The Primacy of Negative Interpretations When Resolving the Valence of Ambiguous Facial Expressions

Maital Neta and **Paul J. Whalen**
Dartmouth College

Abstract

Low-spatial-frequency (LSF) visual information is processed in an elemental fashion before a finer analysis of high-spatial-frequency information. Further, the amygdala is particularly responsive to LSF information contained within negative (e.g., fearful) facial expressions. In a separate line of research, it has been shown that surprised facial expressions are ambiguous in that they can be interpreted as either negatively or positively valenced. More negative interpretations of surprise are associated with increased ventral amygdala activity. In this report, we show that LSF presentations of surprised expressions bias the interpretation of surprised expressions in a negative direction, a finding suggesting that negative interpretations are first and fast during the resolution of ambiguous valence. We also examined the influence of subjects' positivity-negativity bias on this effect.

Keywords

low spatial frequencies; ambiguity; facial expressions; amygdala

Priority processing of visual information has been demonstrated by filtering visual images into different spatial-frequency bands (Carretie, Hinojosa, Lopez-Martin, & Tapia, 2007; Vuilleumier, Armony, Driver, & Dolan, 2003; Winston, Vuilleumier, & Dolan, 2003). For example, low spatial frequencies (LSFs) are consistently found to be processed first and fast (Bar et al., 2006; Hughes, Nozawa, & Kitterle, 1996). Moreover, visual object-recognition studies integrating functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) have shown that a coarse version of a visual stimulus, comprising mainly LSF information, is rapidly projected from early visual processing regions to the orbitofrontal cortex (OFC) before activity occurs in object-processing regions in the inferior temporal cortex (IT; Bar et al., 2006; Kveraga, Boshyan, & Bar, 2007). The LSF image is proposed to activate "initial guesses" as to what object might have given rise to such visual input (Bar et al., 2006). Emotional stimuli can also elicit priority processing, evoking early and strong responses in the amygdala and OFC (Vuilleumier & Pourtois, 2007), as well as in subcortical structures associated with gating attention and eye movements, such as the pulvinar and the superior colliculus (Liddell et al., 2005; Morris et al., 1998).

Facial expressions provide information about the emotions and intentions of other people. Fearful expressions activate both OFC and the amygdala (Eimer & Holmes, 2002;

Corresponding Author: Maital Neta, Department of Psychological & Brain Sciences, Dartmouth College, 6207 Moore Hall, Hanover, NH 03755, maital.neta@dartmouth.edu.

Vuilleumier, Henson, Driver, & Dolan, 2002), and these regions are especially sensitive to LSF versions of fearful faces (Vuilleumier, Armony, Driver, & Dolan, 2001; Winston et al., 2003). Both regions also have connections to motor structures such as the basal ganglia and the superior colliculus, which may subserve the initial, coarse evaluation of input and the triggering of reflexive or learned responses (Kveraga, Ghuman, & Bar, 2007). These results suggest that LSF information within emotional stimuli may subserve an analogous initial-guess process as part of a rapid threat-detection or "early warning" system (de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999; LeDoux, 1996; Morris, Öhman, & Dolan, 1999; Rafal, Smith, Krantz, Cohen, & Brennan, 1990; Sahraie, Weiskrantz, Trevethan, Cruce, & Murray, 2002).

Although some facial expressions, such as fearful expressions, are clear in terms of the valence of the outcomes that they predict, other expressions send a more ambiguous message. For example, when surprised expressions are presented within an experimental context that provides no information to disambiguate their valence, they are interpreted negatively by some people and positively by others (Kim, Somerville, Johnstone, Alexander, & Whalen, 2003; Kim et al., 2004). Thus, surprised expressions offer a means to assess positivity-negativity bias (Neta, Norris, & Whalen, 2009). Previous neuroimaging research has demonstrated that more negative interpretations of surprised expressions are correlated with increased activation of the amygdala (Kim et al., 2003). Subjects who interpreted surprised expressions positively showed lower amygdala activity and also (inversely correlated) higher ventral medial prefrontal cortex (vmPFC) activity, which suggests that a regulatory process is involved.

We speculated that ascribing negative valence to the expression of surprise is a default interpretation (i.e., the *initial-negativity hypothesis*), whereas a more positive interpretation of the expression requires a regulatory override of this default response. If LSF information is processed more rapidly than high-spatial-frequency (HSF) information by the systems that form the neural substrate of this initial-negativity hypothesis (e.g., amygdala), then LSF presentations of surprised expressions will be interpreted more negatively than HSF presentations. Across two experiments, we aimed to determine whether this effect occurs by comparing ratings of surprised expressions with ratings of fearful expressions (which have a similar morphological structure but are more consistently rated as negative; Experiment 1), as well as with more clearly valenced expressions of happiness and anger, which served as positive and negative anchors (Experiment 2).

Experiment 1

Method

Subjects—Thirty-two healthy Caucasian Dartmouth undergraduates (23 female and 9 male; ages 18–23 years, mean age = 19.8 years) volunteered to participate. All subjects had normal or corrected-to-normal vision, used no psychoactive medication, and reported no significant neurological or psychiatric history. None were aware of the purpose of the experiment, and they were all compensated for their participation through monetary payment or course credit. Written informed consent was obtained from each subject before the session, and all procedures were approved by the Dartmouth College Committee for the Protection of Human Subjects. Three subjects were excluded for having Beck Depression Inventory (BDI; Beck, Ward, & Mendelson, 1961) scores that were above our cutoff (13) for normal adults. As a result, the final sample contained 29 subjects (21 females and 8 males). All 29 subjects tested within normal limits for depression ($M = 4.72$, $SE = 0.56$) and anxiety (State-Trait Anxiety Inventory, STAI; Spielberger, Gorsuch, & Lushene, 1970; state STAI: $M = 35.6$, $SE = 1.97$; trait STAI: $M = 37.2$, $SE = 1.68$).

Stimuli—Stimuli consisted of equal numbers of male and female faces posing fearful and surprised expressions. These faces were derived from multiple standardized sets, including the NimStim set (22 individuals; Tottenham et al., 2009), the Pictures of Facial Affect (13 individuals; Ekman & Friesen, 1976), and the Averaged Karolinska Directed Emotional Faces (KDEF) database (31 individuals; Lundqvist, Flykt, & Öhman, 1998). Specifically, there were a total of 66 identities included in the stimuli. Some of these individuals posed both surprised and fearful expressions, and others posed only one of the expressions, for a total of 88 discrete stimuli. All images were transformed to gray-scale images (aside from the Pictures of Facial Affect, which were already in gray-scale) with a resolution of 75 dots per inch. The facial expressions in this stimulus set were validated by a separate set of subjects who labeled each expression; only faces that were correctly labeled by more than 60% of subjects were included.

Spatial-frequency content in the original image (broad spatial frequencies, BSFs) was filtered in order to create two versions of each face: one comprising primarily the HSF information and one comprising primarily the LSF information (see Fig. 1). We used a high-pass cutoff of 24 cycles per image for the HSF information and a low-pass cutoff of 6 cycles per image for the LSF information; these values are consistent with those used in previous work (e.g., Vuilleumier et al., 2003). Moreover, prior to filtering, we adjusted the contrast and luminance of each image in order to equate these elements across stimulus conditions and stimulus sets.

Procedure—Faces of the three frequency types (i.e., BSF, HSF, and LSF) were presented in pseudorandom order within each of four runs that consisted of 44 trials. Each face (a given identity posing a given expression) was presented twice to each subject, for a total of 176 trials. Face identities were counter-balanced, such that each subject viewed a given face as either filtered (the HSF and LSF versions in a counterbalanced order) or intact (two presentations of the BSF version). We avoided presenting the same identity in both BSF and filtered versions to a given subject so that the BSF versions would not affect ratings of the filtered images (see Vuilleumier et al., 2003). Faces were presented, one at a time, at the center of the screen (visual angle = 7°) on a black background. Each face was presented for 200 ms and was followed by an interstimulus interval during which a fixation cross appeared at the center of the screen. The interstimulus interval varied from 1,800 to 5,800 ms (M = 3,800 ms). Subjects were asked to report, as quickly and accurately as possible, their gut reaction as to whether each face had a positive or negative valence.

Results

A repeated measures analysis of variance (ANOVA) showed the predicted significant interaction of expression (surprised, fearful) and spatial frequency (HSF, BSF, LSF), $F(2, 27) = 13.58$, $p < .001$, $\eta^2 = .05$ (see Fig. 2a). Fisher's least significant difference (LSD) pair-wise comparisons revealed that surprised expressions were rated as more negative when presented in their LSF versions than when presented in their HSF versions, $p < .01$, one-tailed, or their BSF versions, $p = .001$. Surprised expressions were also rated as more negative when presented in their HSF versions than when presented in their BSF versions, $p < .05$. There was also a significant main effect of expression, $F(1, 28) = 84.26$, $p < .001$, $\eta^2 = .56$. Fisher's LSD pair-wise comparisons revealed that fearful expressions were rated as more negative than surprised expressions, $p < .001$. Moreover, fearful expressions were rated as more negative when presented in their BSF versions than when presented in their LSF versions, $p = .003$, or HSF versions, $p = .002$.

Experiment 2

Method

Subjects—Thirty-eight healthy Caucasian Dartmouth undergraduates (21 female and 17 male; ages 18–26 years, mean age = 20.4 years) volunteered to participate and met the same requirements as in Experiment 1. Four subjects were removed because they provided nonnormative ratings of facial expressions (e.g., they rated happy expressions as negative or angry expressions as positive on greater than 40% of trials). Another 3 subjects were removed for having BDI scores that were above our cutoff (13) for normal adults. As a result, the final sample contained 31 subjects (16 females and 15 males). All 31 subjects tested within normal limits for depression (BDI: $M = 6.42$, $SE = 0.85$) and anxiety (state STAI: $M = 37$, $SE = 1.55$; trait STAI: $M = 39.4$, $SE = 1.59$).

Stimuli and procedure—Experimental methods were the same as in Experiment 1, and stimuli were validated in the same way as in Experiment 1, except that only surprised, angry, and happy expressions were used. Surprised, angry, and happy expressions were taken from the same standardized sets as in Experiment 1: NimStim (8 individuals; Tottenham et al., 2009), Pictures of Facial Affect (13 individuals; Ekman & Friesen, 1976), and KDEF database (39 individuals; Lundqvist et al., 1998). Of the 60 individuals whose images were included in the experiment, some posed all three expressions, and some posed only one or two of the expressions. We used a total of 99 discrete stimuli, which were presented in three runs of 66 trials each. All stimuli were validated prior to use in this study.

Results

A repeated measures ANOVA showed the predicted significant interaction of expression (surprised, angry, happy) and spatial frequency (HSF, BSF, LSF), $F(4, 27) = 8.01$, $p < .001$, $\eta^2 = .004$ (see Fig. 2b). Fisher's LSD pair-wise comparisons revealed that LSF surprised expressions were rated as more negative than HSF surprised expressions, $p < .025$, one-tailed, and as more negative than BSF surprised expressions, $p = .026$, one-tailed. There was also a significant main effect of expression, $F(2, 29) = 1,019.67$, $p < .001$, $\eta^2 = .94$. Fisher's LSD pair-wise comparisons revealed that angry expressions were rated as more negative than happy expressions, $p < .001$, and as more negative than surprised expressions, $p < .001$. Also, happy expressions were rated as more positive than surprised expressions, $p < .001$. Moreover, BSF angry expressions were rated as more negative than LSF angry expressions, $p < .001$, and as more negative than HSF angry expressions, $p = .03$. In addition, BSF happy expressions were rated as more positive than LSF happy expressions, $p < .04$, and as more positive than HSF happy expressions, $p = .003$. Finally, happy expressions were rated as more positive in their LSF versions than in their HSF versions, $p < .04$.

Combined Results for Surprised Expressions: Positivity-Negativity Bias

In accordance with a study that was run in parallel to the experiments reported here (Neta et al., 2009), another goal of the present study was to examine whether these effects were modulated by an individual's positivity-negativity bias (i.e., tendency to interpret surprised expressions as having a negative, $Bias_{neg}$, or positive, $Bias_{pos}$, valence). As in our previous study, we used a median split (across the two experiments; $N = 60$) to divide subjects on the basis of the percentage of BSF surprised expressions that they rated as negative. Combining the data from the two experiments provided adequate power for this analysis.

First, a one-sample t test revealed that the BSF surprised expressions were rated as significantly more negative than chance ($M = 66.3\%$ negative ratings, $SE = 2.54$, range: 23.3%–100%), $t(59) = 6.42$, $p < .001$. A median split identified the 30 subjects showing the

greatest tendency to interpret these expressions as having a negative valence, (Bias$_{neg}$ group; 12 men, 18 women; $M = 81.8\%$ negative ratings, $SE = 1.87$, range: 67.9%–100%) and the other 30 subjects showing a lesser tendency to interpret these expressions as having a negative valence (Bias$_{pos}$ group; 11 men, 19 women; M 50.8% negative ratings, $SE = 2.50$, range: 23.3%–66.7%). A Spatial Frequency (HSF, BSF, LSF) × Group (Bias$_{neg}$, Bias$_{pos}$) repeated measures ANOVA showed a significant interaction, $F(2, 57) = 5.18$, $p = .006$, $\eta^2 = .08$ (see Fig. 3). Pair-wise comparisons (Fisher's LSD) revealed no difference in the percentage of negative ratings across the spatial-frequency conditions for the Bias$_{neg}$ group. However, the Bias$_{pos}$ group rated LSF surprised expressions as more negative than the HSF surprised expressions, $p = .003$, and the BSF surprised expressions, $p = .001$, and the HSF surprised expressions as more negative than the BSF surprised expressions, $p = .02$. Also, there was an overall significant main effect of spatial frequency, $F(2, 57) = 10.42$, $p < .001$, $\eta^2 = .14$. Pair-wise comparisons (Fisher's LSD) revealed that LSF surprised expressions were rated as more negative than HSF surprised expressions, $p = .002$, and BSF surprised expressions, $p < .001$, but that there was no difference in the percentage of negative ratings between the HSF and BSF surprised expressions.

Discussion

This study showed that people are more likely to interpret a coarse, elemental presentation of an ambiguous surprised facial expression (i.e., LSF information) as negatively valenced than a fine presentation of the same expression (i.e., HSF information). These data are consistent with the notion that a negative interpretation of surprised expressions is more fundamental and more heavily linked to LSF information than a positive interpretation of surprised expressions, whereas a positive interpretation of these expressions is more heavily dependent on HSF information. In contrast, expressions with clear valence were rated more consistently as having that valence (i.e., angry and fearful expressions were rated as negative, and happy expressions were rated as positive) when presented intact (BSF images) than when filtered to include only LSF or HSF information. Finally, subjects' positivity-negativity bias significantly influenced these ratings. Specifically, subjects who tended to rate BSF surprised expressions as negative rated HSF and LSF versions similarly (both negative), whereas subjects who tended to rate BSF surprised expressions as positive showed the expected shift toward negativity in response to LSF versions.

In a previous fMRI study, we showed that subjects who interpreted surprised expressions as negative showed increased amygdala activity that correlated with attenuated vmPFC responses (Kim et al., 2003). Conversely, subjects who interpreted these expressions as positive showed the opposite response pattern (i.e., increased vmPFC and decreased amygdala activity). On the basis of these results, we suggested that the prefrontal cortex is necessary to resolve the dual-valence representation associated with surprised expressions (i.e., the fact that they have predicted both positive and negative outcomes), and that individuals who fail to recruit this region maintain their initial negativity judgment. The present results showing that LSF versions of surprised expressions were rated as more negative than HSF (or BSF) versions provide support for this assertion. Further, across all spatial-frequency conditions, ratings of surprised expressions were more positive than ratings of fearful expressions, even for subjects demonstrating a negativity bias for surprise. These data are consistent with the proposed dual-valence representation associated with surprised expressions, which makes them fundamentally distinct from fear (but see Katsikitis, 1997).

One alternative explanation of the present data is that blurry LSF representations of faces are generally interpreted as more negative than intact facial images, regardless of the specific facial expression. Indeed, previous research has shown that LSF versions of neutral faces are

rated as less trustworthy than HSF and BSF versions of the same faces (Said, Baron, & Todorov, 2008). However, we did not observe a similar effect for fear and anger. In addition, LSF versions of happy expressions were rated as more positive than HSF versions of happy expressions, which suggests that LSF information may be linked to the more basal valence of the expression.

In summary, our a priori prediction was that, among facial expressions, the primacy effect associated with LSF information is contingent upon the ambiguity of valence associated with surprise. This primacy effect was observed across all subjects and was most pronounced in those showing a positivity bias in their ratings of surprise. In future studies, we will investigate whether these findings correlate with prefrontal-amygdala interaction during the processing of surprised expressions to determine if positivity-negativity bias is related to the degree to which subjects recruit their vmPFC.

Acknowledgments

References

Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, et al. Top-down facilitation of visual recognition. Proceedings of the National Academy of Sciences, USA. 2006; 103:449–454.

Beck AT, Ward C, Mendelson M. Beck Depression Inventory (BDI). Archives of General Psychiatry. 1961; 4:561–571. [PubMed: 13688369]

Carretie L, Hinojosa JA, Lopez-Martin S, Tapia M. An electrophysiological study on the interaction between emotional and spatial frequency of visual stimuli. Neuropsychology. 2007; 45:1187–1195.

de Gelder B, Vroomen J, Pourtois G, Weiskrantz L. Non-conscious recognition of affect in the absence of striate cortex. NeuroReport. 1999; 10:3759–3763. [PubMed: 10716205]

Eimer M, Holmes A. An ERP study on the time course of emotional face processing. NeuroReport. 2002; 13:427–431. [PubMed: 11930154]

Ekman, P.; Friesen, WV. Pictures of Facial Affect [Slides]. Palo Alto, CA: Consulting Psychologists Press; 1976.

Hughes HC, Nozawa G, Kitterle F. Global precedence, spatial frequency channels, and the statistics of natural images. Journal of Cognitive Neuroscience. 1996; 8:197–230.

Katsikitis M. The classification of facial expressions of emotion: A multi-dimensional scaling approach. Perception. 1997; 26:613–626. [PubMed: 9488885]

Kim H, Somerville LH, Johnstone T, Alexander A, Whalen PJ. Inverse amygdala and medial prefrontal cortex responses to surprised faces. NeuroReport. 2003; 14:2317–2322. [PubMed: 14663183]

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]

Kveraga K, Boshyan J, Bar M. Magnocellular projections as the trigger of top-down facilitation in recognition. Journal of Neuroscience. 2007; 27:13232–13240. [PubMed: 18045917]

Kveraga K, Ghuman AS, Bar M. Top-down predictions in the cognitive brain. Brain and Cognition. 2007; 65:145–168. [PubMed: 17923222]

LeDoux, JE. The emotional brain. New York: Simon & Schuster; 1996.

Liddell BJ, Brown KJ, Kemp AH, Barton MJ, Das P, Peduto A, et al. A direct brainstem-amygdala-cortical 'alarm' system for subliminal signals of fear. NeuroImage. 2005; 24:235–243. [PubMed: 15588615]

Lundqvist, D.; Flykt, A.; Öhman, A. The Karolinska Directed Emotional Faces—KDEF [CD ROM]. Sweden: Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet, Solna; 1998. Available from

Morris JS, Friston KJ, Buchel C, Frith CD, Young AW, Calder AJ, Dolan RJ. A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain. 1998; 121(1):47–57. [PubMed: 9549487]

Morris JS, Öhman A, Dolan RJ. A subcortical pathway to the right amygdala mediating "unseen" fear. Proceedings of the National Academy of Sciences, USA. 1999; 96:1680–1685.

Neta M, Norris CJ, Whalen PJ. Corrugator muscle responses are associated with individual differences in positivity-negativity bias. Emotion. 2009; 9:640–648. [PubMed: 19803586]

Rafal R, Smith J, Krantz J, Cohen A, Brennan C. Extrageniculate vision in hemianopic humans: Saccade inhibition by signals in the blind field. Science. 1990; 250:118–121. [PubMed: 2218503]

Sahraie A, Weiskrantz L, Trevethan CT, Cruce R, Murray AD. Psychophysical and pupillometric study of spatial channels of visual processing in blindsight. Experimental Brain Research. 2002; 143:249–259.

Said CP, Baron SG, Todorov A. Nonlinear amygdala response to face trustworthiness: Contributions of high and low spatial frequency information. Journal of Cognitive Neuroscience. 2008; 21:519–528. [PubMed: 18564045]

Spielberger, CD.; Gorsuch, RL.; Lushene, RE. STAI: Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1970.

Tottenham N, Tanaka J, Leon AC, McCarry T, Nurse M, Hare TA, et al. The NimStim set of facial expressions: Judgments from untrained research participants. Psychiatry Research. 2009; 168:242–249. [PubMed: 19564050]

Vuilleumier P, Armony JL, Driver J, Dolan RJ. Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. Neuron. 2001; 30:829–841. [PubMed: 11430815]

Vuilleumier P, Armony JL, Driver J, Dolan RJ. Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nature Neuroscience. 2003; 6:624–631.

Vuilleumier P, Henson RN, Driver J, Dolan RJ. Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nature Neuroscience. 2002; 5:491–499.

Vuilleumier P, Pourtois G. Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. Neuropsychologia. 2007; 45:174–194. [PubMed: 16854439]

Winston JS, Vuilleumier P, Dolan RJ. Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. Current Biology. 2003; 13:1824–1829. [PubMed: 14561410]
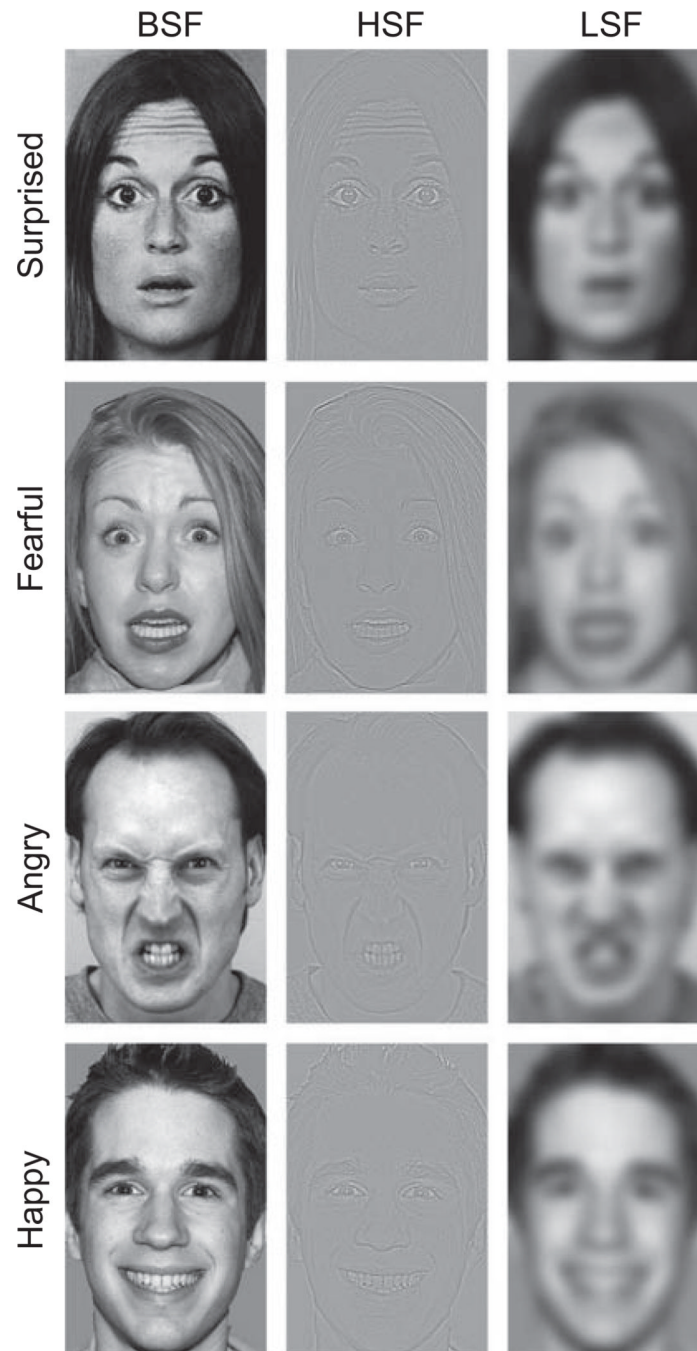
Fig. 1.
Example stimuli. Faces with a normal, broad-spatial-frequency (BSF) content were filtered to contain only high-spatial-frequency (HSF) or low-spatial-frequency (LSF) information. Fearful and surprised expressions (Experiment 1) and angry, happy, and surprised expressions (Experiment 2) were presented in BSF, HSF, or LSF form for 200 ms and followed by a fixation cross that appeared for a variable interval (range from 1,800 to 5,800 ms). The task was to decide whether each expression was positive or negative.
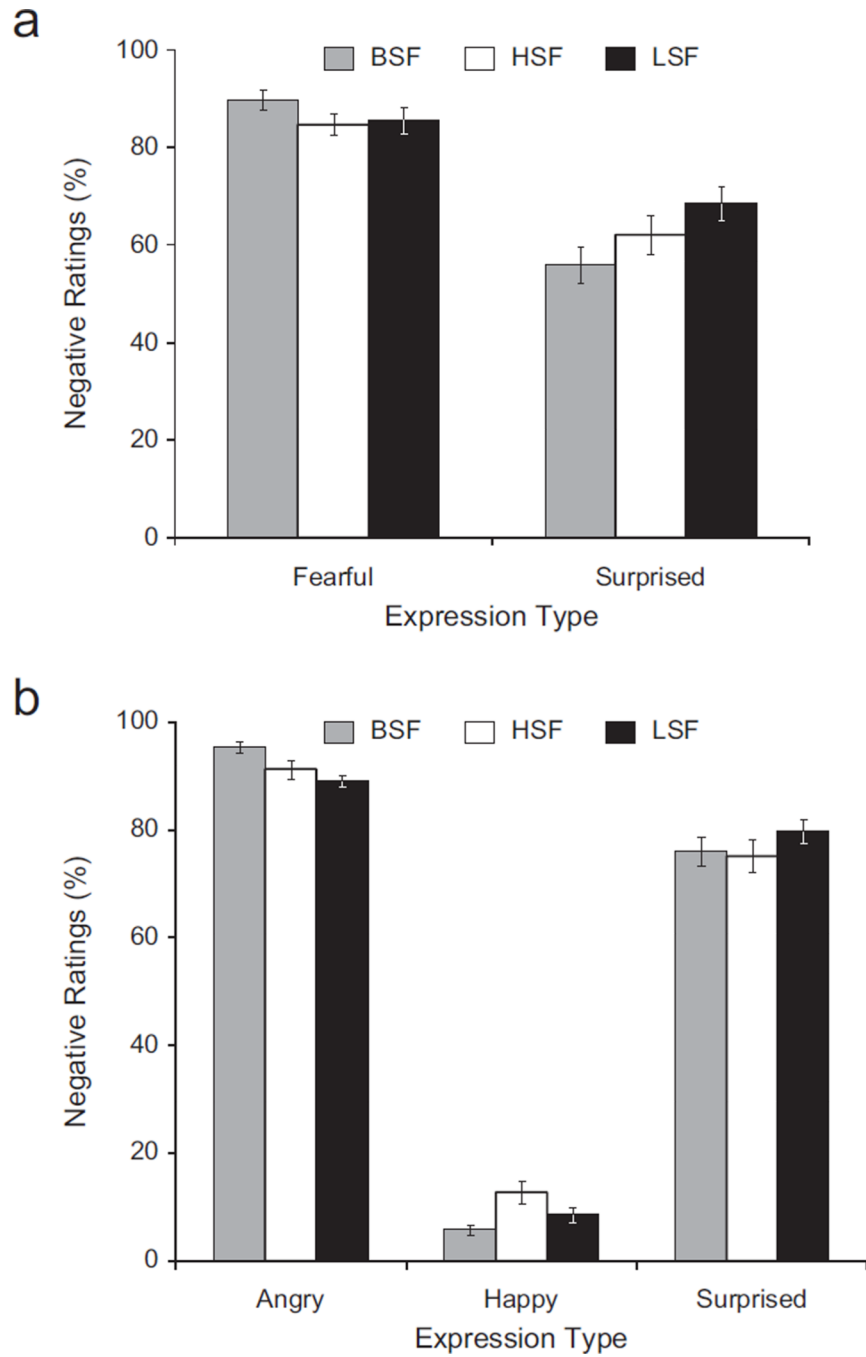
Fig. 2.
Percentage of negative ratings of the facial expressions in (a) Experiment 1 and (b) Experiment 2 as a function of expression type and spatial-frequency information. Fearful and surprised expressions were rated in Experiment 1, and surprised, angry, and happy faces were rated in Experiment 2. Error bars represent standard errors of the mean. BSF = broad spatial frequency (i.e., intact image); HSF = high spatial frequency; LSF = low spatial frequency.
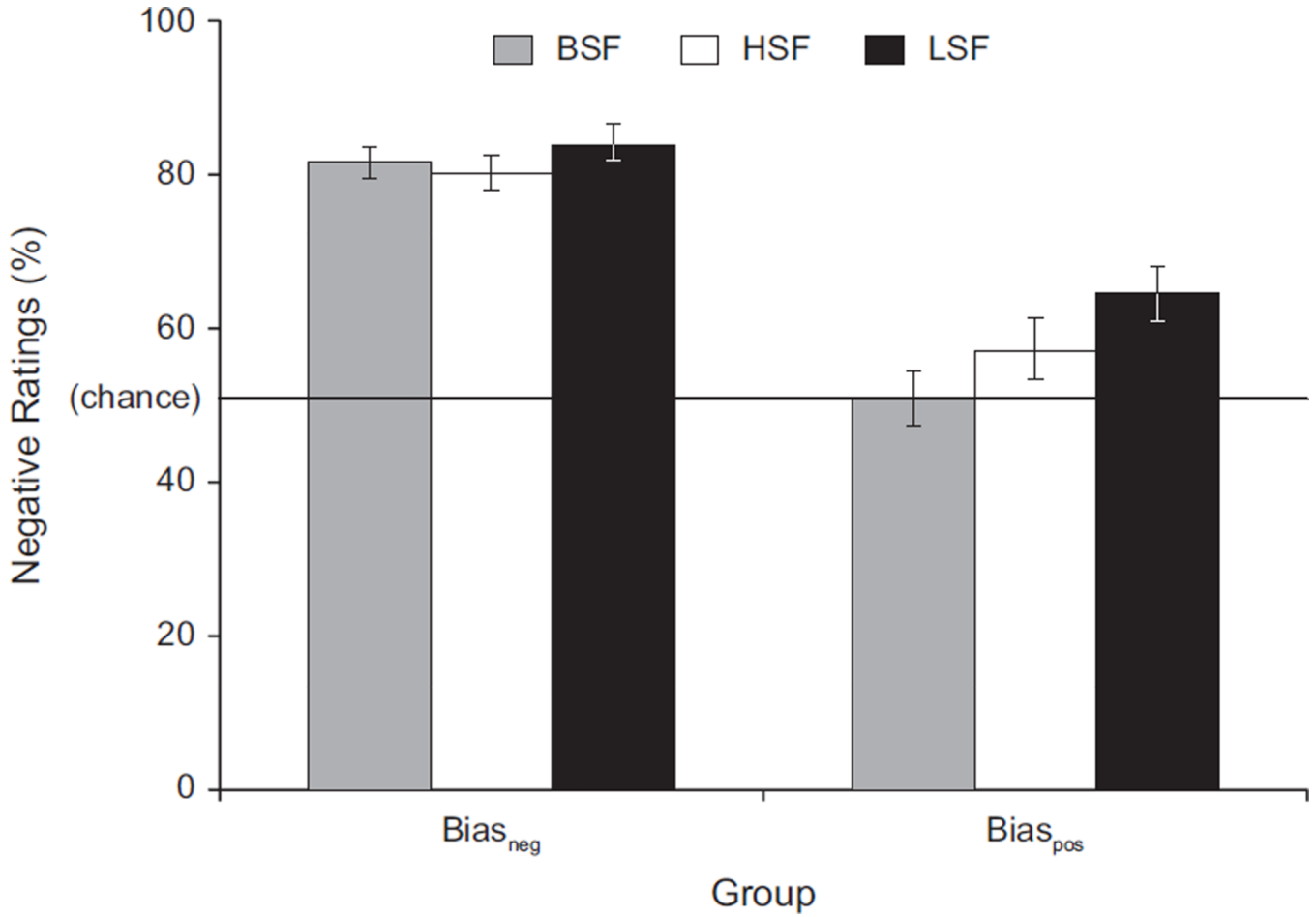
Fig. 3.
Percentage of negative ratings of surprised expressions as a function of spatial frequency for the individuals who tended to interpret surprise as negative (Bias$_{neg}$ group) and those who tended to interpret surprise as positive (Bias$_{pos}$ group). Error bars represent standard errors of the mean. BSF = broad spatial frequency (i.e., intact image); HSF = high spatial frequency; LSF = low spatial frequency.