Within-hemifield perceptual averaging of facial expressions predicted by neural averaging

Timothy D. Sweeny,
Department of Psychology, Evanston, IL, USA

Marcia Grabowecky,
Department of Psychology and Interdepartmental Neuroscience Program, Evanston, IL, USA

Ken A. Paller, and
Department of Psychology and Interdepartmental Neuroscience Program, Evanston, IL, USA

Satoru Suzuki
Department of Psychology and Interdepartmental Neuroscience Program, Evanston, IL, USA

Abstract

High-level visual neurons in the ventral stream typically have large receptive fields, supporting position-invariant object recognition but entailing poor spatial resolution. Consequently, when multiple objects fall within their large receptive fields, unless selective attention is deployed, their responses are averages of responses to the individual objects. We investigated a behavioral consequence of this neural averaging in the perception of facial expressions. Two faces (7°-apart) were briefly presented (100-ms, backward-masked) either within the same visual hemifield (within-hemifield condition) or in different hemifields (between-hemifield condition). Face pairs included happy, angry, and valence-neutral faces, and observers rated the emotional valence of a post-cued face. Perceptual averaging of facial expressions was predicted only for the within-hemifield condition because the receptive fields of 'face-tuned' neurons are primarily confined within the contralateral field; the between-hemifield condition served to control for post-perceptual effects. Consistent with averaging, valence-neutral faces appeared more positive when paired with a happy face than when paired with an angry face, and affective intensities of happy and angry faces were reduced by accompanying valence-neutral or opposite-valence faces, in the within-hemifield relative to the between-hemifield condition. We thus demonstrated within-hemifield perceptual averaging of a complex feature as predicted by neural averaging in the ventral visual stream.

Keywords

object recognition; face recognition; shape and contour; ventral visual pathway; neural averaging; inferotemporal cortex

Introduction

Everyday perception is not like a slow-paced slide-show of single static images. Rather, constant eye movements (e.g., Yarbus, 1967), the vast number of objects in typical

Corresponding author: Satoru Suzuki. satoru@northwestern.edu. Address: Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208, USA..

environments, unpredictable locations of potentially interesting stimuli, and the time necessary to engage voluntary attention (e.g., Chelazzi, Duncan, Miller, & Desimone, 1998; Suzuki & Cavanagh, 1997) all interact such that objects are often seen briefly and without a clear focus of attention. This lack of focused attention poses a potential problem for processing of behaviorally relevant complex visual features such as facial expressions.

Facial expressions are encoded by neurons in temporal visual cortex (e.g., Hasselmo, Rolls, & Baylis,1989; Streit et al., 1999) that have large receptive fields (e.g., Boussaoud, Desimone, & Ungerleider, 1991; Chelazzi et al., 1998; Desimone & Gross, 1979; Niemeier, Goltz, Kuchinad, Tweed, & Vilis, 2005; Op De Beeck & Vogles, 2000). Large receptive fields allow neuronal responses to be relatively independent of stimulus position, thereby achieving position invariant coding of complex patterns. However, large receptive fields also incur a cost of reduced spatial resolution, and this is problematic in a typical visual environment where multiple objects are present. Fortunately, selective attention can be engaged to enhance signals from an attended object while suppressing signals from other objects, so that a high-level ventral visual neuron then responds as if only the attended object is present (e.g., Chelazzi et al., 1998).

When there is insufficient time to engage selective attention, however, as frequently happens in fleeting glances, high-level ventral visual neurons cannot resolve the multiple patterns presented within their large receptive fields. Specifically, they respond as if they 'average'across those patterns (e.g., Kastner, De Weerd, Pinsk, Elizondo, Desimone, & Ungerleider, 2001; Miller, Gochin, & Gross, 1993; Rolls & Tovee, 1995; Sato, 1989; Zoccolan, Cox, & DiCarlo, 2005). For example, suppose a high-level ventral neuron (e.g., one tuned to angry facial expressions) responds to its preferred pattern (e.g., an angry face) at 20 spikes/sec and to a non-preferred pattern (e.g., a valence-neutral face) at 8 spikes/sec, when each pattern is presented alone. This neuron would respond at about 14 spikes/sec (an average of 20 spikes/ sec and 8 spikes/sec) when both patterns are simultaneously presented within its receptive field. We investigated whether this neural averaging prevalent in high-level ventral visual processing results in perceptual averaging of complex visual features. We examined perception of facial expressions because they are both behaviorally important and are known to be encoded by high-level ventral visual neurons in temporal cortex (e.g., Hasselmo et al., 1989; Streit et al., 1999).

We took advantage of the fact that high-level ventral visual neurons that respond to complex patterns, including those that respond to faces, have large receptive fields often encompassing substantial regions of the contralateral visual field (e.g., Boussaoud et al., 1991; Chelazzi et al.,1998; Desimone & Gross, 1979; Niemeier et al., 2005; Op De Beeck & Vogels, 2000). Thus, when a pair of faces is presented within the left or right visual hemifield—the within-hemifield condition—both faces should fall within the same receptive fields of many high-level ventral visual neurons.

Furthermore, although some inferotemporal neurons have receptive fields that extend into the ipsilateral visual field, receptive fields of many inferotemporal neurons are confined within the contralateral visual field (e.g., Boussaoud et al., 1991; Desimone & Gross, 1979; DiCarlo & Maunsell, 2003; Kastner et al., 2001). In particular, when one stimulus is presented in the contralateral field and the other in the ipsilateral field, responses of inferotemporal neurons are nearly completely dominated by the contralateral stimulus, regardless of selective attention (e.g., Chelazzi et al., 1998). Thus, when a pair of faces is presented on opposite sides of the vertical meridian—the between-hemifield condition—each stimulus should drive a different set of neurons in the contralateral cerebral hemisphere.

This between-hemifield condition served to control for post-perceptual effects concerning arousal, emotional biasing, and/or semantic interactions, which should be similar for the within-hemifield and between-hemifield conditions. Only in the within-hemifield condition, in which two faces fell within the same receptive fields of many high-level ventral visual neurons, was perceptual averaging predicted. We thus compared the within- and between-hemifield conditions to evaluate perceptual averaging.

Each experimental trial (Figure 1A) included a pair of simultaneously presented faces. In most experiments (Experiments 1, 2, and 4), one face always had a clear valence of expression, positive (happy) or negative (angry), and the other had a valence-neutral (surprise) expression (Figure 1B). These pairings allowed us to examine the complementary manifestations of perceptual averaging predicted by neural averaging.

Suppose that perception of positive (or negative) valence is mediated by responses of positive-expression-tuned (or negative-expression-tuned) neurons that respond strongly to happy (or angry) faces and weakly to valence-neutral faces. When a happy (or angry) face and a valence-neutral face are simultaneously presented within these neurons' receptive fields, they would respond at an intermediate level, making both faces appear somewhat happy (or angry). Consequently, a valence-neutral face should appear to take on the valence of the paired happy (or angry) face—*perceptual spreading*, and at the same time, the affective intensity of the happy (or angry) face should appear to be reduced—*perceptual reduction*. Both the spreading and reduction aspects of perceptual averaging should occur in the within-hemifield condition (where the face pairs are expected to mostly activate single neural receptive fields) compared to the between-hemifield condition (where individual faces are expected to mostly activate separate neural receptive fields). The perceptual-spreading effect predicted by neural averaging was examined in Experiments 1 and 4 by post-cueing observers to report the perceived valence of valence-neutral (surprise) faces paired with happy or angry faces. The perceptual-reduction effect predicted by neural averaging was examined in Experiment 2 by post-cueing observers to report the perceived valence of the happy or angry faces paired with valence-neutral (surprise) faces. To show that perceptual averaging does not depend on the use of valence-neutral faces and also to control for a potential attention strategy, we paired happy faces with angry faces in Experiment 3 (and predicted reduced affective intensity for both expressions in the within-hemifield condition). Influences of potentially confounding factors of distribution of spatial attention and stimulus configuration were ruled out in Experiment 4.

Experiment 1: Within-hemifield perceptual spreading predicted by neural averaging

Within-receptive-field neural averaging in high-level ventral visual areas predicts that a valence-neutral surprise face should take on the emotional valence of an accompanying happy or angry face, but only when the faces are presented within the same visual hemifield.

Methods

Observers—Twenty-four undergraduate students at Northwestern University participated for course credit. Each observer had normal or corrected-to-normal visual acuity, gave informed consent, and was tested individually in a dimly lit room.

Stimuli—We selected three categories of faces from the Karolinska Directed Emotion Face Set (Lundqvist, Flykt, &Öhman, 1998) according to their emotional expressions: 14 happy, 14 angry, and 28 valence-neutral (surprise) faces. We used surprise faces as valence-neutral faces because of their bistability in perceived valence. For example, a surprise face can appear joyful (e.g., in reaction to a surprise party) or distressed (e.g., in reaction to bad news). This bistable

ambiguity makes surprise faces a sensitive instrument to measure contextual effects that induce positive or negative facial expressions (Kim, Somerville, Johnstone, Polis, Alexander, Shin, & Whalen, 2004).

Stimuli were color photographs of faces of different individuals (half female, half male). Each face was cropped using an elliptical mask to exclude hair (which distracts attention from the emotionally relevant features of the face, Tyler & Chen, 2006). The faces were then scaled to be the same size. Each face subtended 2.75° (width) by 3.95° (height) of visual angle and was embedded in a box of colored Gaussian noise slightly larger than the face, subtending 2.98° (width) by 4.24° (height) of visual angle. The noise boxes also served as backward masks (Figure 1A).

In a pilot experiment, we verified the perceived emotional valence of the three expression categories. Each face was presented for 800 ms at the center of the display monitor with an inter-stimulus interval of 1600-2400 ms and with no mask. Eleven observers rated the faces using a 4-point scale from 1 (most negative) to 4 (most positive). The mean ratings were 3.65 ($SD = 0.14$) for the happy faces, 1.42 ($SD = 0.29$) for the angry faces, and 2.40 ($SD = 0.17$) for the surprise faces. Note that the mean valence rating for the surprise faces was about halfway between the extremes of the rating scale, confirming that our surprise faces appeared neutral in valence.

In the main experiment, a pair of faces was presented either within the same visual hemifield or in opposite hemifields, always at 4.73° eccentricity (from the fixation cross to the center of the face). Placement of faces on an invisible approximate iso-acuity ellipse ensured that all faces were viewed with similar acuity (Rovamo & Virsu,1979). In the within-hemifield condition, faces were vertically arranged within either the left or right visual hemifield (left or right, equiprobable) with an inter-face distance of 7.14° (center to center). In the between-hemifield condition, faces were horizontally arranged across either the upper or lower visual hemifield (upper or lower, equiprobable) with an inter-face distance of 6.22° (center to center). The inter-face distance was deliberately set to be slightly shorter for the between-hemifield condition than for the within-hemifield condition (Figure 1A). In this way, any potential retinotopic interactions would be slightly weaker in the within-hemifield condition so that such interactions could not explain perceptual averaging selectively obtained for the within-hemifield condition.

The within- and between-hemifield conditions were run in separate blocks (1 block per condition, 28 trials per block) to avoid potential confusion at the response stage by minimizing the information content of the post-face cue; the post-face cue pointed either up or down in the within-hemifield block, or pointed either left or right in the between-hemifield block. Because the to-be-rated face was presented randomly in the four quadrants in both the within-hemifield and between-hemifield conditions, it is unlikely that the blocking made the pre-trial distribution of attention different for the two conditions (note that we replicated the results without blocking in Experiment 4). Block order was counterbalanced across observers.

On each trial, a valence-neutral surprise face was paired with either a happy or angry face of the same sex. These face pairings were counterbalanced so that each surprise face was paired with a happy or angry face equally often across observers. Pairings of the faces and the order in which different pairs were presented were kept constant across the two blocks for all observers, so that the within- and between-hemifield conditions could be compared with otherwise identical stimulus conditions. Stimuli were presented on a 19" CRT monitor at a viewing distance of 100 cm.

Procedure—Each trial began with the presentation of a white fixation cross on a black screen for 500 ms (Figure 1A). The experimenter strongly encouraged observers to fixate the central cross, emphasizing that looking at any other location on the screen would make the task more difficult (as the to-be-rated face could appear in any of the four quadrants). Note that potentially unstable central fixation would not have produced a confound, and if anything, would have reduced the magnitude of perceptual averaging by obscuring the within- and between-hemifield placements of the face pairs. A face pair (i.e., a valence-neutral surprise face and a happy or angry face) then appeared for 100 ms and was immediately backward masked with Gaussian noise masks lasting 300 ms. This brief stimulus presentation did not allow time for saccades to the face stimuli.

Upon offset of the noise masks, an arrow replaced the fixation cross for 800 ms, post-cueing the valence-neutral surprise face as the face to be rated on its emotional valence. Observers used the following rating scale:

1. most negative,

2. moderately negative,

3. moderately positive, and

4. most positive.

In order to encourage observers to make subtle discriminations, the rating scale did not include a 'neutral' response option. If observers perceived the surprise faces as valence neutral, their responses across trials should have averaged to the midpoint of the scale (or have shown a response bias for positive or negative responses, but such bias should not differ between the within- and between-hemifield conditions). The experimenter instructed observers to respond as quickly as possible and to rely on their 'gut feeling' if they were unsure. To avoid potential response biases due to spatial placement of response buttons, ratings were spoken and the experimenter recorded the responses while sitting out of sight. Six practice trials preceded each of the two (within-hemifield and between-hemifield) experimental blocks.

The brief stimulus presentation, backward masking, random location of the to-be-rated face, and post-cueing procedure prevented observers from systematically attending to the faces to be rated. This experimental procedure thus met the conditions for neural averaging (i.e., a broad distribution of attention during stimulus presentation).

Results

Valence-neutral (surprise) faces were rated more positively when accompanied by a happy face than when accompanied by an angry face in the within-hemifield condition, $t_{23} = 2.836$, $p < 0.01$, but not in the between-hemifield condition, $t_{23} = 0.241$, n.s. (Figure 2). This pattern of results was confirmed by a significant inter-action in a repeated-measures ANOVA with the experimental condition (within- vs. between-hemifield) and valence of the accompanying face (happy vs. angry) as the two factors, $F_{1,23} = 6.903$, $p < 0.02$.

The degree of perceptual spreading of emotional valence to surprise faces (measured as changes in valence ratings for surprise faces from the between-hemifield condition to the within-hemifield condition) appears to be larger for angry expressions than for happy expressions in Figure 2, but this difference was not significant, $t_{23} = 1.298$, *n.s.* Perceptual spreading was slightly larger for stimuli presented in the right compared to the left visual hemifield (a marginal three-way interaction among the experimental condition [within- vs. between-hemifield], valence of the accompanying face [happy vs. angry], and the visual hemifield [left vs. right] in which the rated valence-neutral faces were presented, $F_{1,23} = 3.545$, $p < 0.073$).

Overall, the results from Experiment 1 demonstrated perceptual averaging in terms of perceptual spreading of a positive (or negative) expression from a happy (or angry) face to a valence-neutral face. Importantly, this perceptual spreading occurred only when the faces were presented within the same visual hemifield, consistent with the prediction based on within-receptive-field neural averaging occurring in high-level ventral visual areas.

Experiment 2: Within-hemifield perceptual reduction predicted by neural averaging

Within-receptive-field neural averaging in high-level ventral visual areas also makes a complementary prediction that the affective intensity of a happy (or angry) face should be reduced by an accompanying valence-neutral face when faces are presented within the same visual hemifield compared to when faces are presented in opposite hemifields.

Methods

Observers, Stimuli, and Procedure—A new group of 20 observers was recruited. The design and procedure were identical to Experiment 1, except that the post-face cue on each trial instructed observers to rate the perceived valence of the happy or angry face instead of the valence-neutral face.

Results

Happy faces were rated less positively and angry faces less negatively when the accompanying valence-neutral face was presented within the same visual hemifield compared to when the valence-neutral face was presented in the opposite hemifield (Figure 3). This pattern of results was confirmed by a significant interaction in a repeated-measures ANOVA with the experimental condition (within- vs. between-hemifield) and valence of the emotional face (happy vs. angry) as the two factors, $F_{1,19} = 4.558$, $p < 0.05$.

As with perceptual spreading, the degree of perceptual reduction of affective intensity (how much the positive rating of a happy face or the negative rating of an angry face was reduced in the within-hemifield condition compared to the between-hemifield condition) did not differ significantly between the happy and angry expressions, $t_{19} = 0.457$, *n.s.* Nor did perceptual reduction depend on the visual hemifield in which the to-be-rated face was presented (a non-significant three-way interaction among the experimental condition [within- vs. between-hemifield], valence of the emotional face [happy vs. angry], and the visual hemifield [left vs. right] in which the rated emotional faces were presented, $F_{1,19} = 0.313$, *n.s.*).

Overall, the results from Experiment 2 demonstrated perceptual averaging in terms of perceptual reduction of affective intensity of a happy (or angry) face due to simultaneous presentation of a valence-neutral face within the same visual hemifield.

Experiment 3: A replication of within-hemifield perceptual reduction without the use of valence-neutral faces and controlling for a potential attention strategy

In Experiments 1 and 2, an emotionally valenced (happy or angry) face was always paired with a valence-neutral (surprise) face. To show that perceptual averaging is not limited to the use of an affectively ambiguous expression, here we used pairs composed of a happy and an angry face. Neural averaging predicts that the affective intensities of the happy and angry faces should both be reduced in the within-hemifield condition compared to the between-hemifield condition.

This experiment also controlled for a potential attention strategy that observers might have employed in Experiments 1 and 2.In Experiment 1 valence-neutral (surprise) faces were always post-cued and in Experiment 2 emotionally valenced faces (happy or angry faces) were always post-cued to be rated on their emotional valence. It is thus possible that within the 100-ms presentation of the faces, observers might have attempted to identify expressions of both faces and shifted attention to the likely target face (i.e., the one with the weaker valence in Experiment 1 and the one with the stronger valence in Experiment 2). We note, however, that it is unlikely that observers would have been able to identify the two expressions, determine to which face to attend, and complete shifting of attention, all within 100 ms. Even if that were possible, shifting of attention to the target face would have only reduced the effect of neural averaging (see the Introduction section), and it is also difficult to argue that the act of shifting attention would have produced the perceptual averaging effects obtained inExperiments 1 and 2. Nevertheless, a demonstration of perceptual averaging in this experiment would alleviate this potential concern. Because each pair included one happy face and one angry face and either expression was post-cued with an equal probability, it would have been impossible for observers to use any strategy to shift attention to the to-be-rated face in this experiment.

Methods

Observers, Stimuli, and Procedure—A new group of 26 observers was recruited. The design and procedure were identical to those in Experiments 1 and 2, except that the valence-neutral (surprise) faces were replaced by happy and angry faces, so that each face pair included a happy face and an angry face of the same gender. As in Experiments 1 and 2, each face pair was presented twice in both the within- and between-hemifield conditions, so that in each condition observers were postcued to rate both faces from each pair across trials.

Results

Happy faces were rated less positively and angry faces less negatively when the accompanying face of opposite emotional valence was presented within the same visual hemifield compared to when the accompanying face was presented in the opposite hemifield (Figure 4). This pattern of results was confirmed by a significant interaction in a repeated-measures ANOVA with the experimental condition (within- vs. between-hemifield) and valence of the rated face (angry vs. happy) as the two factors, $F_{1,25} = 4.827$, $p < 0.05$.

As with Experiment 2, the degree of perceptual reduction of affective intensity (in the within-hemifield condition compared to the between-hemifield condition) did not differ significantly between the happy and angry expressions, $t_{25} = 1.704$, *n.s.* Perceptual reduction, however, was slightly larger for stimuli presented in the left compared to the right visual hemifield (a marginal threeway interaction among the experimental condition [within- vs. between-hemifield], valence of the rated [happy vs. angry] face, and the visual hemifield [left vs. right] in which the rated faces were presented, $F_{1,25} = 3.291$, $p < 0.09$).

Overall, the results from Experiment 3 replicated the perceptual-reduction aspect of within-hemifield averaging in that the affective intensity of a happy (or angry) face was reduced when an accompanying face of opposite valence was simultaneously presented within the same visual hemifield compared to when it was presented in the opposite hemifield. Thus, the phenomenon of within-hemifield perceptual averaging of facial expressions is not limited to the use of valence-neutral (surprise) faces. Furthermore, because it was impossible for observers to determine the to-be-rated face prior to presentation of the post-face cue, any attention strategy other than spreading attention across the display prior to presentation of a face pair and attending to both faces would have been ineffective. This suggests that the perceptual-averaging effects obtained in Experiments 1 and 2 are unlikely to be an artifact of an attention strategy.

Experiment 4: A replication of within-hemifield perceptual spreading, with the between- and within-hemifield conditions intermixed and face alignment jittered

We have demonstrated within-hemifield perceptual averaging consistent with within-hemifield neural averaging by high-level ventral visual neurons, in terms of perceptual spreading (a happy or angry expression of a face spreading to a paired valence-neutral [surprise] face; Experiment 1) and perceptual reduction (the affective intensity of a happy or angry face being reduced by an accompanying valence-neutral face, Experiment 2, or by an accompanying face of opposite valence, Experiment 3). In this experiment we attempted to replicate the perceptual-spreading result while controlling for two factors that could have potentially confounded our results in Experiments 1-3.

First, the within-hemifield and between-hemifield conditions were blocked in Experiments 1-3. Consequently, the face pairs were always vertically arranged in the within-hemifield block and horizontally arranged in the between-hemifield block. We blocked these conditions to simplify the post-cue (always pointing up or down in the within-hemifield block, and pointing left or right in the between-hemifield block). Although the post-cued faces were unpredictably presented across the four quadrants, the blocking allowed observers to anticipate either the vertical or horizontal arrangement, and our results might potentially depend on such anticipation generating different spatial attention distributions in the two conditions. In this experiment, we randomly intermixed the within- and between-hemifield presentations of face pairs, so that the arrangement of the face pair as well as the location of the to-be-rated face were unpredictable on a trial-to-trial basis. This ensured that spatial attention was similarly and evenly distributed in the two conditions.

Second, the faces within each pair were always perfectly aligned in Experiments 1-3 (Figure 1A). Because the faces were horizontally aligned in the between-hemifield condition and vertically aligned in the within-hemifield condition, it is possible that perceptual averaging might depend on vertical alignment of facial features (eyes, nose, mouth, etc.) rather than on within-hemifield presentations of faces. To evaluate this potential confound, we randomly jittered the *X* and *Y* coordinates of each face so that the faces were always misaligned in an unpredictable manner. If the within-hemifield perceptual-spreading effect obtained in Experiment 1 depended on vertical alignments of facial features, perceptual spreading should be eliminated (or substantially reduced) in this experiment. In contrast, if the effect was due to within-hemifield neural averaging, it should be unaffected by misalignment of faces.

Methods

Observers, Stimuli, and Procedure—A new group of 22 observers was recruited. The design and procedure were identical to those in Experiment 1, except that (1) within- and between-hemifield trials were randomly intermixed and run in a single block, and (2) the location of each face on each trial was randomly jittered both vertically and horizontally around the locations used in Experiment 1 by an amount randomly sampled from a uniform distribution between ±0.52° (jitter amounts were independently sampled for *X* and *Y* coordinates). The maximum displacement for a vertical pair, for example, resulted in an eye of one face being aligned with the nose of the other face, thus substantially disrupting vertical alignments of facial features.

Results

Valence-neutral (surprise) faces were rated more positively when accompanied by a happy face than when accompanied by an angry face in the within-hemifield condition, $t_{21} = 2.152$,

$p < 0.05$, but not in the between-hemifield condition, $t_{21} = 0.842$, *n.s.* (Figure 5). This pattern of results was confirmed by a significant interaction in a repeated-measures ANOVA with the experimental condition (within- vs. between-hemifield) and valence of the accompanying face (happy vs. angry) as the two factors, $F_{1,21} = 6.745$, $p < 0.05$.

As with Experiment 1, the degree of perceptual spreading of emotional valence to surprise faces (in the within-hemifield condition compared to the between-hemifield condition) did not differ significantly between the happy and angry expressions $t_{21} = 0.776$, n.s. Perceptual spreading was larger for stimuli presented in the right compared to the left visual hemifield (a significant three-way interaction among the experimental condition [within-vs. between-hemifield], valence of the accompanying face [happy vs. angry], and the visual hemifield [left vs. right] in which the rated valence-neutral faces were presented, $F_{1,21} = 8.333$, $p < 0.01$). We note that there was little consistency across the four experiments in terms of the degree of perceptual spreading or reduction as a function of the valence (positive or negative) of the rated or accompanying faces. Nor was there a reliable across-experiment consistency in terms of a right or left hemifield advantage for perceptual spreading or reduction (except that the right-hemifield advantage for perceptual spreading was marginal in Experiment 1 and significant in Experiment 4). Our results are thus inconclusive regarding these secondary effects.

Overall, the results from Experiment 4 replicated the perceptual-spreading aspect of within-hemifield averaging demonstrated in Experiment 1. The fact that we replicated perceptual spreading while randomly intermixing the within- and between-hemifield conditions suggests that our results from Experiments 1-3 (in which these conditions were blocked) are unlikely to be attributable to different distributions of spatial attention that could have been employed in the within- and between-hemifield conditions. Furthermore, the successful replication while faces were randomly misaligned suggests that the within-hemifield averaging effects we obtained in Experiments 1-3 are unlikely to be attributable to vertical alignments of facial features.

Discussion

We demonstrated perceptual averaging of emotional valence using pairs of faces that were simultaneously and briefly presented, without allowing sufficient time for observers to systematically focus attention on a single face (a condition necessary to induce neural averaging in high-level ventral visual areas). Specifically, we demonstrated two complementary manifestations of perceptual averaging, (1) perceptual spreading in which the valence-neutral (surprise) face appeared to take on the emotional valence of the accompanying happy or angry face, and (2) perceptual reduction in which the happy face appeared less happy and the angry face appeared less angry due to the accompanying valence-neutral or opposite-valence face. Perceptual spreading and reduction together indicate that perception of both valence-neutral and emotional faces moved toward the average expression for each pair.

Crucially, this perceptual averaging occurred only when the two faces were presented within the same visual hemifield. Furthermore, averaging occurred when the two faces were not in close proximity but were separated by a distance of 7° visual angle, indicating a long-range interaction within each visual hemifield. These results are consistent with prior evidence of neural averaging in high-level ventral visual processing. High-level ventral visual neurons that encode complex visual features such as facial expressions have large, primarily contralateral receptive fields. When these neurons are activated by multiple stimuli in the absence of attention focused on one stimulus, their firing rates reflect the average of those in response to the individual stimuli (e.g., Chelazzi et al.,1998; Kastner et al., 2001; Miller et al., 1993; Sato, 1989; Zoccolan et al., 2005). Our finding of within-hemifield perceptual averaging of facial

expressions can thus be explained on the basis of within-receptive-field averaging by high-level ventral visual neurons whose receptive fields are large and primarily contralateral.

Nevertheless, one might argue for an alternative explanation based on differential attention resources between and within hemifields. For example, there is evidence suggesting that each cerebral hemisphere has a relatively independent resource for visual spatial attention and short-term memory, especially for split-brain individuals but also for normal individuals to various degrees depending on the behavioral task (e.g., Alvarez & Cavanagh, 2005; Delvenne, 2005; Duncan, Bundesen, Olson, Humphreys, Chavda, & Shibuya, 1999; Luck, Hillyard, Mangun, & Gazzaniga, 1989). It is thus possible that in the within-hemifield condition, the two faces might have competed for a more limited attention resource within one cerebral hemisphere, whereas in the between-hemifield condition each face might have been processed by a relatively separate attention resource in the contralateral cerebral hemisphere. Although our results do not definitively reject this attention-resource hypothesis, we favor the neural-averaging hypothesis because it naturally predicts the phenomenon of perceptual averaging, whereas the attention-resource hypothesis must postulate an additional mechanism whereby competition for a limited attention resource results in perceptual averaging (rather than, for example, perceptual degradation).

Other alternative explanations of our results based on post-perceptual processing, attentional strategies, and spatial interactions are unlikely. Our results cannot be explained in terms of post-perceptual effects such as arousal, emotional biasing, and/or semantic interactions because these effects should have been operative to a similar extent in both the within-hemifield and between-hemifield conditions.

Nor can the results be explained in terms of a top-down attention strategy. Observers could not have adjusted their spatial distribution of attention prior to stimulus presentation because the to-be-rated (i.e., post-cued) face was always unpredictably presented in any of the four quadrants on each trial. Especially in Experiment 4 where the within- and between-hemifield conditions were randomly intermixed, observers could not have anticipated the spatial configuration (vertical or horizontal) of each face pair. Thus, neither pre-stimulus distribution of spatial attention nor anticipation of spatial configuration could account for our perceptual-averaging results. It is also unlikely that observers adopted a post-stimulus attention strategy. Each face pair was only briefly presented (100 ms and backward masked) at an unpredictable location; it is thus unlikely that observers had sufficient time to identify the expressions of the two peripherally flashed faces and to complete shifting of attention to the type of face that would be post-cued. Even if observers were able to do this to a limited degree in Experiments 1, 2, and 4, focusing attention on the to-be-rated face would only have reduced the effect of neural averaging (e.g., Chelazzi et al., 1998), and it is unclear how an act of shifting attention itself would have caused perceptual averaging. Furthermore, in Experiment 3, the to-be-rated face was completely unpredictable until the post-cue on each trial, but we still obtained evidence of perceptual averaging. Thus, neither pre-stimulus distribution of spatial attention nor post-stimulus shifting of attention could account for our perceptual-averaging effects.

The results are also unlikely to be attributable to retinotopic or spatial interactions. For example, the inter-face distance was slightly shorter in the between-hemifield condition than in the within-hemifield condition, but we obtained evidence of perceptual averaging only in the within-hemifield condition. Furthermore, perceptual spreading occurred whether the face locations were aligned (Experiment 1) or randomly jittered (Experiment 4), indicating that perceptual averaging is mediated by mechanisms that are relatively insensitive to random shifts in location. This is consistent with our hypothesis that within-hemifield perceptual averaging reflects within-receptive-field averaging by high-level ventral visual neurons whose responses

are primarily confined to the contralateral visual hemifield but are otherwise largely insensitive to small changes in stimulus locations.

Because neural averaging occurs throughout the ventral visual pathway whenever multiple stimuli fall within a single receptive field and no particular stimulus is selectively attended (e.g., Chelazzi et al., 1998; Kastner et al., 2001; Miller et al., 1993; Reynolds, Chelazzi, & Desimone, 1999; Sato, 1989; Zoccolan et al., 2005), it is plausible that perceptual averaging is a ubiquitous phenomenon affecting perception of simple as well as complex visual features. In fact, Parkes, Lund, Angelucci, Solomon, and Morgan (2001) demonstrated short-range (0.47° visual angle) perceptual averaging of local orientation consistent with small receptive fields in low-level visual areas. Because neural receptive fields become larger in higher level ventral visual areas that tend to process more complex visual features (see Suzuki, 2005 and Tanaka, 1996 for reviews), the spatial extent of perceptual averaging is likely to be larger for more complex features (at least 7° for facial expressions based on our current results) than for simpler features. Furthermore, neural tunings for complex features in high-level ventral visual neurons tend to develop in response to behavioral demands (e.g., Kobatake, Wang, & Tanaka, 1998; Logothetis, Pauls, & Poggio, 1995; Sigala & Logothetis, 2002; Young & Yamane, 1992). Thus, longrange averaging of complex features in fleeting glances might be advantageous by allowing people to rapidly perceive the gist of behaviorally relevant information, prior to localizing individual objects.

Acknowledgments

References

Alvarez GA, Cavanagh P. Independent resources for attentional tracking in the left and right visual hemifields. Psychological Science 2005;16:637–643. [PubMed: 16102067]

Boussaoud D, Desimone R, Ungerleider LG. Visual topography of area TEO in the macaque. Journal of Comparative Neurology 1991;306:554–575. [PubMed: 1712794]

Chelazzi L, Duncan J, Miller EK, Desimone R. Responses of neurons in inferior temporal cortex during memory-guided visual search. Journal of Neurophysiology 1998;80:2918–2940. [PubMed: 9862896]

Delvenne JF. The capacity of visual short-term memory within and between hemifields. Cognition 2005;96:B79–B88. [PubMed: 15996557]

Desimone R, Gross CG. Visual areas in the temporal cortex of the macaque. Brain Research 1979;178:363–380. [PubMed: 116712]

DiCarlo JJ, Maunsell JH. Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. Journal of Neurophysiology 2003;89:3265–3278.

Duncan J, Bundesen C, Olson A, Humphreys G, Chavda S, Shibuya H. Systematic analysis of deficits in visual attention. Journal of Experimental Psychology: General 1999;128:450–478. [PubMed: 10650583]

Hasselmo ME, Rolls ET, Baylis GC. The role of expression and identity in the faceselective responses of neurons in the temporal visual cortex of the monkey. Behavioral Brain Research 1989;32:203–218.

Kastner S, De Weerd P, Pinsk MA, Elizondo MI, Desimone R, Ungerleider LG. Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. Journal of Neurophysiology 2001;86:1398–1411. [PubMed: 11535686]

Kim H, Somerville LH, Johnstone T, Polis S, Alexander AL, Shin LM, Whalen PJ. Contextual modulation of amygdala responsivity to surprised faces. Journal of Cognitive Neuroscience 2004;16:1730–1745. [PubMed: 15701225]

Kobatake E, Wang G, Tanaka K. Effects of shape-discrimination training on the sensitivity of inferotemporal cells in adult monkeys. Journal of Neurophysiology 1998;80:324–330. [PubMed: 9658053]

Logothetis NK, Pauls J, Poggio T. Shape representation in the inferior temporal cortex of monkeys. Current Biology 1995;5:552–563. [PubMed: 7583105]

Luck SJ, Hillyard SA, Mangun GR, Gazzaniga MS. Independent hemispheric attentional systems mediate visual search in split-brain patients. Nature 1989;342:543–545. [PubMed: 2586625]

Lundqvist, D.; Flykt, A.; Öhman, A. The Karolinska directed emotional faces. Psychology Section, Department of Clinical Neuroscience, Karolinska Institute; Stockholm: 1998.

Miller EK, Gochin PM, Gross CG. Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. Brain Research 1993;616:25–29. [PubMed: 8358617]

Niemeier M, Goltz HC, Kuchinad A, Tweed DB, Vilis T. A contralateral preference in the lateral occipital area: Sensory and attentional mechanisms. Cerebral Cortex 2005;15:325–331. [PubMed: 15269109]

Op De Beeck H, Vogels R. Spatial sensitivity of macaque inferior temporal neurons. Journal of Comparative Neurology 2000;426:505–518. [PubMed: 11027395]

Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. Nature Neuroscience 2001;4:739–744.

Reynolds JH, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4. Journal of Neuroscience 1999;19:1736–1753. [PubMed: 10024360]

Rolls ET, Tovee MJ. The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. Experimental Brain Research 1995;103:409–420.

Rovamo J, Virsu V. An estimation and application of the human cortical magnification factor. Experimental Brain Research 1979;37:495–510.

Sato T. Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. Experimental Brain Research 1989;77:23–30.

Sigala N, Logothetis NK. Visual categorization shapes feature selectivity in the primate temporal cortex. Nature 2002;415:318–320. [PubMed: 11797008]

Streit M, Ioannides AA, Liu L, Wölwer W, Dammers J, Gross J, et al. Neurophysiological correlates of the recognition of facial expressions of emotion as revealed by magnetoencephalography. Cognitive Brain Research 1999;7:481–491. [PubMed: 10076093]

Suzuki, S. High-level pattern coding revealed by brief shape aftereffects. In: Clifford, C.; Rhodes, G., editors. Fitting the mind to the world: Adaptation and aftereffects in high-level vision (Advances in Visual Cognition Series. Vol. 2. Oxford University Press; New York, New York: 2005. p. 135-172.

Suzuki S, Cavanagh P. Focused attention distorts visual space: An attentional repulsion effect. Journal of Experimental Psychology: Human Perception and Performance 1997;23:443–463. [PubMed: 9104004]

Tanaka K. Inferotemporal cortex and object vision. Annual Review of Neuroscience 1996;19:109–139.

Tyler CW, Chen CC. Spatial summation of face information. Journal of Vision 2006;6(10):11, 1117–1125.http://journalofvision.org/6/10/11/, doi:10.1167/6.10.11

Yarbus, AL. Eye movements and vision. Plenum Press; New York: 1967.

Young MP, Yamane S. Sparse population coding of faces in the inferotemporal cortex. Science 1992;256:1327–1331. [PubMed: 1598577]

Zoccolan D, Cox DD, DiCarlo JJ. Multiple object response normalization in monkey inferotemporal cortex. Journal of Neuroscience 2005;25:8150–8164. [PubMed: 16148223]
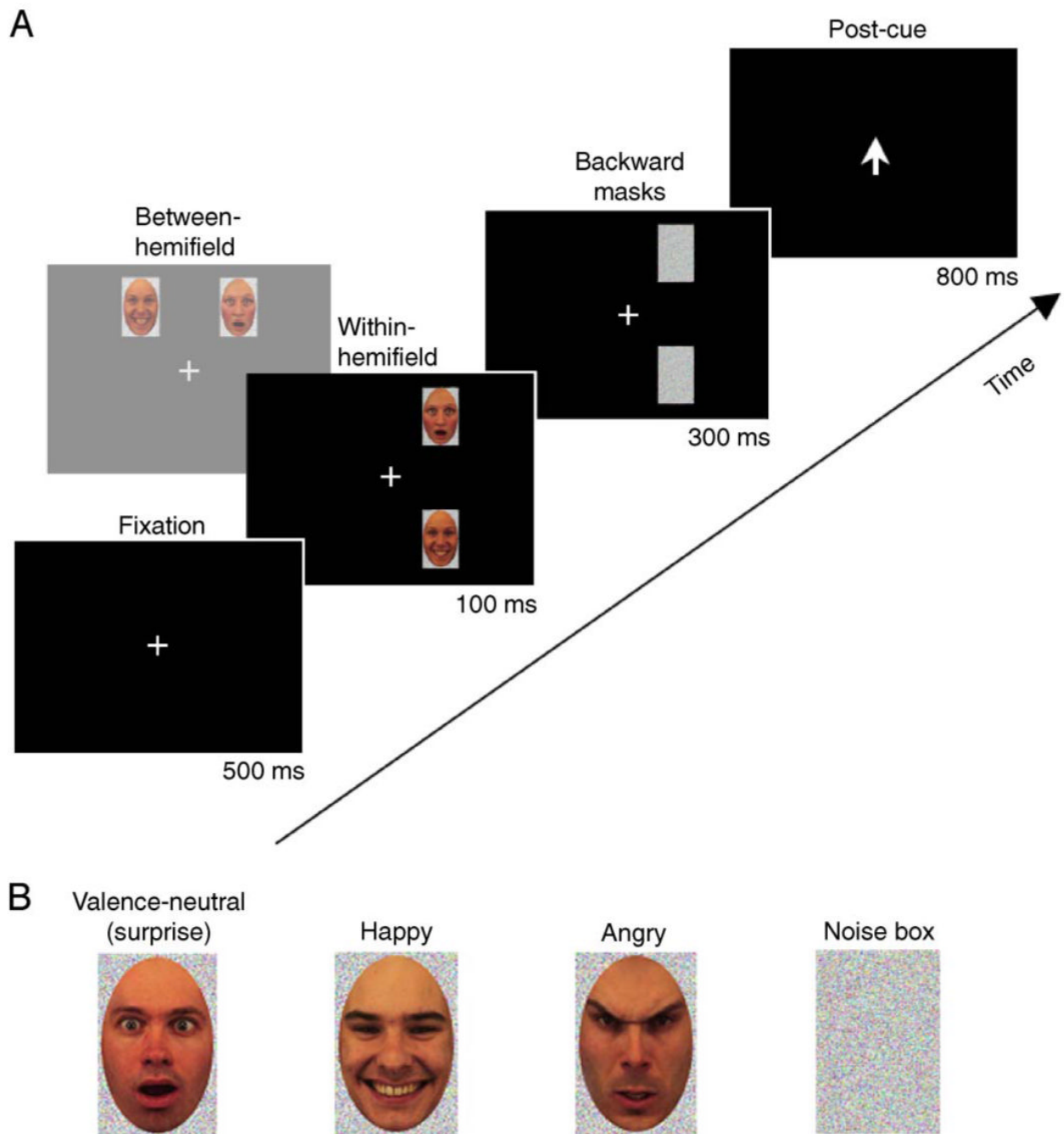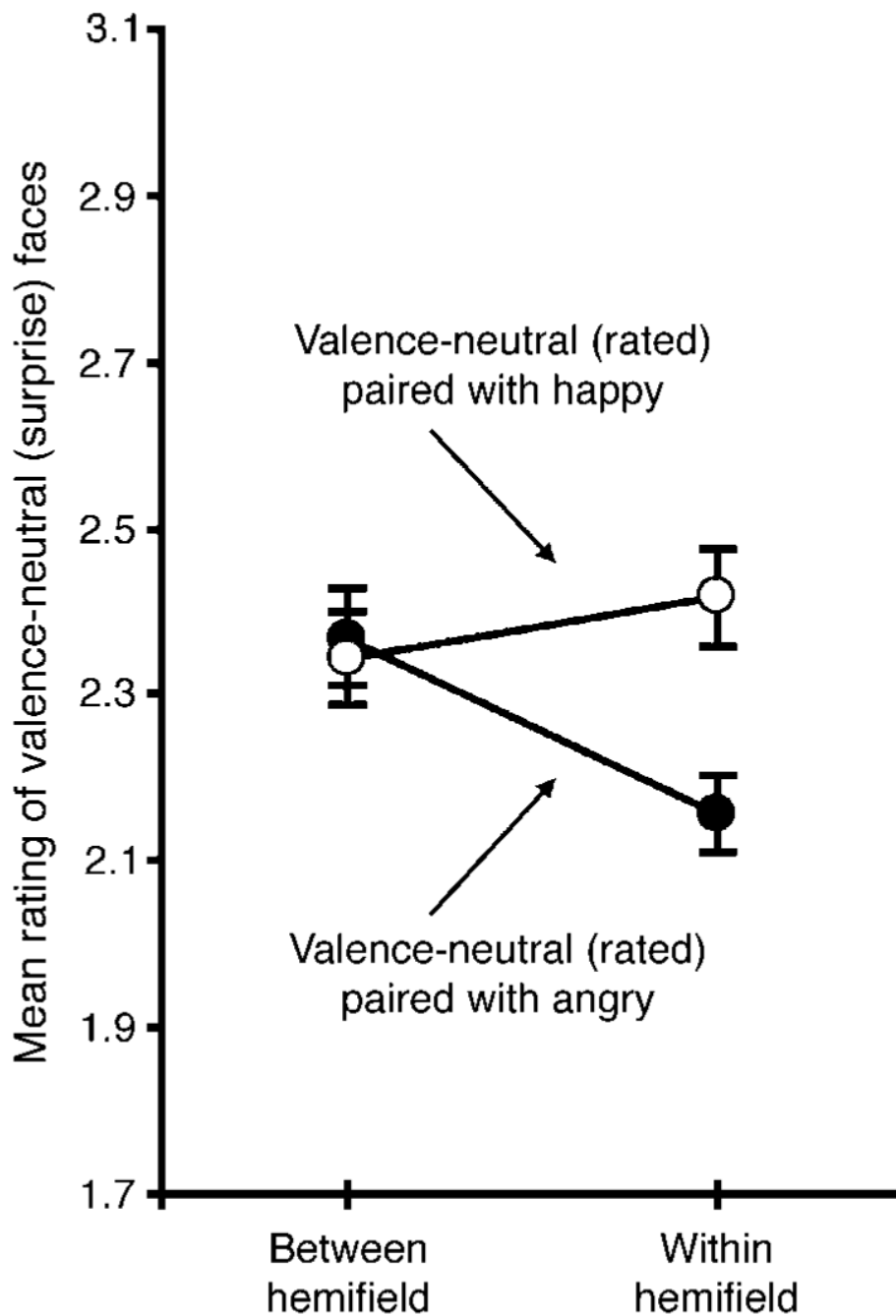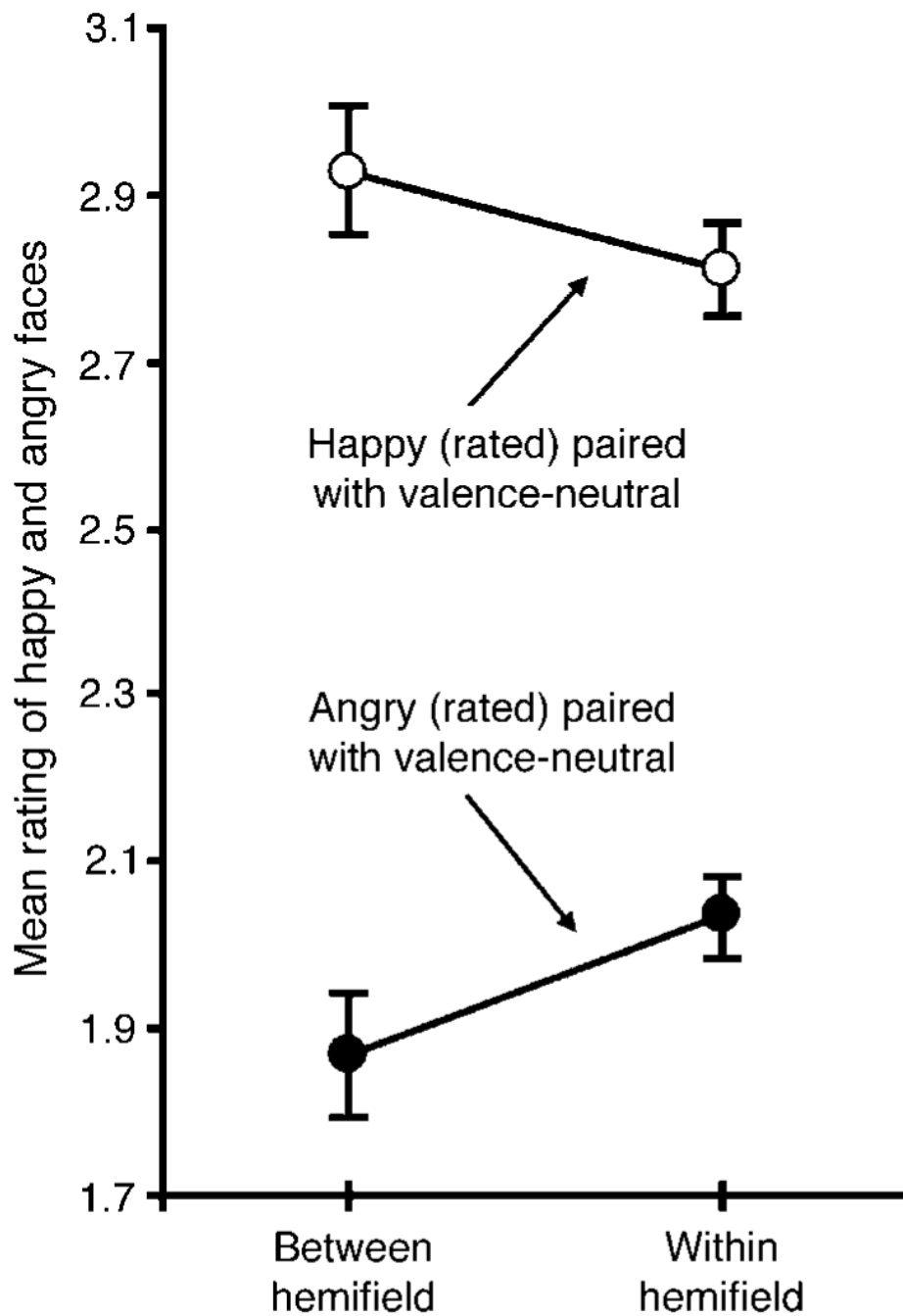
Figure 1.
(A) A typical trial sequence. Each trial in Experiments 1, 2, and 4 contained two faces, one expressing either happiness (positive valence) or anger (negative valence), and the other expressing surprise (neutral valence). Each trial in Experiment 3 contained one happy face and one angry face. In the within-hemifield condition, faces were presented either within the left or right visual hemifield. In the between-hemifield condition, faces were presented in opposite visual hemifields either above or below fixation. The within- and between-hemifield conditions were blocked in Experiments 1-3 and were randomly intermixed in Experiment 4. The post-face cue, a central arrow, indicated which face the observer should rate for its emotional

valence. (B) Examples of faces with valence-neutral (surprise), happy, and angry expressions, and a noise box that served as a backward mask.
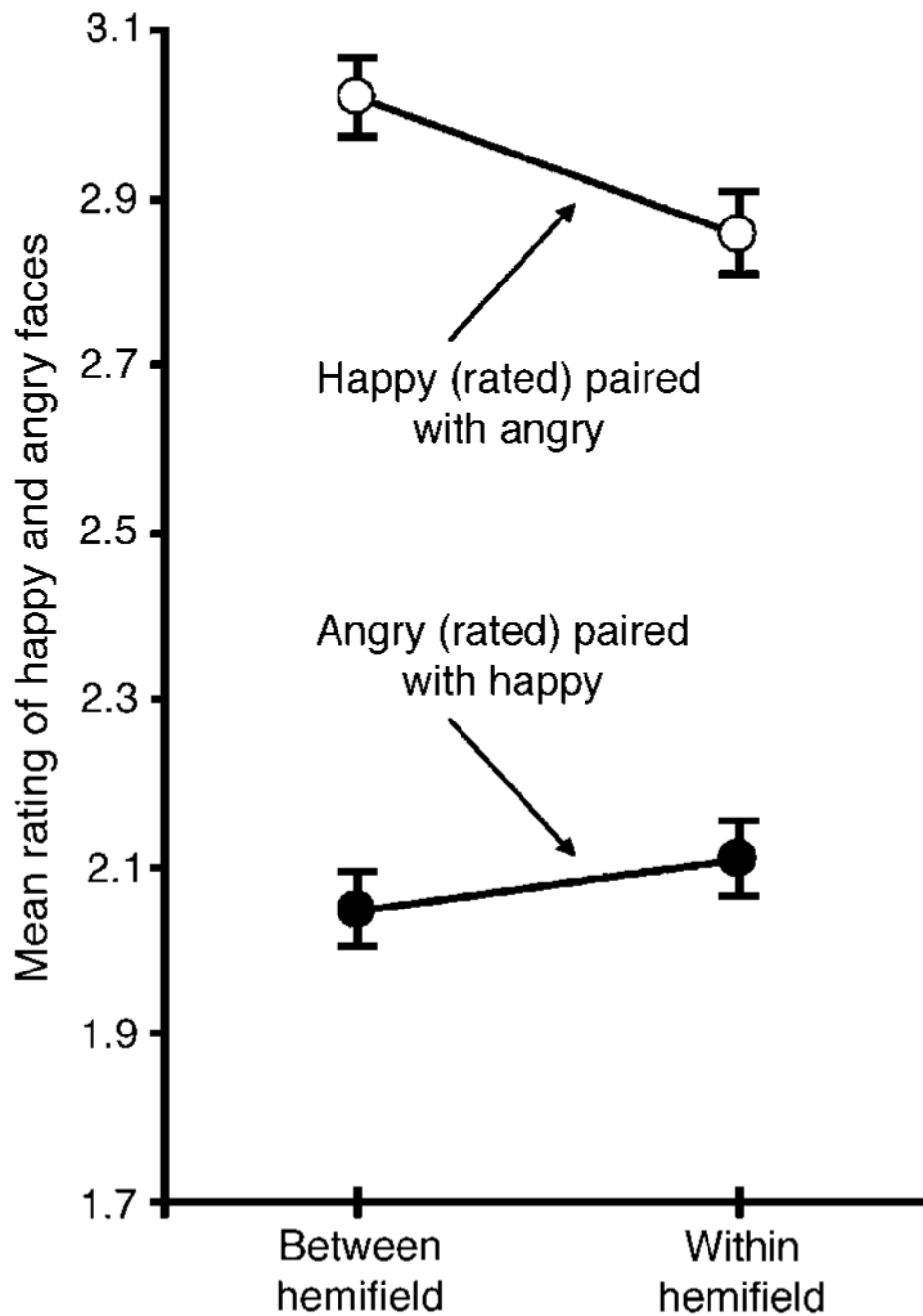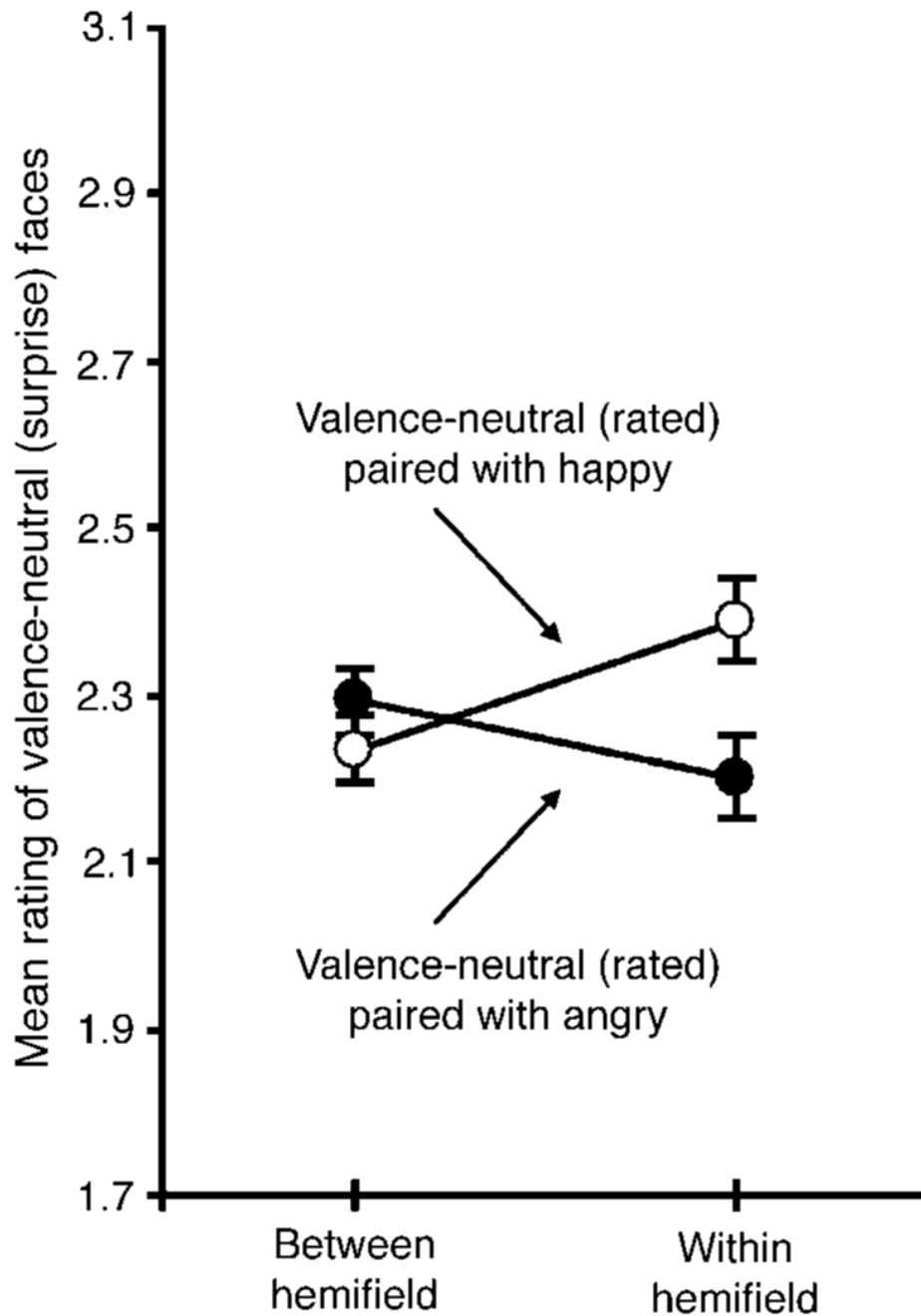
Figure 2.
Emotion ratings (ranging from most negative, 1, to most positive, 4) for a valence-neutral (surprise) face when an accompanying happy or angry face was presented in the opposite visual hemifield (between-hemifield condition) or in the same visual hemifield (within-hemifield condition). The error bars represent ±1 *SEM* with the baseline individual variability removed (i.e., repeated-measures error bars).

Figure 3.
Emotion ratings (ranging from most negative, 1, to most positive, 4) for a happy or angry face when an accompanying valence-neutral face was presented in the opposite visual hemifield (between-hemifield condition) or in the same visual hemifield (within-hemifield condition). The error bars represent ±1 *SEM* with the baseline individual variability removed (i.e., repeated-measures error bars).

Figure 4.
Emotion ratings (ranging from most negative, 1, to most positive, 4) for a happy or angry face when an accompanying face of opposite valence was presented in the opposite visual hemifield (between-hemifield condition) or in the same visual hemifield (within-hemifield condition). The error bars represent ±1 *SEM* with the baseline individual variability removed (i.e., repeated-measures error bars).

Figure 5.
Emotion ratings (ranging from most negative, 1, to most positive, 4) for a valence-neutral (surprise) face when an accompanying happy or angry face was presented in the opposite visual hemifield (between-hemifield condition) or in the same visual hemifield (within-hemifield condition). The error bars represent ±1 *SEM* with the baseline individual variability removed (i.e., repeated-measures error bars). Unlike in Experiment 1, the between- and within-hemifield conditions were randomly intermixed rather than blocked, and the face locations were randomly jittered rather than aligned.