017) Summary statistics in the attentional blink. Atten Percept Psychophys 79(1):100–116. https://doi.org/10.3758/s13414-016-1216-2

Miller JE, Carlson LA, McAuley JD (2013) When what you hear influences when you see: listening to an auditory rhythm influences the temporal allocation of visual attention. Psychol Sci 24(1):11–18

Neri P, Morrone MC, Burr DC (1998) Seeing biological motion. Nature 395(6705):894–896

Pan J, Zhang J, Ma P, Liang H, Wang H, Tao J, Wen S, Zhang J (2006) The utility of Liebowitz Social Anxiety Scale in the patients with social anxiety disorder in Chinese. Chin J Nerv Ment Dis 32(3):206–210

Pavlova MA (2011) Biological motion processing as a hallmark of social cognition. Cereb Cortex 22(5):981–995 doi:bhr156[pii]10.1093/cercor/bhr156

Pavlova MA, Sokolov AN, Bidet-Ildei C (2014). Sex di erences in the neuromagnetic cortical response to biological motion. Cereb Cortex. doi:bhu175[pii]10.1093/cercor/bhu175

Pell DG, Cavanagh P, Desimone R, Tjan B, Treisman A (2007) Crowding: including illusory conjunctions, surround suppression, and attention. J Vis 7(2):1

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10(4):437–442

Piazza EA, Sweeny TD, Wessel D, Silver MA, Whitney D (2013) Human use summary statistics to perceive auditory sequences. Psychol Sci 24(8):1389–1397

Pollick FE, Kay JW, Heim K, Stringer R (2005) Gender recognition from point-light walkers. J Exp Psychol Hum Percept Perform 31(6):1247–1265. doi:2005-15838-006[pii]10.1037/0096-1523.31.6.1247

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230

Rizzolatti G, Fabbri-Destro M (2010) Mirror neurons: from discovery to autism. Exp Brain Res 200(3–4):223–237. https://doi.org/10.1007/s00221-009-2002-3

Rytwinski NK, Fresco DM, Heimberg RG, Coles ME, Liebowitz MR, Cissell S, Hofmann SG (2009) Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. Depress Anxiety 26(1):34–38. https://doi.org/10.1002/da.20503

Safren SA, Heimberg RG, Horner KJ, Juster HR, Schneier FR, Liebowitz MR (1999) Factor structure of social fears: the Liebowitz Social Anxiety Scale. J Anxiety Disord 13(3):253–270

Schouten B, Troje NF, Brooks A, van der Zwan R, Verfaillie K (2010) The facing bias in biological motion perception: e ects of stimulus gender and observer sex. Atten Percept Psychophys 72(5):1256–1260. https://doi.org/10.3758/APP.72.5.1256

Schouten B, Troje NF, Verfaillie K (2011) The facing bias in biological motion perception: structure, kinematics, and body parts. Atten Percept Psychophys 73(1):130–143. https://doi.org/10.3758/s13414-010-0018-1

Shim W, Alvarez G, Jiang Y (2008) Spatial separation between targets constraints maintenance of attention on multiple objects. Psychon Bull Rev 15(2):390–397

Singer N, Eapen M, Grillon C, Ungerleider LG, Hendler T (2012) Through the eyes of anxiety: Dissecting threat bias via emotional-binocular rivalry. Emotion 12(5):960–969. https://doi.org/10.1037/a0027070

Thornton IM, Rensink RA, Shi rar M (2002) Active versus passive processing of biological motion. Perception 31(7):837–853

Troje NF, Westho C (2006) The inversion e ect in biological motion perception: evidence for a "life detector"? Curr Biol 16(8):821–824. doi:S0960-9822(06)01281-4[pii]10.1016/j.cub.2006.03.022

Van de Cruys S, Schouten B, Wagemans J (2013) An anxiety-induced bias in the perception of a bistable point-light walker. Acta Psychol (Amst) 144(3):548–553. doi:S0001-6918(13)00215-1[pii]10.1016/j.actpsy.2013.09.010

Vanrie J, Verfaillie K (2004) Perception of biological motion: a stimulus set of human point-light actions. Behav Res Methods Instrum Comput 36(4):625–629

Walker D, Vul E (2014) Hierarchical encoding makes individuals in a group seem more attractive. Psychol Sci 25(1):230–235

Weech S, McAdam M, Troje NF (2014) What causes the facing-the-viewer bias in biological motion? JVis 14:1–15

Westheimer G (1982) The spatial grain of the perifoveal visual field. Vision Res 22(1):157–162

Whitney D, Levi DM (2011) Visual crowding: a fundamental limit on conscious perception and object recognition. Trends Cogn Sci 15(4):160–168. https://doi.org/10.1016/j.tics.2011.02.005

Yiltiz H, Chen L (2015) Tactile input and empathy modulate the perception of ambiguous biological motion. Front Psychol 6:161. https://doi.org/10.3389/fpsyg.2015.00161

Ying HJ, Xu H (2017). Rapid sequential visual presentation and ensemble coding in facial expression adaptation. J Vi 17(15). https://doi.org/10.1167/1117.1161.1115