Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Categorical processing of emotion in ensemble perception and the foveal input bias

Yu R. Dandan
Peking University
Jing Cui
Southeast University
Fang Fang
Peking University
Bilge Sayim
bilge.sayim@cnrs.fr

École Normale Supérieure (ENS), PSL University, Centre National de la Recherche Scientifique (CNRS)

Research Article

Keywords: Ensemble perception, Foveal input bias, Categorical perception

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11467246/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Categorical processing of emotion in ensemble perception and the foveal input bias

Yu R. Dandan[1], Jing Cui[2], Fang Fang[1,3,4,5*], Bilge Sayim[6*]

[1] School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Brain-Computer Interface and Mental Health Modulation, Peking University, Beijing 100871, China

[2] School of Foreign Languages, Southeast University, Nanjing, China

[3] International Data Group/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

[4] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

[5] Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China

[6] Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, École Normale Supérieure (ENS), PSL University, Centre National de la Recherche Scientifique (CNRS), Paris, France

Abstract

When perceiving the emotionality of groups of faces, fixated faces weigh more than faces not fixated, potentially influencing social judgments and subsequent behavior. This effect—the foveal input bias (FI-bias)—reveals location-dependent weighting in ensemble perception that may originate from differences in early sensory encoding, with foveal input providing more precise sensory information than peripheral input. Although facial expressions vary continuously, observers perceive them as discrete emotional categories. Whether such categorical representations modulate location-dependent weighting underlying the FI-bias remains unclear. In two experiments (N=64), participants reported the average emotion of face groups with the fixated ('foveal') face either belonging to the same (within-category) or a different category (across-category) than the other faces ('flankers'). The FI-bias was robust in both within- and across-category conditions. The relative contributions of the foveal face and flankers to ensemble perception varied as a function of the difference in (perceived) emotional intensity in the across-category condition, but remained comparatively stable in the within-category condition. Crossing categorical boundaries reduced the FI-bias and increased the tendency to report the ensemble as "neutral", the categorical boundary between the two emotional categories. These findings suggest that categorical information modulates the spatial weighting profile in ensemble perception without eliminating the FI-bias.

Keywords

Ensemble perception, Foveal input bias, Categorical perception

Introduction

The processing of facial expressions has an important role in our social life. Facial expressions are a universal signaling system that reflects a person's emotional state in real time (Ekman & Friesen, 1971; Russell et al., 1994; Elfenbein & Ambady, 2002). Within this context, group perception is particularly important in social interactions. The visual system's ability to extract the overall emotional information from groups of faces is called ensemble emotion perception (Haberman & Whitney, 2007, 2009; Whitney & Yamanashi Leib, 2018). For example, when entering a classroom, we can quickly estimate the overall emotional state of a group of students. Ensemble emotion perception is remarkably efficient. Observers could extract the average emotion of a group of faces successfully with short stimulus durations (e.g., as short as 50 ms; Li et al., 2016), large set sizes (Wolfe et al., 2015), and dynamic (Elias et al., 2017), ambiguous (e.g., Mooney face, Han et al., 2021), and sequentially presented facial expressions (Haberman et al., 2009; Hubert-Wallander & Boynton, 2015; Ji et al., 2018; Goldenberg et al., 2022).

It is widely accepted that ensemble perception requires the integration of multiple items (for reviews, see Alvarez, 2011; Whitney & Yamanashi Leib, 2018; Corbett et al., 2023); however, how individual items contribute to the perceived ensemble remains a topic of ongoing debate. A growing body of evidence indicates that ensemble estimates arise from a non-uniform integration in which some items are weighted more heavily than others (De Fockert & Marchant, 2008; Jung, Bülthoff, & Armann, 2017; Pascucci, Ruethemann, & Plomp, 2021; Tiurina et al., 2024; Ni & Stocker, 2023, 2026; Hochmitz et al., 2026). A key example is location-dependent compulsory weighting. In particular, items at (or close to) fixation weigh more than others in the group, revealing a foveal input bias ('FI-bias'; e.g., Pascucci et al., 2021; Dandan et al., 2023; Tiurina et al., 2024). This spatial anisotropy is partly driven by cortical magnification, whereby foveal input has a disproportionately greater cortical representation, resulting in higher precision and lower sensory uncertainty than for peripheral input (for cortical

magnification, see Virsu & Rovamo, 1979; Duncan & Boynton, 2003; Harvey & Dumoulin, 2011; Himmelberg et al., 2022).

Notably, although FI-bias may arise from differences in sensory reliability between foveal and peripheral inputs, it remains unclear to what extent this location-dependent weighting can be modulated by higher-level representations. Facial expressions provide a particularly useful test case because they convey socially meaningful information that can be represented categorically, while varying continuously in their basic visual features. For example, two facial expressions that differ by the same physical magnitude can be perceived as either belonging to the same or different emotional categories (Etcoff & Magee, 1992; Schyns & Oliva, 1999; Kiffel et al., 2005; Fugate, 2013). We therefore asked whether emotional categories influence the location-dependent weighting underlying the FI-bias in ensemble perception. The FI-bias can be explained as a consequence of inhomogeneous encoding of foveal and peripheral items (Ni & Stocker, 2023, 2026). In contrast, the retrospective Bayesian decoding theory (Ding et al., 2017) suggests that the brain prioritizes the decoding of stable, high-level categorical information of stimuli, which then retrospectively constrain low-level sensory representations. Based on this assumption, the FI-bias in ensemble perception may not be fixed during the early encoding stage. Instead, it may reflect a flexible weighting mechanism that can be modulated by later-stage categorical and decisional processes.

The primary goal of this study is to investigate whether the strength of the FI-bias is modulated by higher-level representations, specifically whether the fixated (‘foveal’) face and non-fixated faces in the group (‘flankers’) belong to the same or different emotional categories (“categorical consistency”). Stimuli consisted of a 3×3 matrix of faces with the foveal face (at the center) and the flankers either from the same emotional category (‘within-category’) or from different emotional categories (‘across-category’). The task was to evaluate the average emotional intensity of the face group using a rating scale. A core behavioral signature of categorical perception is the

across-category advantage, whereby two stimuli are discriminated more accurately when they fall on opposite sides of a category boundary than when they differ by the same physical amount but fall within the same category (for reviews, see Goldstone & Hendrickson, 2010; Fugate, 2013; Timeo, Farroni, & Maass, 2016). If categorical structure modulates how individual items are integrated during ensemble perception, two competing predictions arise for the integration of a foveal face with surrounding flankers. The foveal face may be more perceptually distinct from the flankers in the across-category condition, consistent with the across-category advantage (for reviews, see Goldstone & Hendrickson, 2010; Fugate, 2013; Timeo et al., 2016). If this increased perceptual distinctiveness leads to greater weighting of the more reliable sensory source, the contribution of the foveal face should be greater in the across-category condition than in the within-category condition. Alternatively, when the foveal face and flankers fall into different emotional categories, the competing categorical information may reduce the dominance of either category during ensemble integration. This competition may shift the integrated percept toward an intermediate response near the category boundary, rather than toward either emotional category. Under this account, the ensemble estimation should be less strongly driven by the foveal face, resulting in a smaller contribution of the foveal face in the across-category condition than in the within-category condition. By contrast, if ensemble perception is governed solely by low-level, reliability-weighted sensory pooling and is unaffected by categorical consistency, foveal weighting should be comparable in within- and across-category conditions. Distinguishing among these predictions can clarify whether categorical representations modulate the weighting and integration process underlying ensemble perception.

Experiment 1: The FI-bias when perceiving within- and across-category face ensembles

Method

Participants

34 undergraduate students (18-30 years, 19 females, 15 males) from Peking University participated in the current study. All participants provided written consent before their participation. All participants were right-handed and had normal or corrected-to-normal visual acuity. The study was approved by the institutional review board at Peking University. A power analysis of linear mixed models via simulation (SIMR package in R; Green & MacLeod, 2016), based on data from a recent study on the foveal input bias (Dandan et al., 2023), indicated that 17 participants would suffice to achieve 99.99% power in our analysis. In the power analysis, we used category information (within-category vs. across-category) as the predictor, participants' ensemble performance (as indicated by discriminability (d′)) as the dependent variable, and participant ID as the random effect.

Apparatus

The experiment was conducted on a standard 19" LCD display (E196FP, DELL) at a refresh rate of 60Hz and a resolution of 1280 × 1024. Participants were seated at a distance of 57 cm from the screen. E-prime 3.0 was used for stimulus presentation and data collection. A chin rest was used during the experiment.

Stimuli and Procedure

We used three facial expressions (happy, disgusted, neutral) of one female from the NimStim face set (Tottenham et al., 2009) to morph different intensity levels of facial expressions. Disgusted and happy expressions were used to generate positive and negative emotions. Previous studies indicate that happiness and sadness elicit varying intensities of arousal, while arousal differences between happiness and disgust seem less evident (Sutton et al., 2019; To et al., 2019). Hence, to reduce the

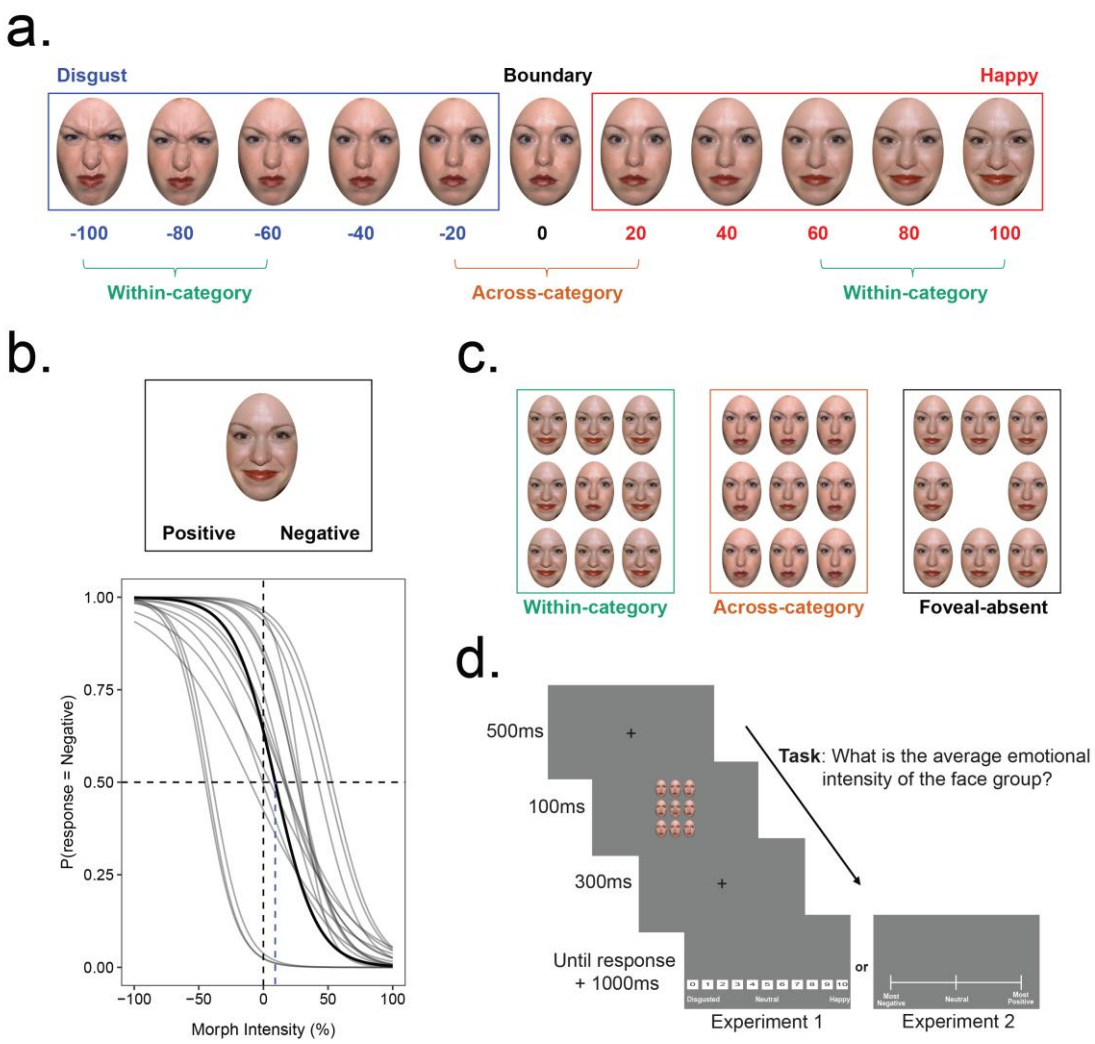
probability of different intensities of arousals due to different positive and negative emotions, we used disgust (and not sadness). Specifically, we created a series of 11 facial expression levels by morphing happy and disgusted faces with a neutral expression, ranging from 100% disgusted to 100% happy in 20% increments, with neutral as the midpoint (Figure 1a). All the external face contours and all hair were removed (using GIMP v2.10). The background of each image was replaced with a transparent background. Each face subtended a visual angle of 1.49° × 2.21° and was separated by 0.30° horizontally and 0.15° vertically from neighboring faces (the edge-to-edge distance). The whole face group (group size: 1, 8, or 9) subtended a visual angle of 5.07° × 6.93°. To examine the categorical boundary, we reanalyzed the data from a previous study (Dandan et al., 2023). In this identification task (N = 17), a single face was presented at the center of the screen for 100 ms. Participants were asked to identify the facial expression as either “positive” or “negative”. The categorical boundary — defined as the morph intensity associated with chance-level performance — occurred at 9% happiness, as indicated by the blue dashed line in Figure 1b.

We used a within-subjects design in which each participant viewed and rated the facial expressions of a face group containing 8 or 9 faces, or a single face. The task was to report the perceived average emotion of the face group (or the emotion of the single face) by using the mouse to click the corresponding labels: An 11-point rating scale was used in the experiment (from 0 to 10; ‘0’ representing ‘disgusted’, ‘5’ representing ‘neutral’, ‘10’ representing ‘happy’). In the foveal-present condition, one of the 11 different facial expressions was used for the foveal face and one for the flankers (121 trials: 11 foveal faces × 11 flankers). The 8 flankers in each face group were identical (with the same emotional expression). The foveal face and the flankers belonged to either the same or different emotional categories (see Figure 1c for an illustration). In the within-category condition, both the foveal face and the flankers were either both happy or both disgusted. In the across-category condition, the foveal face

was happy when the flankers were disgusted and vice versa. In the foveal-absent condition, only flankers were presented, repeating the 11 facial expressions 11 times (121 trials: 11 flankers × 11 repetitions) to match the trial number of the foveal-present condition. The comparison between the foveal-present and foveal-absent conditions was used as an indicator for the bias induced by the foveal face. In the single-face condition, a single face was presented at the center of the screen (121 trials: 11 faces × 11 repetitions). The single-face condition was used to obtain each participant's perceived emotional intensity of each of the 11 expressions. In total, there were 3 blocks with 121 trials each (foveal-present, foveal-absent, and single-face condition). The different blocks were presented in random order for each participant. Participants practiced at the beginning of the experiment without receiving feedback. In 12 practice trials, a face group containing 8 or 9 faces in separate blocks was presented at the center of the screen and participants were required to report the average emotion of the face group.

On each trial, a black fixation cross was presented for 500 ms, followed by a face group containing 1, 8 or 9 faces. Stimuli were presented for 100 ms with the foveal face in the center of the screen. After stimulus offset, a blank screen was presented for 300 ms, followed by the response screen. Participants were asked to report either the face group's average emotion or the single face's emotion on the rating scale (from 0=disgusted to 10=happy in Experiment 1; from -100=most negative to 100=most positive in Experiment 2). After participants' responses, an inter-trial interval of 1000 ms was inserted before the next trial (Figure 1d).

Figure 1



Note. **(a)** Facial stimuli used in the experiment. **(b)** Categorical boundary. The gray lines represent logistic curves fitted to individual participant's data and the black line represents the model fit to the group average. **(c)** Stimulus examples: Within-category (20% happy foveal face and 100% happy flankers), Across-category (40% happy foveal face and 40% disgusted flankers), Foveal-absent (60% happy, flankers only). In the example, the emotional contrast is 80% for both the within-category and across-category conditions. **(d)** Sequence of one trial: A face group (here 9 faces) was presented in the center of the screen for 100 ms. In Experiment 1, participants were asked to report either the average emotion of the face group (or the single face's emotion) by clicking the corresponding number keys (0-10). In Experiment 2, they were

instructed to click any point on a rating scale ranging from most negative to most positive with neutral as the midpoint (-100 to +100).

Transparency and Openness

All data from the current study are available online via the Open Science Framework (https://osf.io/ne37r). This study's design and its analysis were not pre-registered.

Analysis

We calculated the perceived emotional intensity for each single face to determine the perceptual scaling between physical morph level and perceived intensity (Figure 2a). The analysis served as a cross-validation of the emotional intensity manipulation. Responses, originally collected on a 0-10 scale, were linearly rescaled to range from -100 (negative) to 100 (positive). To assess the unique contributions of the foveal face and the flankers to the perceived ensemble, we performed a linear regression analysis with the perceived emotional intensity of the foveal face and flankers as predictors and participants' responses as the dependent variable (Figure 2b). The raw weights were normalized using RMS-based normalization (see Dandan et al., 2026). The magnitude of the FI-bias in the within-category and across-category conditions was assessed by measuring the response difference when the foveal face was present compared to when it was absent (Figure 2c). Specifically, we calculated the absolute deviation in each condition,

$$Absolute\ Deviation = \frac{\sum_{i=1}^{n} |\ Perceived\ average\ with\ foveal\ face_i - Without\ foveal\ face_i|}{\mathrm{n}}\ , \quad (1)$$

where parameter i is the trial index and parameter n is the total number of trials in the corresponding condition. Larger absolute deviation indicates stronger FI-bias. Trials with the absolute deviation either below or above 3SD from the mean were categorized as outliers and excluded from the formal analysis (0.8% of all trials). To further evaluate the directionality of the observed FI-bias, we further calculated the relative deviation in each condition,

$$Relative\ Deviation = \frac{\sum_{i=1}^{n} Perceived\ average\ with\ foveal\ face_i - Without\ foveal\ face_i}{\mathrm{n}}, \quad (2)$$

where i is the trial index and n is the total number of trials in each condition. Positive (negative) values represent participants perceived the face group as more positive (negative) compared to when there was no foveal face. The average responses across participants and trials for each combination of the foveal face and flankers are shown in Figure 2e.

We performed linear mixed model analyses for the absolute deviation and relative deviation indexes (lmer function from lme4 package in R; Bates et al., 2015). Models were fitted with the restricted maximum likelihood (REML) to produce unbiased parameter estimates. We employed a stepwise procedure to simplify the model, selecting the optimal model based on the Likelihood Ratio Test (anova function in LMMs). The selection criteria included decreases in the Akaike Information Criterion (AIC) and a stopping threshold of $p < 0.05$. *T*-tests on individual parameter estimates were used for posterior comparisons (emmeans package in R; *p*-value adjustment: Bonferroni method). The degrees of freedom of the denominator, required for *F* and *t*-tests, were approximated with Satterthwaite's method (lmerTest package in R; Kuznetsova et al., 2017). Partial R^2 for each predictor in the model was calculated to estimate effect sizes (r2glmm package; Edwards et al., 2008; Jaeger et al., 2016).

Results

Perceptual intensity vs. morph intensity of the single face

A linear mixed model was performed to evaluate whether the perceived intensity of the single face was positively associated with the morph intensity in two emotional valence conditions (positive vs. negative). The final model included only the main effects of morph intensity (b = 0.63 [0.59, 0.68], SE = 0.02, $t(304)$ = 29.58, $p < 0.001$, R^2 = 0.74) and valence (b = -5.48 [-7.85, -3.10], SE = 1.21, $t(304)$ = -4.52, $p < 0.001$, R^2 = 0.06). The morph intensity significantly predicted increases in participants' perceived intensity of the single face. Based on the model fit, the perceived emotional

intensity of positive faces was 5.48% lower than that of negative ones, suggesting that observers perceived negative faces as more emotionally intense than positive faces. Additionally, as shown in Figure 2a, perceived emotional intensity was generally lower than the morph intensity, especially at higher morph levels, revealing a scale compression effect.

Contributions of the foveal face and the flankers to the perceived ensemble

We used linear regression, estimated using the Maximum Likelihood Estimation (MLE) approach, to obtain the best-fitting sets of weights for each participant under each category condition (Figure 2b). Note that we used the perceptual intensity from the single face condition for each participant as the emotional intensity of the foveal face and the flankers (not the morph intensities). The regression analysis was based on participants' perceived intensities to estimate the weighting of internally encoded facial representations while controlling for perceptual scaling. We then conducted a linear mixed model analysis with RMS-normalized weight as the dependent variable and category (within vs. across) and location (foveal face vs. flankers) as the two predictors. Results showed a significant main effect of location (b = -0.30 [-0.54, -0.06], SE = 0.12, $t(132)$ = -2.39, p = 0.02, R^2 = 0.10). The main effect of category (b = 0.08 [-0.17, 0.32], SE = 0.12, $t(132)$ = 0.61, p = 0.54, R^2 = 0.001) and the interaction effect (b = -0.08 [-0.42, 0.27], SE = 0.18, $t(132)$ = -0.44, p = 0.66, R^2 = 0.001) were not significant. In both category conditions, the foveal face was weighted more strongly than the flankers, revealing a pronounced FI-bias.

The categorical effect on the FI-bias

We then examined the magnitude of the FI-bias in both within-category and across-category conditions (Figure 2c). A linear mixed model with participant ID as the random factor, category as the predictor, and absolute deviation as the dependent variable revealed that the error was greater in the across-category condition compared to the within-category condition (b = 23.37 [18.49, 28.24], SE = 2.46, $t(33)$ = 9.52, p < 0.001, R^2 = 0.67). Additionally, the absolute deviation in both category conditions was

significantly different from zero (one-sample *t-test*: within-category: $t(33) = 23.18$, $p < 0.001$, $d = 3.97$; across-category: $t(33) = 48.36$, $p < 0.001$, $d = 2.75$). As a control analysis, we quantified ensemble estimation error as the absolute deviation between participants' ensemble estimates and the actual average of the group, and obtained the same pattern of results (Figures S1a and S1b in the Appendix).

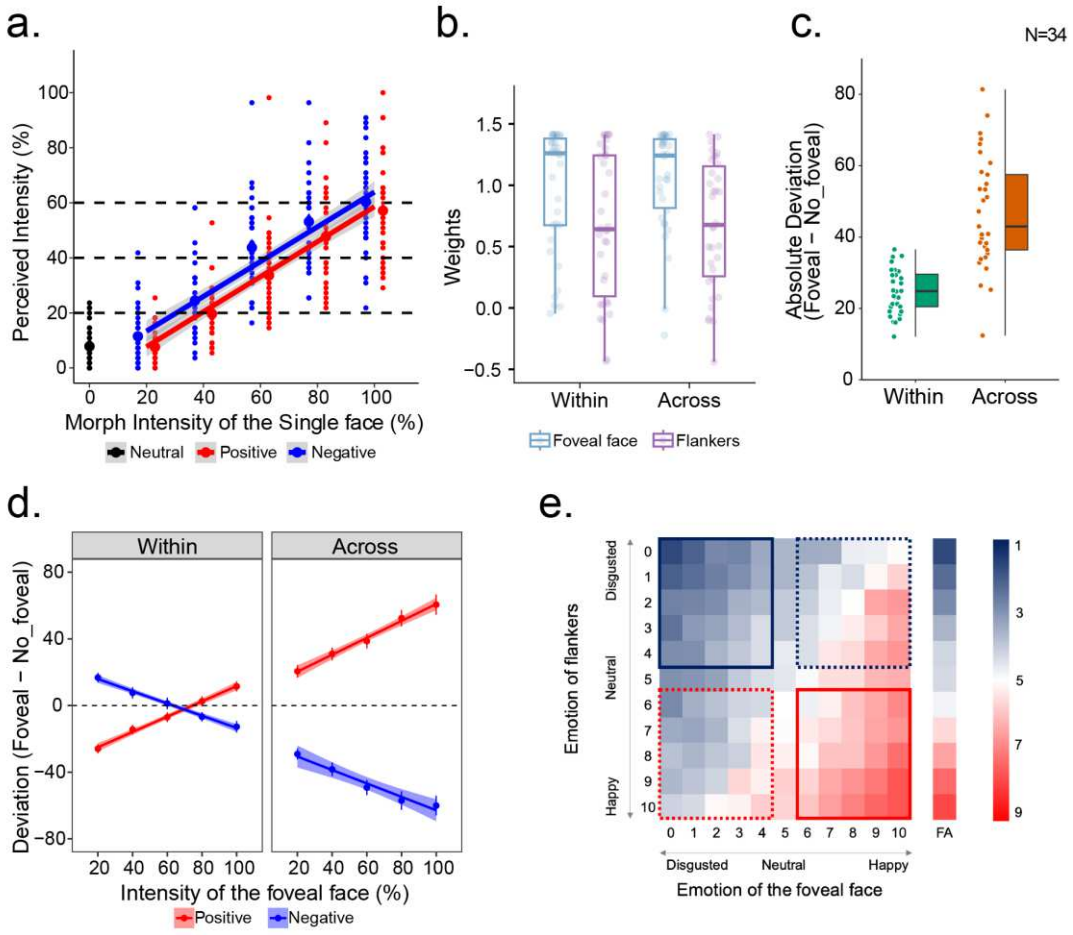
To assess whether the magnitude of the FI-bias was modulated by the emotional valence and the intensity of the foveal face, we compared the relative deviation as a function of the emotional valence and intensity of the foveal face in both within-category and across-category conditions (Figure 2d). The full model included category (within, across), valence (positive, negative), intensity (20%, 40%, 60%, 80%, 100%), and their interactions. Participant ID was added as the random factor. The final model with the lowest AIC included the main effects of the three factors (category: $b = -47.97$ [-52.42, -43.52], $SE = 2.27$, $t(641) = -21.10$, $p < 0.001$, $R^2 = 0.00$; valence: $b = -59.98$ [-67.99, -51.96], $SE = 4.10$, $t(641) = -14.63$, $p < 0.001$, $R^2 = 0.02$; intensity: $b = -0.39$ [-0.46, -0.31], $SE = 0.04$, $t(641) = -9.58$, $p < 0.001$, $R^2 = 0.01$), along with the interactions between category and valence ($b = 95.28$ [88.99, 101.57], $SE = 3.22$, $t(641) = 29.63$, $p < 0.001$, $R^2 = 0.58$), and between valence and intensity ($b = 0.87$ [0.76, 0.98], $SE = 0.06$, $t(641) = 15.27$, $p < 0.001$, $R^2 = 0.27$). The category $\times$ valence interaction showed that when the foveal face was negative (positive), participants perceived the entire group as more negative (positive) in both category conditions, relative to the foveal-absent condition. The valence $\times$ intensity interaction revealed that participants perceived the face group as more negative (positive) with increased emotional intensity of negative (positive) foveal faces.

Figure 2e plots the average response for each combination of foveal face and flankers, showing how responses varied when the emotion of the foveal face and flankers differed or belonged to the same category. The emotion of the foveal face is shown on the x-axis; the emotion of the flankers on the y-axis, both varied from 100%

disgusted to 100% happy. A value ‘0’ represents reporting the face group as 100% disgusted, ‘5’ represents reporting the face group as neutral, and ‘10’ represents reporting the face group as 100% happy. The results showed that the ensemble performance was strongly biased by the emotion of the foveal face. As illustrated in Figure 2e, in the across-category condition, the flankers had an emotional valence opposite to the foveal face (as indicated by dashed rectangles). Hence, participants would have achieved greater accuracy by reporting the emotional intensity of the flankers. Instead, they frequently reported the emotionality of the foveal face as the group average. In the within-category condition where the foveal face had the same emotion as the flankers but different emotional intensity, participants were biased by the foveal face’s emotion—they reported the face group as more intense with the increased emotional intensity of the foveal face when the corresponding flankers were the same. Together, the findings revealed a pronounced FI-bias in both within- and across-category face groups.

Figure 2

Results of Experiment 1



Note. **(a)** The perceived emotional intensity plotted against the morph intensity and emotional valence of the single face. The dashed lines represent the corresponding morph intensity when the perceived intensity is 20%, 40%, and 60% positive/negative. **(b)** RMS-normalized weights corresponding to each predictor (foveal face vs. flankers) shown for within-category and across-category conditions. Each data point represents the distributed weight of one observer. **(c)** Absolute deviation separated for within-category and across-category conditions. **(d)** Relative deviation in the positive and negative foveal face conditions across the 5 emotional intensities of the foveal face (from 20% to 100%). Dots represent averaged relative deviation across participants at different valence and intensity levels. Lines represent fitted linear regression lines. Error bars represent standard errors of the mean; the colored ribbons represent the

95% confidence interval of the model fit. **(e)** Each cell in the matrix represents the average rating for all combinations of foveal faces and flankers. Blue (red) rectangles surround the cells where “disgusted” (“happy”) was the correct response. Dashed rectangles surround cells of the across-category condition where the foveal face had a different emotion than the flankers; solid rectangles surround cells of the within-category condition where the foveal face had the same emotion as the flankers but with different emotional intensity. “FA” represents the Foveal-absent condition.

Discussion

The findings of Experiment 1 revealed a FI-bias both within and across emotional categories. In the experiment, the eight flankers were the same and the foveal face was the statistical outlier which was expected be ignored or at least down-weighted based on the robust averaging weighting account (see for example, De Gardelle & Summerfield, 2011; Li et al., 2017; Ni & Stocker, 2023, 2026). However, we consistently observed a pronounced FI-bias: when the foveal face was positive (negative), participants perceived the face group as more positive (negative) with increased emotional intensity of the foveal face (Figures 2d, 2e). According to the calculated spatial weighting profile (Figure 2b), the foveal face was weighted more strongly than the flankers in both category conditions, demonstrating a pronounced FI-bias. Together, these results demonstrate a location-based compulsory weighting mechanism in ensemble perception, such that foveal items receive disproportionate weight in ensemble judgments, even when this weighting is detrimental to task performance.

Notably, the spatial weighting profile showed comparable foveal face weight in both within- and across-category conditions (Figure 2b), however, the deviation results revealed more deleterious ensemble performance in the across-category condition compared to the within-category condition. One plausible explanation is that in the current experiment, the emotional contrast between the foveal face and flankers varied from 0% (both the foveal face and flankers had the same emotional intensity) to 200%

(the foveal face was 100% positive and flankers were 100% negative or vice versa), and the contrast in the across-category condition was on average higher than the within-category condition. Therefore, the bigger response errors in the across-category condition may simply be driven by the greater emotional contrast between the foveal face and flankers rather than the perceived categorical difference. In an exploratory analysis, we evaluated the magnitude of the FI-bias as a function of the difference in perceived emotional intensity between the foveal face and the flankers and found a positive correlation (Figure S1C in the Appendix). To control for the influence of the magnitude of emotional contrast, in the second experiment, we investigated the categorical effect on the observed FI-bias under conditions in which the emotional contrast remained constant between the two category conditions.

Experiment 2: Constant contrast: the categorical effect on FI-bias

Method

Participants

We recruited 32 new participants (undergraduate students, 19-28 years, 21 females, 11 males) for the current experiment. Two participants were excluded from the analysis because their ratings of individual faces did not scale linearly with the physical (morph) emotional intensity of the faces.

Stimuli and design

As in Experiment 1, we used perceived intensity rather than morph intensity of the faces in the current experiment. The perceptual intensity of each face was estimated using a linear regression model fitted to the single face condition in Experiment 1 (see Figure 2a). For instance, faces perceived as 20% happy corresponded to 39% happy morph intensity, and faces perceived as 20% disgusted corresponded to 31% disgusted morph intensity (see Figure 2a). There were 8 face images of the same person (the same face identity used in Experiment 1) included in the current

experiment: perceptual 10%, 20%, 30%, 50% happy/disgusted, corresponding to morph levels of 24%/15%, 39%/31%, 55%/47%, 87%/78% happy/disgusted. We used grey-scaled faces to diminish the effect induced by low-level features (e.g., the redness of lips). The experimental design was the same as in Experiment 1 except for the emotional contrast manipulation. Three emotional contrast levels (i.e., the emotional intensity difference between the foveal face and the flankers) were used: 20%, 30%, and 40%. For each contrast level, we created 8 face combinations (2 categories $\times$ 2 emotional valences $\times$ 2 face combinations) based on the foveal face and flankers' emotional intensity (see Table 1). The 8 flankers were identical (same as Experiment 1). In the control condition, only 8 flankers were presented in each trial ('foveal-absent' condition). There were 6 blocks in total (3 conditions (foveal-present, foveal-absent, single-face) $\times$ 2 repetitions), and each block consisted of 96 trials (4 face combinations $\times$ 2 categories $\times$ 3 contrast levels $\times$ 4 repetitions). The 6 blocks were presented in random order. Participants were required to judge the average emotional intensity of the face group or the single face's emotional intensity. They practiced before the experiment to familiarize themselves with the task demands. The experimental procedure was the same as in Experiment 1, with the exception that a continuous rating scale ranging from negative ('-100') to positive ('+100'), with the midpoint ('0') denoting 'neutral' was used (Figure 1b).

Table 1

Face combinations under three contrast levels and two emotional categories

Contrast	Category	Face combinations [Foveal face, Flankers]	
20%	Within	[10, 30]	[30, 10]
		[-10, -30]	[-30, -10]
	Across	[10, -10]	
		[-10, 10]	
30%	Within	[20, 50]	[50, 20]
		[-20, -50]	[-50, -20]
	Across	[10, -20]	[-20, 10]
		[-10, 20]	[20, -10]
40%	Within	[10, 50]	[50, 10]
		[-10, -50]	[-50, -10]
	Across	[10, -30]	[30, -10]
		[-10, 30]	[-30, 10]

*N*ote: Positive values represent a happy (positive) expression, while negative values represent a disgusted (negative) expression; the value itself represents the perceptual emotional intensity.

Data analysis

The main analyses were identical to those used in Experiment 1. Trials with absolute deviations either below or above 3SD from the mean were categorized as outliers and excluded from the formal analysis (0.4% of all trials). Furthermore, we performed Gaussian mixture modeling to characterize the latent response sources contributing to the observed response distribution. Specifically, we modeled participants' responses as a mixture of three components centered on the foveal face, flankers, and the categorical boundary ('perceived neutral'). To capture the best-fitting models, we compared three models: (1) a fixed 1-component model, in which the response distribution was modeled as a single Gaussian centered on the group-average response; (2) a fixed 2-component model, with two Gaussians centered on the foveal face and flankers, respectively; (3) a fixed 3-component model, which

additionally included a Gaussian centered on the perceived neutral boundary. The formulas and interpretations are presented below (Table 2).

Table 2

Gaussian mixture models

Model	Formula	κ	Interpretation
1-component	$p(y_i) = \mathcal{N}(y_i \mid A_i, \sigma_A^2)$	1	Responses are centered on the group average (A_i)
2-component	$p(y_i) = \omega_F \mathcal{N}(y_i \mid F_i, \sigma_F^2) + \omega_L \mathcal{N}(y_i \mid L_i, \sigma_L^2)$	3	Responses reflect the relative influence of foveal (F_i) and flanker (L_i) information
3-component	$p(y_i) = \omega_F \mathcal{N}(y_i \mid F_i, \sigma_F^2) + \omega_L \mathcal{N}(y_i \mid L_i, \sigma_L^2) + \omega_0 \mathcal{N}(y_i \mid 0, \sigma_0^2)$	5	Responses reflect the relative influence of foveal information (F_i), flanker information (L_i), and boundary-centered response tendency (0)

Note: In the 2-component model, the mixture weights were constrained using a logistic transformation, whereas in the 3-component model, they were constrained using a softmax transformation. These transformations ensured that all component weights were positive and summed to 1. k represents the number of free parameters estimated in each model. y_i denotes the observed response on trial i; A_i, F_i, and L_i denote the group average, foveal face value, and flankers value on trial i, respectively. $\mathcal{N}(y_i \mid \mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2. ω_F, $\omega_{L,}$ ω_0 denote the mixture weights of the foveal-centered, flanker-centered, and boundary-centered components, respectively. σ_F^2, σ_L^2, σ_0^2 denote the corresponding component variances, and σ_A^2 denotes the variance of the group average component.

Results

Unique contributions of the foveal face and flankers to the perceived ensemble

We fitted the linear regression model with the emotional intensity of the foveal face and flankers as predictors, and participants' responses as the dependent variable. Weights were estimated separately for each contrast and category condition at the individual level. We further conducted a linear regression analysis on the best-fitting

sets of RMS-based weights, with location (foveal face vs. flankers), category (within vs. across), contrast (20%, 30%, 40%), and their interactions as fixed factors (Figure 3a). The best-fitting model included the main effects of location (b = 0.63 [0.50, 0.75], SE = 0.07, t(355) = 9.64, p < 0.001, R^2 = 0.21), category (b = 1.36 [0.87, 1.84], SE = 0.25, t(355) = 5.48, p < 0.001, R^2 = 0.08), contrast (b = 0.03 [0.02, 0.05], SE = 0.01, t(355) = 6.14, p < 0.001, R^2 = 0.06), as well as a significant interaction effect between the category and contrast (b = -0.03 [-0.05, -0.02], SE = 0.01, t(355) = -3.92, p < 0.001, R^2 = 0.04). Specifically, in the within-category condition, the weights assigned to the foveal face and the flankers were comparable across the three contrast levels (F (2, 353) = 0.18, p = 1.00). In the across-category condition, however, the weights assigned to both the foveal face and the flankers increased as emotional contrast increased (F (2, 353) = 19.74, p < 0.001). Furthermore, we calculated the weight deviation between the across- and within- category conditions for each combination of location and contrast, and then performed the one-sample Wilcoxon test to determine whether these deviations differed significantly from zero (Figure 3b). For the foveal face, the deviation was significantly different from zero only at 20% contrast level (20% contrast: W = 58, p < 0.001, effect size r = 0.66; 30% contrast: W = 198, p = 0.49, r = 0.13; 40% contrast: W = 316, p = 0.09, r = 0.31). For the flankers, however, the deviation was significantly different from zero at all contrast levels (20% contrast: W = 56, p < 0.001, r = 0.66; 30% contrast: W = 78, p < 0.001, r = 0.58; 40% contrast: W = 58, p < 0.001, r = 0.66).

The categorical effect on the FI-bias

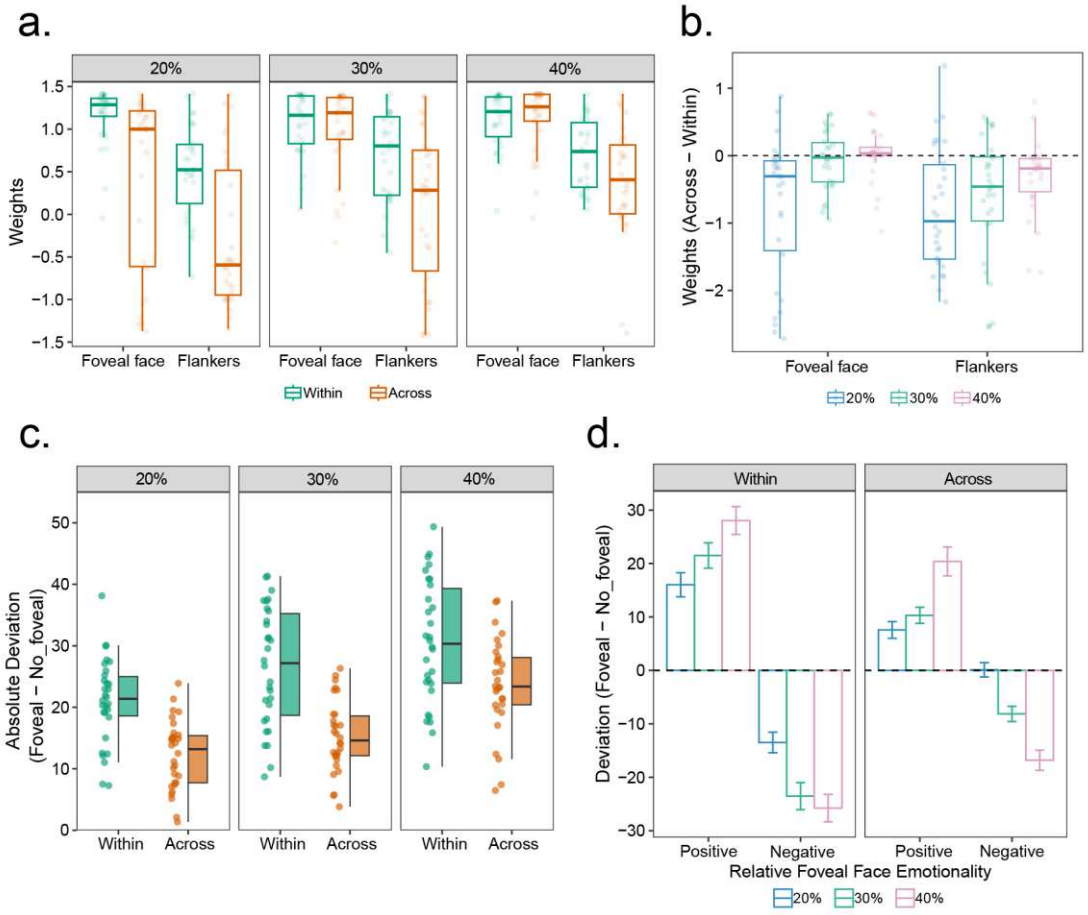
We measured the contrast × category interaction effect by incorporating both factors and their interaction as fixed effects in the linear mixed model, with the absolute deviation serving as the dependent variable (Figure 3c). Random intercept and random slope were added for participant ID. The best-fitting model included only the main effects of contrast (b = 0.52 [0.44, 0.61], SE = 0.04, t(148) = 12.10, p < 0.001, R^2 = 0.50) and category (b = 8.98 [7.60, 10.36], SE = 0.71, t(148) = 12.73, p < 0.001, R^2 =

0.52). Pairwise comparison results showed that the absolute deviation was bigger in the within-category condition compared to the across-category condition (estimate across – within = -8.98, $t(147) = -12.75$, $p < 0.001$), and bigger with increased emotional contrast (estimate contrast 20 – contrast 30 = -4.31, $t(147) = -5.00$, $p < 0.001$; estimate contrast 20 – contrast 40 = -10.46, $t(147) = -12.12$, $p < 0.001$; estimate contrast 30 – contrast 40 = -6.14, $t(147) = -7.12$, $p < 0.001$).

Furthermore, we calculated relative deviations to further examine the FI-bias for each category, contrast, and face combination level. We merged all possible face combinations into two subgroups based on the emotionality of the foveal face relative to the flankers: a positive condition, in which the foveal face was more positive than the flankers, and a negative condition, in which the foveal face was more negative than the flankers. As shown in Figure 3d, in both category conditions, participants perceived the face group as more positive when the foveal face was more positive than the flankers, and vice versa. For each combination of category and relative foveal face emotionality, deviation magnitude increased as contrast increased (within + positive: linear slope $b = 6.01$, $p < 0.001$; within + negative: $b = -6.14$, $p < 0.001$; across + positive: $b = 6.41$, $p < 0.001$; across + negative: $b = -8.46$, $p < 0.001$).

Figure 3

Respective contributions of the foveal face and flankers



Note. **(a)** RMS-based weights under each category, contrast, and location condition. **(b)** Weight deviations between the across- and within-category conditions are plotted as a function of category and contrast. Positive values indicate greater weights in the across-category condition than in the within-category condition. **(c)** Absolute deviation as a function of category and contrast levels. Dots represent individuals' data. The box plot shows the data distribution. **(d)** The deviation is plotted against category, contrast, and the emotionality of the foveal face relative to the flankers. Positive indicates that the foveal face was more positive than the flankers, whereas negative indicates that the foveal face was more negative than the flankers.

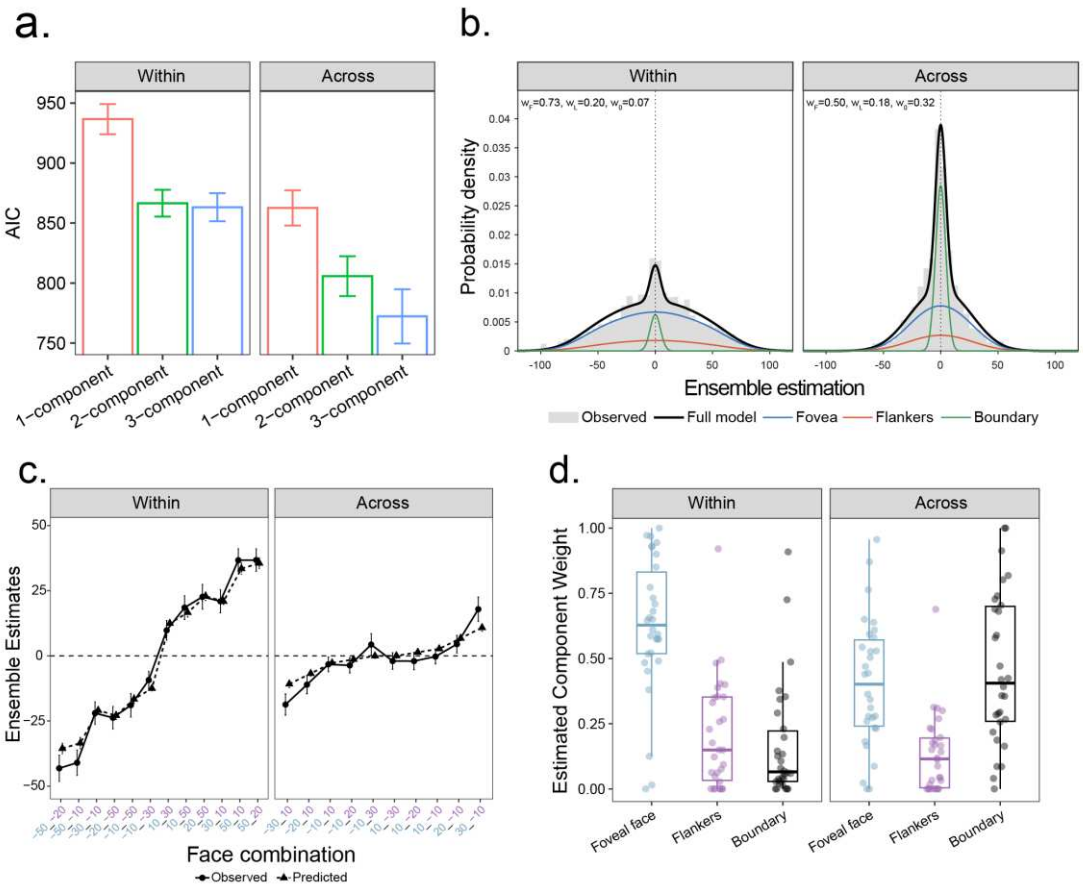
Model fitting

The model was fit at both the individual-participant level and the group level for each category condition, yielding consistent results. Model comparison was conducted using AIC at the group level. In the within-category condition, the 3-component model (863 +/- 11.7) had AIC values comparable to those of the 2-component model (867 +/- 11.2), and both were lower than those of the 1-component model (937 +/- 12.6). In the across-category condition, the 3-component model had the smallest AIC value (772 +/- 22.6) compared to the 2-component (806 +/- 16.6) and 1-component (863 +/- 14.7) models (Figure 4a). Figure 4b shows the group-level three-component model fits for each category condition. Figure S2 (in the Appendix) shows the corresponding model fits for each face combination condition. To further illustrate model fit, we compared participants' observed responses with the responses predicted by the 3-component model. Specifically, for each participant and category condition, we calculated the trial-level predicted response based on the previously fitted 3-component model. We then averaged the observed and model-predicted responses separately for each participant, category, and face combination conditions. These participant-level means were averaged across participants to obtain the group-level observed and predicted responses for each category and face combination condition. The results demonstrate a good model fit (Figure 4c). Last, we examined the estimated weights of the foveal face, flankers, and categorical boundary ('perceived neutral') components in each category condition under the 3-component model (Figure 4d). The linear regression model with estimated component weight as the dependent variable, category and component as the predictors revealed a pronounced interaction effect (b = 0.52 [0.35, 0.69], SE = 0.09, $t(174)$ = 6.04, $p < 0.001$, R^2 = 0.17). The estimated weight of the foveal face component was significantly lower in the across-category condition than in the within-category condition ($p < 0.001$). In contrast, the flanker-component weight was comparable between the two conditions (p = 0.70), whereas the boundary

component weight was lower in the within-category condition than in the across-category condition ($p < 0.001$).

Figure 4

Results of model comparison



Note. **(a)** AIC values for each Gaussian mixture model in the within- and across-category conditions. Error bars represent standard errors of the mean. **(b)** Probability density function (PDF) of ensemble estimations under each category condition and the corresponding best model fit. **(c)** Observed mean ensemble estimates and model-predicted estimates derived from the 3-component model fit, shown separately by category and face combination condition. **(d)** Estimated component weights for the foveal face, flankers, and categorical boundary, derived from the 3-component model.

Discussion

In Experiment 2, we investigated whether the categorical consistency between foveated and non-foveated faces modulates the FI-bias across different levels of emotional contrast. The linear regression model revealed that, in both within- and across-category conditions, the foveal face exerted a stronger influence on ensemble judgments than the flankers, demonstrating a pronounced FI-bias (Pascucci et al., 2021; Dandan et al., 2023; Tiurina et al., 2024). Performance analysis based on the absolute and relative deviations further confirmed the FI-bias (Figures 3c, d). Importantly, emotional contrast did not affect the weights assigned to the foveal face or flankers in the within-category condition, but it strongly modulated these weights in the across-category condition. Specifically, in the across-category condition, the weights of both the foveal face and flankers increased as emotional contrast increased, whereas the difference between them decreased with increasing emotional contrast (Figure 3a). When the foveal face and flankers belonged to the same emotional category, observers appeared to integrate them in a relatively stable manner, with a dominant contribution from the foveal face (Figures 3a, 4d). However, when the faces belonged to different emotional categories, greater contrast may have made both the foveal face and flankers more salient or more perceptually separable. As a result, observers may have relied more on both sources of information, leading to increased contributions from both the foveal face and flankers. Intriguingly, in the across-category condition, observers also assigned more weight to the categorical boundary. The results demonstrate that categorical processing plays an important role in shaping the integration process underlying ensemble perception.

General Discussion

We have previously shown that ensemble perception is shaped by a spatially anisotropic weighting mechanism, whereby foveal items are overweighted than the peripheral items in the same group—a phenomenon termed the foveal input bias (FI-

bias; Dandan et al., 2023). This bias may partly reflect early sensory encoding advantages for foveal information, given the enhanced spatial resolution associated with central vision. In the present study, we extended this account by examining whether and how the FI-bias is modulated by categorical consistency. In Experiment 1, the foveal and the peripheral faces ('flankers') were presented at mixed contrast levels. The results showed that the FI-bias emerged in both within- and across-category conditions, with comparable weights assigned to the foveal face across these conditions. In Experiment 2, we used constant contrast levels to more directly examine the categorical effect. The results showed that (1) the weight of the foveal face was greater than that of the flankers in both category conditions, replicating a pronounced FI-bias; (2) the weighting pattern differed between the two category conditions: in the within-category condition, the weights assigned to the foveal face and flankers remained stable and were unaffected by emotional contrast levels, whereas in the across-category condition, the weights of both components increased as emotional contrast increased; (3) in the across-category condition, observers assigned greater weight to the categorical boundary ('perceived neutral') than in the within-category condition, while assigning less weight to the foveal face. The categorical effect on the observed FI-bias suggests that the bias might not be fixed at the early sensory encoding stage, but remains susceptible to higher-order perceptual influences.

Categorical effect on FI-bias

The present findings suggest that the FI-bias in ensemble perception is not determined solely by lower-level sensory reliability, whereby foveal input receives greater weight because of its higher sensory precision. Rather, it is modulated by higher-level categorical representations, with the categorical relationship between individual items influencing how their information is weighted and integrated into the perceived ensemble. Across both category conditions, foveal information exerted a privileged influence on ensemble judgments, replicating a pronounced FI-bias. However, the magnitude of the FI-bias varied across categorical contexts. The three-

component Gaussian mixture model fits indicated that the across-category condition was characterized by a reduced weighting of the foveal face and increased weighting of the (perceived) neutral point relative to the within-category condition. These findings suggest that ensemble perception does not arise solely from feedforward pooling of local sensory signals. Instead, higher-level categorical representations may influence how information from individual items is weighted and integrated into the perceived ensemble. This interpretation is consistent with the retrospective Bayesian decoding framework (Ding et al., 2017), which proposes that stable, higher-level representations are inferred before being fed back to constrain lower-level sensory representations. Our findings are consistent with this account, suggesting that higher-level categorical representations may retrospectively shape the integration of individual items within the group. Thus, ensemble perception may not arise solely from feedforward pooling of local sensory signals but may also reflect the influence of higher-level categorical representations on lower-level information.

One alternative explanation is that ensemble perception is purely a feedforward pooling process, in which local visual features are automatically combined according to their sensory reliability without recurrent or contextual influences (Utochkin et al., 2024). Under this account, the observed FI-bias would arise solely from the higher spatial resolution and lower sensory uncertainty of foveal vision. However, the present findings are difficult to reconcile with this view. Although the low-level visual properties and retinal locations of the stimuli were matched across two category conditions, the relative weighting of foveal and peripheral faces differed between these conditions. Hence, the observed modulation of the FI-bias by categorical consistency cannot be explained by feedforward pooling alone.

Categorization & Ensemble perception

Previous work has shown that observers can extract ensemble representations from facial crowds even when the expressions span the categorical boundary (Yang et al., 2024). Meanwhile, Utochkin (2015) proposed that ensemble perception provides

the basis for rapid categorization, with the visual system inferring whether an ensemble contains one or multiple categories from the shape of the pooled feature distribution. The present findings reveal a complementary direction of influence: rather than categorization simply emerging from ensemble representations, the established categorical structure also constrains the ensemble representation by modulating the weighting of individual items during perceptual integration.

Information integration vs. segmentation

The current study offers a useful perspective on how categorical structure modulates the integration and segmentation of foveal and peripheral information. According to causal inference accounts (e.g., Alais & Burr, 2004; Körding et al., 2007; Rohe & Noppeney, 2015), signals originating from a common source should be integrated, with each signal weighted according to its relative sensory reliability, whereas signals arising from independent sources should be processed separately. Critically, because the observer does not have direct access to the underlying causal structure, it must be inferred from available cues, including spatiotemporal consistency or higher-order information such as semantic congruency. In this framework, when the foveal and peripheral faces belong to the same category, they are more likely to be inferred as arising from a common source and therefore to be integrated into a unified ensemble representation. In contrast, crossing a categorical boundary may increase the probability that they are treated as distinct sources, promoting separate representations. Under this account, one would expect that responses in the within-category condition would be better captured by a single Gaussian component centered on the ensemble average, whereas responses in the across-category condition would be better described by a two-component Gaussian model reflecting separate foveal and peripheral representations. However, our findings did not support this prediction. Although the two-component Gaussian mixture model provided a fit comparable to that of the three-component model in the within-category condition, response distributions in the across-category condition were better captured by the three-component model,

with an additional component centered at the perceived neutral. These findings suggest that the categorical consistency does not simply determine whether foveal and peripheral information are integrated or segregated. Rather, it modulates the likelihood of responses centered on the categorical boundary, indicating that category membership provides a cue for causal inference that shapes perceptual integration.

The underlying mechanism of FI-bias

The FI-bias likely reflects a reliability-weighted pooling process in which visual signals are integrated according to their relative sensory precision. The human visual system allocates disproportionately greater neural resources to the foveal region as demonstrated by cortical magnification (Carrasco & Frieder, 1997; Carrasco et al., 2004; Kupers et al., 2022). Consequently, foveal input is represented with higher spatial resolution and lower sensory uncertainty than peripheral inputs, making it more likely to receive greater weight during ensemble perception, despite all items being equally relevant. From this perspective, the bias does not necessarily indicate a fixed preference for centrally located information, but rather an adaptive weighting strategy that favors the most reliable signals. It is important to clarify that although the foveal input is given more weight, this does not mean that information from peripheral regions is entirely disregarded, nor does it imply that summary statistics cannot be accurately derived without individual items located at or near the fovea.

Another possible explanation for the FI-bias is an attentional bias toward the foveal face, with greater weighting of the foveal face than of the other face members, despite the task demanding to report the average of all faces. Previous studies demonstrated that attended stimuli are preferentially weighted over unattended stimuli in ensemble perception (de Fockert & Marchant, 2008; Im et al., 2015; Ying, 2022). Here, by presenting only a single face at fixation instead of two or more faces, we sought to maximize attention to the foveal location. The pronounced FI-bias is consistent with the possibility that attention to a single face at fixation strongly interfered with ensemble emotion perception. Increased attention may also contribute

to stronger and/or more accurate visual working memory (VWM) of the foveal face (Lepsien & Nobre, 2007; Thomas et al., 2014; Goldenberg et al., 2022). This in turn may shape the estimation of the average. In the current study, participants were instructed to fixate on a single face, potentially strengthening its representation in visual working memory and thereby increasing it influence on the reported ensemble emotion. Therefore, in addition to differences in sensory reliability, attentional prioritization of foveal information may contribute to the FI-bias.

Limitations and future directions

One limitation of the current study is that we cannot draw firm conclusions about the temporal dynamics of foveal and categorical information processing. Recently, Sama and colleagues (2024) speculated that information from the foveal face may be processed before categorical information is extracted. This account implies that the initial weighting of foveal information should occur largely independently of the categorical relationship between foveal and peripheral inputs. In contrast, previous studies have shown that contextual relationships between foveal and peripheral items can influence the processing of both foveal (e.g., Kroell & Rolfs, 2022, 2025) and peripheral targets (e.g., Sayim et al., 2014). For example, when the foveal and peripheral items match, perception of the peripheral target is enhanced, resulting in reduced crowding (Sayim et al., 2014), an effect that persists even when the foveal item is only imagined (Yildirim-Keles et al., 2026). Such findings indicate that foveal and peripheral information may not be processed independently before their relationship is evaluated. Although this evidence does not establish when categorical information becomes available, it raises the possibility that the categorical relationship between foveal and peripheral inputs may influence their perceptual weighting earlier than what would be predicted by a strictly sequential account. Further research incorporating neural measures would be necessary to clarify the temporal dynamics of foveal and categorical processing and how these processes interact over time. A second limitation is that we cannot rule out the possibility that observers simply

subsampled only a small number of items, such as the foveal face and one of the flankers, rather than integrating information across the entire face group. To examine this possibility, we conducted a control analysis comparing the mean squared error (MSE) between observers' responses and four candidate estimates: the actual mean, the subsampling mean of the foveal face and one randomly selected flanker, the foveal face alone, and the perceived neutral (Figure S3 in the Appendix). The results showed that the smallest errors were obtained for the subsampling and fovea-only estimates, with no significant difference between them. Thus, the control analysis leaves open the possibility that the pronounced FI-bias reflects subsampling rather than graded weighting of information across the entire face group. Specifically, because the fovea-only and subsampling estimates produced similarly small errors, we cannot determine whether observers integrated information from the entire group while assigning disproportionately greater weight to the foveal face, or instead primarily relied on the foveal face and a small number of other items. Further investigation is needed to distinguish between strong foveal weighting and subsampling of the foveal face.

Funding

This work was supported by the Fujian Provincial Social Science Foundation of China (FJ2025B195) to Yu R. Dandan, the Agence Nationale pour la Recherche (ANR-17-EURE-0017 Frontcog and ANR10-IDEX-0001-02 PSL) to Bilge Sayim, and the National Science and Technology Innovation 2030 Major Program (2022ZD0204802) to Fang Fang.

Declarations

Ethics approval and consent to participate The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Psychological and Cognitive Sciences, Peking University (Ref. No. 2024-12-01). All participants provided written informed consent prior participation. Participation

was entirely voluntary, and no incentives were provided. Only participants' gender and age were collected as personal information. All data were collected, stored, and analyzed securely, and no personally identifiable information was disclosed.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

Correspondence

Fang Fang, School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Brain-Computer Interface and Mental Health Modulation, Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China.

Email: ffang@pku.edu.cn

Bilge Sayim, Laboratoire de Sciences Cognitives et Psycholinguistique (LSCP), Département d'Études Cognitives, École Normale Supérieure (ENS), CNRS, 29 rue d'Ulm, 75005 Paris France

Email: bilge.sayim@cnrs.fr

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology*, *14*(3), 257-262. https://doi.org/10.1016/j.cub.2004.01.029

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in cognitive sciences*, *15*(3), 122-131. https://doi.org/10.1016/j.tics.2011.01.003

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967. https://doi.org/10.48550/arXiv.1506.04967

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision research*, *37*(1), 63-82.

https://doi.org/10.1016/S0042-6989(96)00102-2

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature neuroscience*, *7*(3), 308-313. https://doi.org/10.1038/nn1194

Corbett, J. E., Utochkin, I., & Hochstein, S. (2023). The Pervasiveness of Ensemble Perception: Not Just Your Average Review (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009222716

Dandan, Y. R., Ji, L., Song, Y., & Sayim, B. (2023). Foveal vision determines the perceived emotion of face ensembles. *Attention, Perception, & Psychophysics*, *85*(1), 209-221. https://doi.org/10.3758/s13414-022-02614-z

Dandan, Y. R., Li, L., & Sayim, B. (2026). Spatial anisotropy in crowding and ensemble perception. *Journal of Vision*, *26*(7), 15-15. https://doi.org/10.1167/jov.26.7.15

De Fockert, J. W., & Marchant, A. P. (2008). Attention modulates set representation by statistical properties. *Perception & Psychophysics*, *70*(5), 789-794. https://doi.org/10.3758/PP.70.5.789

De Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences*, *108*(32), 13341-13346. https://doi.org/10.1073/pnas.1104517108

Ding, S., Cueva, C. J., Tsodyks, M., & Qian, N. (2017). Visual perception as retrospective Bayesian decoding from high-to low-level features. *Proceedings of the National Academy of Sciences*, *114*(43), E9115-E9124. https://doi.org/10.1073/pnas.1706906114

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*(4), 659-671. https://doi.org/10.1016/S0896-6273(03)00265-4

Edwards, Lloyd J, Keith E Muller, Russell D Wolfinger, Bahjat F Qaqish, and Oliver Schabenberger. (2008). An R2 Statistic for Fixed Effects in the Linear Mixed Model. *Statistics in Medicine* 27 (29): 6137–57. https://doi.org/10.1002/sim.3429

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. https://doi.org/10.1037/h0030377

Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble perception of dynamic emotional groups. *Psychological Science*, *28*(2), 193-203. https://doi.org/10.1177/0956797616678188

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin*, *128*(2), 203. https://doi.org/10.1037/0033-2909.128.2.203

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227-240. 10.1016/0010-0277(92)90002-Y

Fugate, J. M. (2013). Categorical perception for emotional faces. *Emotion Review*, *5*(1), 84-89. https://doi.org/10.1177/1754073912451350

Goldenberg, A., Schöne, J., Huang, Z., Sweeny, T. D., Ong, D. C., Brady, T. F., ... & Gross, J. J. (2022). Amplification in the evaluation of multiple emotional expressions over time. *Nature Human Behaviour*, 1-9. https://doi.org/10.1038/s41562-022-01390-y

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493-498. 10.1111/2041-210X.12504

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current biology*, *17*(17), R751-R753. https://doi.org/10.1016/j.cub.2007.06.039

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of vision*, *9*(11), 1-1. https://doi.org/10.1167/9.11.1

Han, L., Yamanashi Leib, A., Chen, Z., & Whitney, D. (2021). Holistic ensemble perception. *Attention, Perception, & Psychophysics*, *83*(3), 998-1013. https://doi.org/10.3758/s13414-020-02173-1

Harvey, B. M., & Dumoulin, S. O. (2011). The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. *The Journal of Neuroscience*, *31*(38), 13604-13612. https://doi.org/10.1523/JNEUROSCI.2572-11.2011

Himmelberg, M. M., Winawer, J., & Carrasco, M. (2022). Linking individual differences in human primary visual cortex to contrast sensitivity around the visual field. *Nature communications*, *13*(1), 3309. https://doi.org/10.1038/s41467-022-31041-9

Hochmitz, I., Yeshurun, Y., & Yashar, A. (2026). Temporal dynamics of integration and individuation: insights from temporal averaging and crowding. *Cognition*, *268*, 106374. https://doi.org/10.1016/j.cognition.2025.106374

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of vision*, *15*(4), 5-5. https://doi.org/10.1167/15.4.5

Im, H. Y., Park, W. J., and Chong, S. C. (2015). Ensemble statistics as units of selection. *Journal of Cognitive Psychology*, 27(1): 114-127. https://doi.org/10.1080/20445911.2014.985301

Jaeger, Byron C., Lloyd J. Edwards, Kalyan Das, and Pranab K. Sen. 2016. “An (R^2) statistic for fixed effects in the generalized linear mixed model.” *Journal of Applied Statistics* 0 (0): 1–20. https://doi.org/10.1080/02664763.2016.1193725.

Ji, L., Chen, W., Loeys, T., & Pourtois, G. (2018). Ensemble representation for multiple facial expressions: Evidence for a capacity limited perceptual process. *Journal of vision*, *18*(3), 1-19. https://doi.org/10.1167/18.3.17

Jung, W., Bülthoff, I., & Armann, R. G. (2017). The contribution of foveal and peripheral visual information to ensemble representation of face race. *Journal of Vision*, *17*(13), 11-11. https://doi.org/10.1167/17.13.11

Kiffel, C., Campanella, S., and Bruyer, R. (2005). Categorical perception of faces and facial expressions: the age factor. *Exp. Aging Res.* 31, 119–147.

https://doi.org/10.1080/03610730590914985

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS one*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Kroell, L. M., & Rolfs, M. (2022). Foveal vision anticipates defining features of eye movement targets. *Elife*, *11*, e78106. https://doi.org/10.7554/eLife.78106

Kroell, L. M., & Rolfs, M. (2025). The magnitude and time course of pre-saccadic foveal prediction depend on the conspicuity of the saccade target. *Elife*, *12*, RP91236. https://doi.org/10.7554/eLife.91236.4

Kupers, E. R., Benson, N. C., Carrasco, M., & Winawer, J. (2022). Asymmetries around the visual field: From retina to cortex to behavior. *PLoS computational biology*, *18*(1), e1009771. https://doi.org/10.1371/journal.pcbi.1009771

Lepsien, J., & Nobre, A. C. (2007). Attentional modulation of object representations in working memory. *Cerebral cortex*, *17*(9), 2072-2083. https://doi.org/10.1093/cercor/bhl116

Li, H., Ji, L., Tong, K., Ren, N., Chen, W., Liu, C. H., & Fu, X. (2016). Processing of individual items during ensemble coding of facial expressions. *Frontiers in psychology*, 7, 1332. https://doi.org/10.3389/fpsyg.2016.01332

Ni, L., & Stocker, A. A. (2023). Efficient sensory encoding predicts robust averaging. *Cognition*, *232*,105334. https://doi.org/10.1016/j.cognition.2022.105334

Ni, L., & Stocker, A. A. (2026). Dynamics of efficient ensemble coding. *Cognition*, *271*, 106432. https://doi.org/10.1016/j.cognition.2025.106432

Pascucci, D., Ruethemann, N., & Plomp, G. (2021). The anisotropic field of ensemble coding. *Scientific reports*, *11*(1), 8212. https://doi.org/10.1038/s41598-021-87620-1

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS biology*, *13*(2), e1002073.

https://doi.org/10.1371/journal.pbio.1002073

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental brain research*, *37*(3), 495-510. https://doi.org/10.1007/BF00236819

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological bulletin*, *115*(1), 102. https://psycnet.apa.org/doi/10.1037/0033-2909.115.1.102

Sayim, B., Greenwood, J. A., & Cavanagh, P. (2014). Foveal target repetitions reduce crowding. *Journal of Vision*, *14*(6), 4-4. https://doi.org/10.1167/14.6.4

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*(3), 243-265. https://doi.org/10.1016/S0010-0277(98)00069-9

Sutton, T. M., Herbert, A. M., & Clark, D. Q. (2019). Valence, arousal, and dominance ratings for facial stimuli. Quarterly Journal of Experimental Psychology, 72(8), 2046-2055. https://doi.org/10.1177/1747021819829012

Thomas, P. M., Jackson, M. C., & Raymond, J. E. (2014). A threatening face in the crowd: effects of emotional singletons on visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 253. https://psycnet.apa.org/doi/10.1037/a0033970

Timeo, S., Farroni, T., & Maass, A. (2017). Race and color: Two sides of one story? Development of biases in categorical perception. *Child development*, *88*(1), 83-102. https://doi.org/10.1111/cdev.12564

Tiurina, N. A., Markov, Y. A., Whitney, D., & Pascucci, D. (2024). The functional role of spatial anisotropies in ensemble perception. *BMC biology*, *22*(1), 28. https://doi.org/10.1186/s12915-024-01822-3

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from

untrained research participants. *Psychiatry research*, *168*(3), 242-249. https://doi.org/10.1016/j.psychres.2008.05.006

To, M. P., Carvey, K. M., Carvey, R. J., & Liu, C. H. (2019). Averaging sets of expressive faces is modulated by eccentricity. *Journal of Vision*, *19*(11), 2-2. https://doi.org/10.1167/19.11.2

Utochkin, I. S. (2015). Ensemble summary statistics as a basis for rapid visual categorization. *Journal of Vision*, *15*(4), 8-8. https://doi.org/10.1167/15.4.8

Utochkin, I. S., Choi, J., & Chong, S. C. (2024). A population response model of ensemble perception. *Psychological review*, *131*(1), 36. https://psycnet.apa.org/doi/10.1037/rev0000426

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual review of psychology*, *69*(1), 105-129. https://doi.org/10.1146/annurev- psych- 010416- 044232

Wolfe, B. A., Kosovicheva, A. A., Leib, A. Y., Wood, K., & Whitney, D. (2015). Foveal input is not required for perception of crowd facial expression. *Journal of Vision*, *15*(4), 11-11. https://doi.org/10.1167/15.4.11

Yang, Z., Wu, Y., Liu, S., Zhao, L., Fan, C., & He, W. (2024). Ensemble coding of crowd with cross-category facial expressions. *Behavioral Sciences*, *14*(6), 508. https://doi.org/10.3390/bs14060508

Ying, H. (2022). Attention modulates the ensemble coding of facial expressions. *Perception*, *51*(4), 276-285. https://doi.org/10.1177/03010066221079686

Yildirim-Keles, F. Z., Aschwanden, R., & Sayim, B. (2026). improvement of peripheral visual discrimination through mental imagery. *Iscience*, *29*(7). https://doi.org/10.1016/j.isci.2026.116437

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx