Emotional Intensity Modulates the Integration of Bimodal Angry Expressions: ERP Evidence

Zhihui Pan[1], Xi Liu[2], Yangmei Luo[1] and Xuhai Chen[1]*

[1] *Key Laboratory of Behavior and Cognitive Psychology in Shaanxi Province, School of Psychology, Shaanxi Normal University, Xi'an, China,* [2] *State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, School of Brain Cognitive Science, Beijing Normal University, Beijing, China*

Integration of information from face and voice plays a central role in social interactions. The present study investigated the modulation of emotional intensity on the integration of facial-vocal emotional cues by recording EEG for participants while they were performing emotion identification task on facial, vocal, and bimodal angry expressions varying in emotional intensity. Behavioral results showed the rates of anger and reaction speed increased as emotional intensity across modalities. Critically, the P2 amplitudes were larger for bimodal expressions than for the sum of facial and vocal expressions for low emotional intensity stimuli, but not for middle and high emotional intensity stimuli. These findings suggested that emotional intensity modulates the integration of facial-vocal angry expressions, following the principle of Inverse Effectiveness (IE) in multimodal sensory integration.

Keywords: facial, vocal, anger, emotional intensity, integration, ERP

INTRODUCTION

Successful social interaction requires a precise understanding of the feelings, intentions, thoughts, and desires of other people (Sabbagh et al., 2004). Human beings have to integrate sensory cues came from facial expressions and vocal intonation to create a coherent, unified perception (Klasen et al., 2012), given that the evaluation of emotion is rarely based on the expression of one modality alone. Although, the phenomenon of bimodal emotion integration has been clearly depicted (Klasen et al., 2012; Chen et al., 2016a,b), the influence of emotional features thereof, such as emotional intensity, has seldom been tested. This is of importance as emotional intensity has been proved to be a key factor influencing emotional perception (Sprengelmeyer and Jentzsch, 2006; Yuan et al., 2007; Dunning et al., 2010; Wang et al., 2015) and the modulation of emotional intensity is one important way to test the principle of Inverse Effectiveness (IE), that is, multisensory integration is more effective when its constituent modality is less salient (Collignon et al., 2008; Stein and Stanford, 2008; Stein et al., 2009; Jessen et al., 2012). Therefore, the current study aims to test modulation of emotional intensity on bimodal anger integration using the technic of electroencephalography (EEG) with high temporal resolution.

The interaction of bimodal emotional cues, particularly the emotional signals delivered by facial and vocal expressions, has been widely reported (Klasen et al., 2012). One line of studies showed that bimodal emotions led to shorter response times and higher response accuracy, demonstrating a facilitation effect of bimodal emotional cues (Dolan et al., 2001; Klasen et al., 2012; Schelenz et al., 2013). Additionally, it was observed that bimodal emotional cues reduced the amplitude of the early

auditory N1 compared to unimodal cues (Jessen and Kotz, 2011; Kokinous et al., 2015), while the redundant emotional information in the auditory domain reduced the amplitude of the P2 and N3 components compared to faces alone (Paulmann et al., 2009). These attenuated components were considered the neural substrate of facilitated perceptual processing (Klasen et al., 2012). Some other studies compared the brain responses for congruent bimodal emotions with those for incongruent bimodal emotions, and found that congruent bimodal emotions enhanced the N1 (De Gelder et al., 2002) and P2 (Balconi and Carrera, 2011) amplitudes relative to incongruent ones, suggesting an integration of bimodal emotional cues indirectly. Conversely, some other studies tested the bimodal integration directly, by comparing the bimodal responses with the summed unimodal responses, which results in either sub- or supra-additive response (Stein and Stanford, 2008; Hagan et al., 2009). For instance, Hagan et al. (2009) observed a significant super-additive gamma oscillation within the first 250 ms for congruent audiovisual emotional perception. Likewise, Jessen and Kotz (2011) found stronger beta suppression for multimodal than for the summed unimodal conditions. Moreover, recent evidence shows a robust superadditivity in P3 amplitudes (Chen et al., 2016a,b) and theta band power (Chen et al., 2016b) during bimodal emotional change perception.

Taken together, the interaction and/or integration of facial-vocal emotional cues have been well-examined. However, some studies argued that the principle of Inverse Effectiveness is the most valid way to test the multimodal integration (Stein and Stanford, 2008; Stein et al., 2009). Concerning the integration of multimodal emotions, it was reported that bimodal emotional cues affected each other, especially when the target modality was less reliable (Collignon et al., 2008) and multimodal emotions elicited earlier N1 than unimodal emotions under high noisy background but not low noisy background (Jessen et al., 2012). Nevertheless, all these studies tested IE principle by manipulating the noisy level of the background. To precisely delineate the integration of bimodal emotion integration, it is necessary to manipulating the saliency of the emotional stimulus itself. One of the effective ways to manipulating emotional saliency is changing emotional intensity, which has been proved to be a key factor influencing emotional perception (Sprengelmeyer and Jentzsch, 2006; Yuan et al., 2007; Dunning et al., 2010; Wang et al., 2015). For instance, Dunning et al. (2010) reported that as the percentages of anger in faces increased from 0 to 100%, faces were perceived as increasingly angry. Moreover, it was reported that the N170 amplitudes increased as the augment of the emotional intensity (Sprengelmeyer and Jentzsch, 2006; Wang et al., 2013). Likewise, it was found that the rates of happiness and P300 amplitudes increased as emotional intensity increased in vocal emotion decoding (Wang et al., 2015).

Therefore, the present study examined the integration of facial and vocal emotion cues by manipulating emotion intensity. Specifically, we asked participants to perform emotion identification task on facial, vocal, and bimodal emotion expressions varied in emotional intensity. As some studies (Yuan et al., 2007, 2008) indicated that emotional intensity effect is more prominent in negative emotions, only angry expressions were studied in the current study.

Based on the findings that emotional intensity modulates emotion processing in single modality (Sprengelmeyer and Jentzsch, 2006; Yuan et al., 2007; Dunning et al., 2010; Wang et al., 2015), and bimodal integration effect was more salient under high noisy background than under low noisy background (Jessen et al., 2012), we hypothesize that bimodal emotion integration effect was more conspicuous for stimuli with low emotional intensity stimuli than for those with high emotional intensity, following the principle of IE. As the integration effect was mainly denoted by N1 and P2 components (Paulmann et al., 2009; Balconi and Carrera, 2011; Jessen and Kotz, 2011; Jessen et al., 2012; Ho et al., 2015; Kokinous et al., 2015, 2017), we predict that the modulation of emotional intensity should be manifested on these components. To circumvent the confound of the physical difference in the direct comparison between bimodal and unimodal stimuli, we compare the brain responses associated bimodal emotions with the sum of those linked with both vocal and facial emotions (Brefczynski-Lewis et al., 2009; Hagan et al., 2009; Stevenson et al., 2012; Chen et al., 2016a,b). Based on the amplitude reduction of the early N1 component for bimodal emotions vs. unimodal emotions reported in previous studies (Jessen et al., 2012; Kokinous et al., 2015), we predicted that bimodal emotions should elicited reduced N1 relative to unimodal emotions. However, given the controversy regarding the P2 amplitude associated multimodal integration, specifically, while many studies reported a suppression P2 components in the bimodal compared with the unimodal condition (Pourtois et al., 2002; Paulmann et al., 2009; Kokinous et al., 2015), some other studies found an enhanced P2 components for bimodal relative to unimodal condition (Jessen and Kotz, 2011; Schweinberger et al., 2011), no prior hypotheses were made with regard to the increase or decrease of the P2 amplitude.

METHOD

Participants

Twenty-five university students (12 women, aged 19–27, mean 23.6 years) were recruited to participate in the experiment. All the participants reported normal auditory and normal or corrected-to-normal visual acuity, and were free of neurological or psychiatric problems. The study was approved by the Local Review Board for Human Participant Research and written informed consent was obtained prior to the study. The experimental procedure was in accordance with the ethical principles of the 1964 Declaration of Helsinki (World Medical Association Declaration of Helsinki, 1997). All participants were reimbursed ¥50 for their time. Five participants were excluded from EEG data analysis because of extensive artifacts.

Materials

Angry and neutral expressions for "嘿/hei/" and "喂/wei/" were chosen from our emotional expression database (the development of the database can be found in the Supplementary Materials). Based on the original expressions, vocally neutral-angry continua (33, 67, and 100% anger) were created for each

interjection using STRAIGHT (Kawahara et al., 2008), while facially neutral-angry continua were created using Morpheus Photo Morpher 3.16 (Morpheus Software LLC, Santa Barbara, CA, USA). And the bimodal expressions were created through presenting the vocal and facial expressions together. To verify the validity of the emotional intensity manipulation, 23 participants were asked to decide whether the synthesized expression expresses anger and to rate the emotional intensity level of the stimulus on a 9-point scale, with 1 being "not intense" and 9 being "highly intense." The mean ratios of anger and intensity level scores are presented in **Table 1**. The repeated measures ANOVA with Intensity (33, 67, and 100%) and Modality (Facial, Vocal, and Bimodal) as within subject factors on ratios of anger and intensity level confirmed the validity of our emotional intensity manipulation. Specifically, the analysis on ratios of anger showed significant main effect of Intensity [$F_{(2, 44)} = 53.32$, $p < 0.001$, $\eta^2_p = 0.71$], with the ratios of anger higher for 100% than for 67 and 33% ($p < 0.01$ and $p < 0.001$, respectively), and higher 67% than for 33% ($p < 0.001$). Similarly, the analysis on intensity levels showed significant main effect of Intensity [$F_{(2, 44)} = 85.37$, $p < 0.001$, $\eta^2_p = 0.80$], and significant interaction between Modality and Intensity [$F_{(4, 88)} = 9.76$, $p < 0.01$, $\eta^2_p = 0.31$]. Simple effect analyses showed that intensity level were larger for 100% than for 67 and 33% ($ps < 0.001$) under all modalities, and larger for 67% than for 33% under facial and bimodal conditions ($ps < 0.01$), while no difference was found between 67 and 33% under vocal condition ($p > 0.1$).

PROCEDURE AND DESIGN

Each participant was seated comfortably in a sound-attenuated room. Stimulus presentation was controlled using E-prime software. Auditory stimuli were presented via loudspeakers placed at both sides of the monitor while the facial stimuli presented on the monitor simultaneously at a central location on a black background (visual angle of 3.4° [width] × 5.7° [height] from a viewing distance of 100 cm). As illustrated in **Figure 1**, each trial was initiated by a 1,000 ms presentation of a white cross on the black computer screen. Then, a blank jitter varied randomly between 800 and 1,200 ms was presented, followed by the stimuli. Participants were required to decide whether the stimuli expressed anger by pressing the "J" or "F" button on the keyboard as accurately and quickly as possible. The response buttons assigned for "yes" or "no" was counterbalanced across participants. The presentation time for facial expressions and bimodal expressions were the same as its corresponding vocal expressions. Each response was followed by 2,000 ms of a blank screen. The stimuli consisted with facial, vocal and bimodal expressions varied in emotional intensity. Based on the orthogonal combination of 3 modals, 3 intensity levels, and 2 interjections, 18 emotional interjections were include. To enhance the signal to noise ratio, each stimulus was repeated 20 times, and therefore 360 emotional expressions were presented to participants as key stimuli. Moreover, to increase the ecological validity and balance the response of "yes" and "no" in the current study, 120 neutral expressions were presented to the participants

as filler. The total 480 stimuli were randomized and split up into five blocks and participants were given a short self-paced break to rest between blocks. Pre-training with 48 trials was included in order to make subjects familiar with the procedure.

EEG RECORDING

Brain activity was recorded at 64 scalp sites using tin electrodes mounted in an elastic cap (Brain Product, Munich, Germany) according to the modified expanded 10–20 system, each referenced on-line to FCZ. Vertical electrooculograms (EOGs) were recorded supra-orbitally and infra-orbitally from the right eye. The horizontal EOG was recorded as the left vs. right orbital rim. The EEG and EOG were amplified using a 0.05–100 Hz bandpass and continuously digitized at 1,000 Hz for offline analysis. The impedance of all electrodes was kept <5 kΩ.

Data Analysis
Behavioral Data
The rates of response anger and reaction times were collected as dependent variable, and were subjected to repeated measures ANOVA with *Modality* (F, V, and FV) and *Intensity* (33, 67, and 100%) as within-subject factors, to test the emotional intensity effect across modalities. To investigate the modulation of emotional intensity on bimodal integration, then, the predicted rates of response anger (pFV) were calculated following the equation [pFV = p(F) + p(V) − p(F)*p(V)] by Stevenson et al. (2014)[1], and subjected to ANOVA with *Intensity* (33, 67, and 100%) and *Modality-type* (FV, pFV) as within subjects factors. And whether the RTs in the bimodal condition exceeded the statistical facilitation predicted by probability summation of two unimodal conditions (Miller, 1982) was tested using the algorithm implemented in RMITest software (Ulrich et al., 2007; see **Figure 2** for graphic illustration). The degrees of freedom of the F-ratio were corrected according to the Greenhouse–Geisser method and multiple comparisons were Bonferroni adjusted. The effect sizes were shown as partial eta squared (η^2_p).
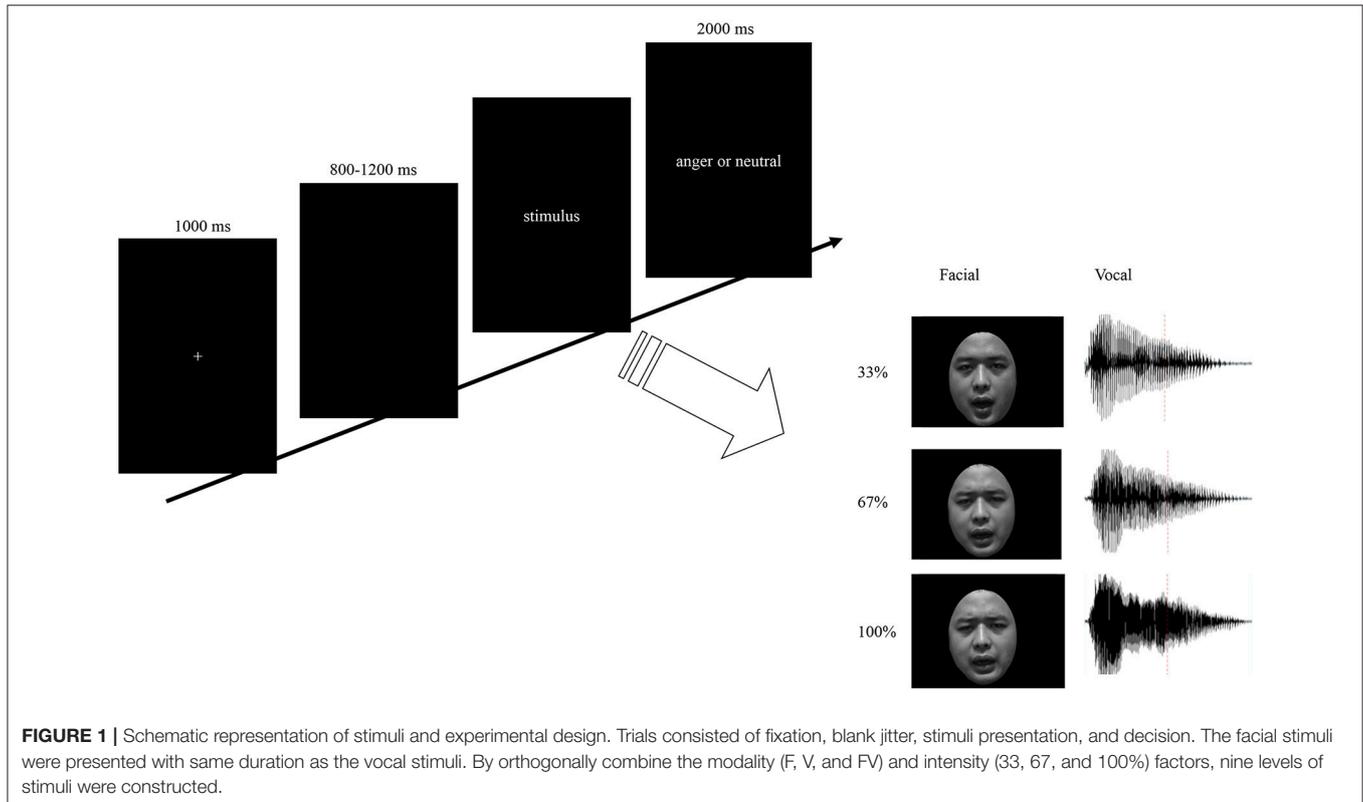
ERP Data
EEG data were preprocessed using EEGLAB (Delorme and Makeig, 2004), an open source toolbox running under the MATLAB environment. The data were down sampled at 250 Hz and then were high pass filtered at 0.1 Hz and low pass filtered at 40 Hz. After re-referenced offline to bilateral mastoid electrodes, the data were segmented to 1,000 ms epochs time-locked to the stimuli onset, starting 200 ms prior to stimuli onset. The epochs data were baseline corrected using the 200 ms before the stimuli onset. EEG epochs with large artifacts (exceeding ±100 μV) were removed and channels with poor signal quality were interpolated. Using the independent component analysis algorithm (Makeig et al., 2004), trials contaminated by eye blinks and other artifacts were detected and automatically excluded with the toolbox ADJUST (Mognon et al., 2011). More than 85% of the trials (34.64 ± 1.37, 34.28 ± 1.33, 33.88 ± 1.61, 34.28 ±

---

[1]Where p(V) is the rates of response anger for vocal expressions and p(F) is the rates of response anger for facial expressions.

TABLE 1 | Rates of response anger and Intensity level as function of emotional intensity (M ± SD).

	Rates of response anger			Intensity level		
	Vocal	**Facial**	**Bimodal**	**Vocal**	**Facial**	**Bimodal**
33%	0.41 ± 0.21	0.52 ± 0.27	0.61 ± 0.31	4.93 ± 1.25	4.28 ± 1.18	4.50 ± 1.12
67%	0.73 ± 0.17	0.73 ± 0.21	0.89 ± 0.15	5.05 ± 1.19	5.79 ± 1.21	5.71 ± 1.30
100%	0.89 ± 0.17	0.80 ± 0.13	0.93 ± 0.14	6.07 ± 1.31	6.46 ± 1.30	6.75 ± 1.22

FIGURE 1 | Schematic representation of stimuli and experimental design. Trials consisted of fixation, blank jitter, stimuli presentation, and decision. The facial stimuli were presented with same duration as the vocal stimuli. By orthogonally combine the modality (F, V, and FV) and intensity (33, 67, and 100%) factors, nine levels of stimuli were constructed.

1.48, 34.48 ± 1.53, 33.60 ± 1.36, 34.24 ± 1.36, 34.84 ± 1.37 and 33.80 ± 1.52 for nine levels, respectively) were retained in further analysis. ERP waveforms were computed separately as function of condition. The extracted average waveforms for each participant and condition were used to calculate grand-average waveforms (see **Figure 3** for ERP).
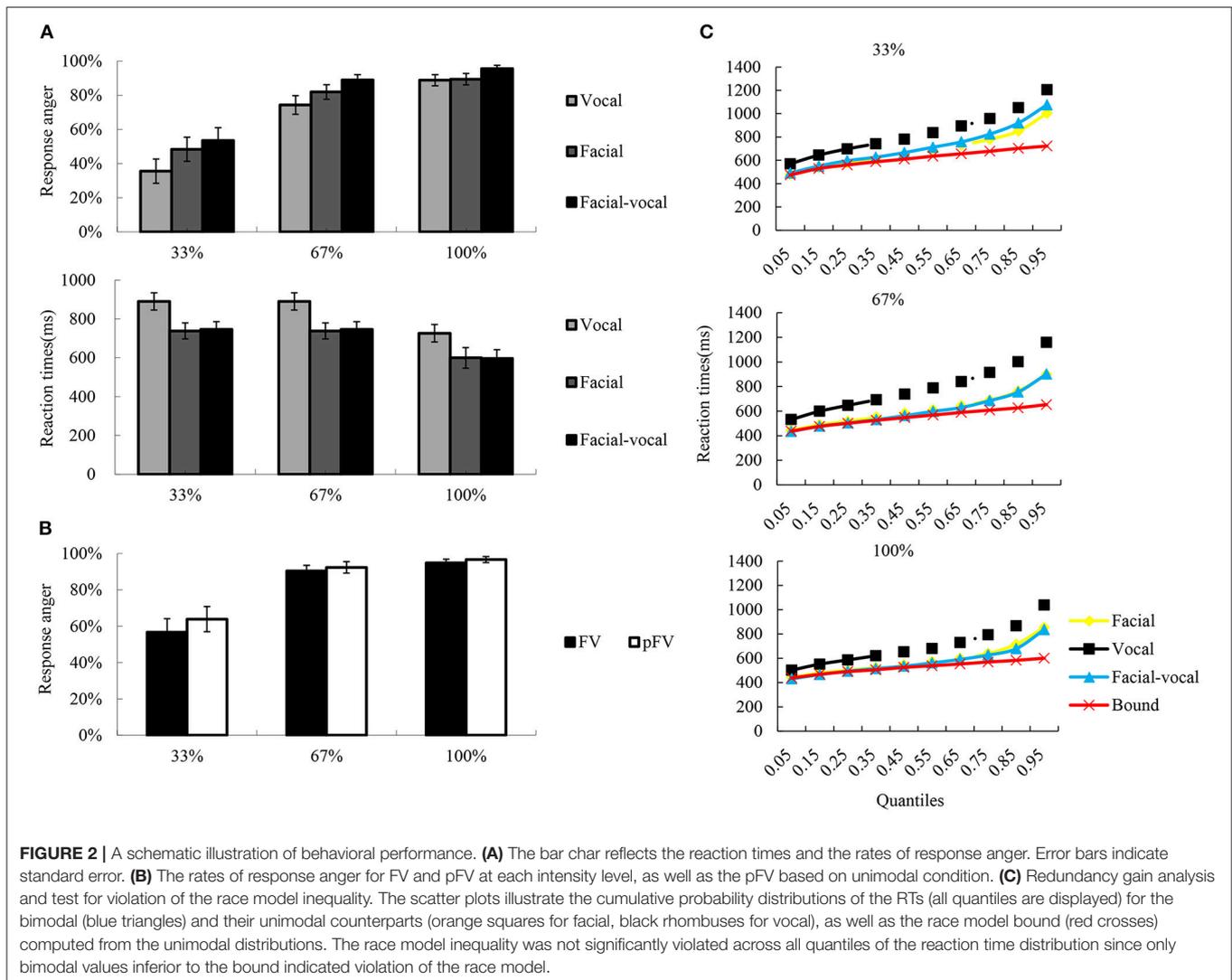
Given the simultaneously presentation of facial and vocal expressions in the current study and the strong visual processing component in the perception of audiovisual stimuli (Paulmann et al., 2009; Jessen and Kotz, 2011), the visual N1 and P2 were selected as the component of interested. Following the previous study addressing visual components associated with multimodal emotion integration (Paulmann et al., 2009) and visual inspection of the grand ERP waves, 80–120 ms and 120–180 ms were defined for N1 and P2, respectively. Since the N1 and P2 is found at anterior electrode-sites (Federmeier and Kutas, 2002; Paulmann et al., 2009), the anterior electrode-sites, were clustered as region of interest (Left anterior: F3, F5, F7, FC3, FC5, and FT7; middle anterior: F1, FZ, F2, FC1, FCZ and FC2; right anterior: F4, F6,

F8, FC4, FC6 and FT8; middle central: C1, CZ, C2, CP1, CPZ, and CP2). To exam the influence of intensity on bimodal emotion integration, bimodal effect (FV), and summed unimodal effect (F + V) were calculated and tested with the additive model (Besle et al., 2004; Stevenson et al., 2014) by carrying out repeated measures ANOVA with *Modality-type* (FV vs. F+V), *Intensity*, and *SROI* as within-subject factors.

RESULT

Behavioral Performance

The descriptive results of the rates of response anger and reaction times were depicted in **Figure 2A**. The two-way repeated measures ANOVA on rates of response anger showed significant main effect *Intensity* [$F_{(2, 38)} = 74.97$, $p < 0.001$, $\eta^2_p = 0.80$], with the rates of response anger higher for 100% (0.90 ± 0.03) than for 67% (0.83 ± 0.04, $p < 0.01$) and 33% (0.48 ± 0.06, $p < 0.001$), and higher for 67% than for 33% ($p < 0.001$). Also significant was the main effect of *Modality* [$F_{(2, 38)} = 8.71$, $p < 0.01$,
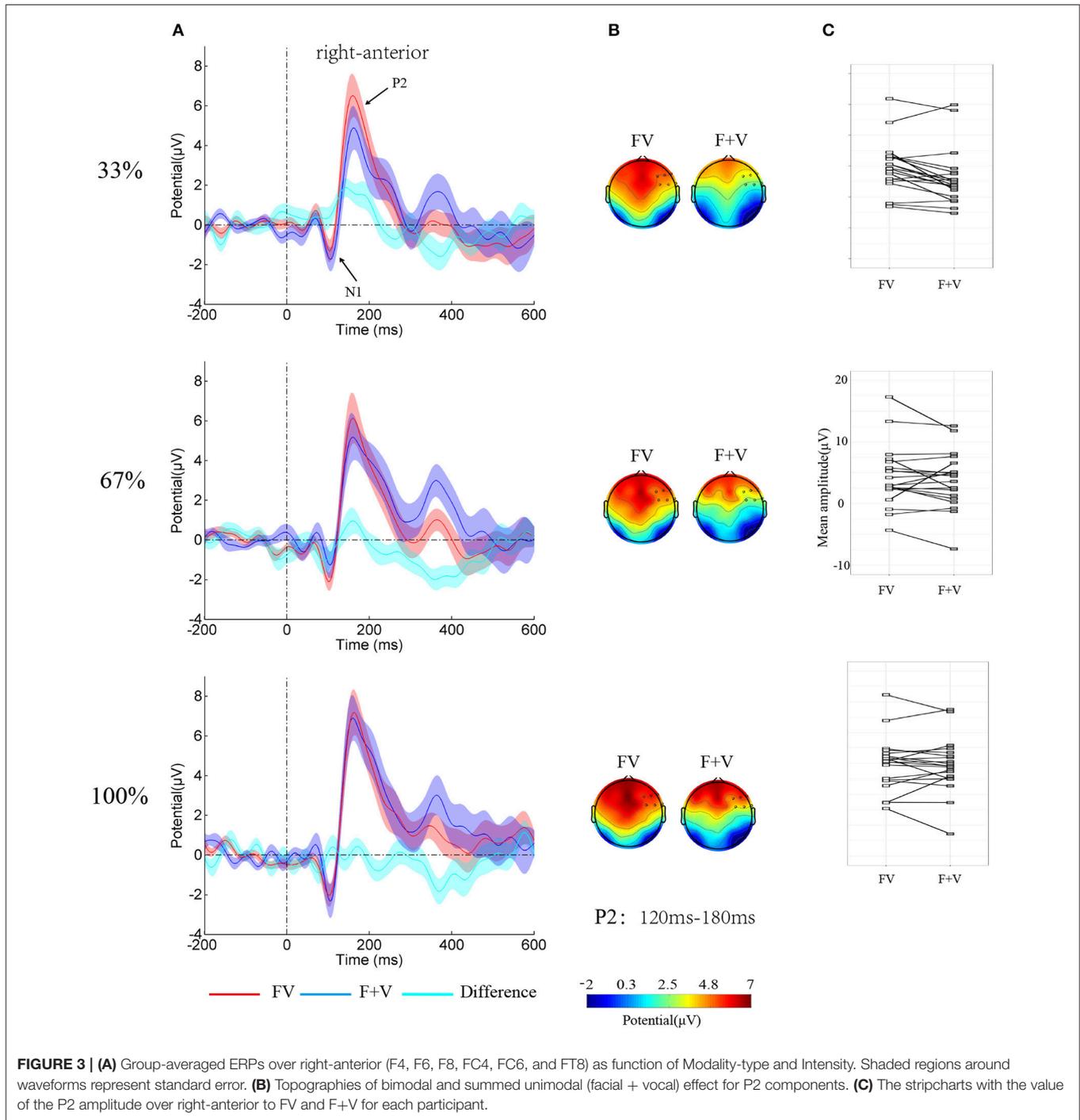
FIGURE 2 | A schematic illustration of behavioral performance. **(A)** The bar char reflects the reaction times and the rates of response anger. Error bars indicate standard error. **(B)** The rates of response anger for FV and pFV at each intensity level, as well as the pFV based on unimodal condition. **(C)** Redundancy gain analysis and test for violation of the race model inequality. The scatter plots illustrate the cumulative probability distributions of the RTs (all quantiles are displayed) for the bimodal (blue triangles) and their unimodal counterparts (orange squares for facial, black rhombuses for vocal), as well as the race model bound (red crosses) computed from the unimodal distributions. The race model inequality was not significantly violated across all quantiles of the reaction time distribution since only bimodal values inferior to the bound indicated violation of the race model.

$\eta^2{}_p = 0.31$], with the rates of response anger was higher for bimodal (0.81 ± 0.04) than for facial (0.72 ± 0.05, $p < 0.01$) and vocal (0.69 ± 0.05, $p < 0.001$), while the latter two conditions did not show significant differences ($p > 0.1$). However, as can be seen in **Figure 2B**, the integration analysis on rates of response anger showed no significant facilitation effects were observed under intensity levels.

Similarly, a two-way repeated measure ANOVA on reaction times revealed a significant main effect of *Modality* [$F_{(2, 38)} = 51.24$, $p < 0.001$, $\eta^2{}_p = 0.73$], a significant main effect of *Intensity* [$F_{(2, 38)} = 25.72$, $p < 0.001$, $\eta^2{}_p = 0.58$], and two-way interaction of *Modality* $\times$ *Intensity* [$F_{(4, 76)} = 7.29$, $p < 0.01$, $\eta^2{}_p = 0.28$]. Simple effect analysis show that the reaction times were faster for 100% (634 ± 45 ms) than for 67% (719 ± 52 ms, $p < 0.01$) and 33% (769 ± 60 ms, $p < 0.001$) under vocal condition, while the latter two were not significantly differentiated ($p > 0.1$). In facial condition, the reaction times were faster for 100% (531 ± 33 ms) than for 67% (546 ± 37 ms, $p < 0.01$) and 33% (622 ± 50 ms, $p < 0.01$) under vocal condition, while the former two

were not significantly differentiated ($p > 0.1$). The reaction times were faster for 100% (512 ± 39 ms) than for 67% (543 ± 40 ms, $p < 0.01$) and 33% (654 ± 52 ms, $p < 0.001$) under bimodal condition, and faster for 67% than for 33% ($p < 0.001$). However, as can be seen in **Figure 2C**, the integration analysis on reaction times showed no violation of the race model prediction over all quantiles of the reaction time distribution across three intensity levels.

ERP Results

The analysis on N1 amplitudes showed a significant two-way interaction of *Modality-type* $\times$ *SROI* [$F_{(3, 57)} = 3.64$, $p < 0.05$, $\eta^2{}_p = 0.16$]. Simple effect analysis showed no significant integration effect of *Modality-type* for each *ROI* ($ps > 0.05$). However, The analysis on the P2 amplitudes showed a significant main effect of *Modality-type* [$F_{(1, 19)} = 10.12$, $p < 0.01$, $\eta^2{}_p = 0.35$], a significant main effect of *Intensity* [$F_{(2, 38)} = 7.02$, $p < 0.01$, $\eta^2{}_p = 0.27$], a significant two-way interaction of *Modality-type* $\times$ *SROI* [$F_{(3, 57)} = 14.79$, $p < 0.01$, $\eta^2{}_p = 0.44$], and a significant three-way

FIGURE 3 | (A) Group-averaged ERPs over right-anterior (F4, F6, F8, FC4, FC6, and FT8) as function of Modality-type and Intensity. Shaded regions around waveforms represent standard error. **(B)** Topographies of bimodal and summed unimodal (facial + vocal) effect for P2 components. **(C)** The stripcharts with the value of the P2 amplitude over right-anterior to FV and F+V for each participant.

interaction of *Modality-type* × *Intensity* × *SROI* [$F_{(6, 114)} = 3.11$, $p < 0.05$, $\eta^2_p = 0.14$]. To break down the three-way interaction, separate ANOVAs were performed for each *ROI*. A significant interaction between *Intensity* and *Modality-type* [$F_{(2, 38)} = 3.61$, $p < 0.05$, $\eta^2_p = 0.16$] was found over *right-anterior area*, while no significant interactions were observed over other areas. Simple effect analysis showed that the amplitude larger for FV (4.77 ± 0.93 µV) than for F+V (3.07 ± 1.00 µV, $p < 0.01$). However, the

ANOVAs for 67% and 100% did not reveal any significant effects of *Modality-type*.

DISCUSSION

The present study tested the modulation of emotional intensity on facial-vocal emotion, thereby examined whether bimodal emotion integration follows the principle of IE. It was found

that the rates of response anger and reaction speed increased as emotional intensity across modalities. Critically, the P2 amplitudes elicited by bimodal condition were larger than the sum of two unimodal conditions for low emotional intensity stimuli, but nor for middle or high intensity stimuli. The significance of these findings will be addressed in the following discussion.

The rates of response anger and reaction speed for anger discrimination increased as the augment of emotional intensity across modalities. These findings were consistent with emotional intensity effect in facial emotion decoding, in which emotional intensity significantly correlated with classification of emotional expression (Herba et al., 2007; Dunning et al., 2010). The current findings were also in line with the study in vocal emotion (Juslin and Laukka, 2001; Wang et al., 2015), which reported that portrayals of the same emotion with different levels of intensity were decoded with different degrees of accuracy. Extending these studies, the current study reported the emotional intensity effect in bimodal emotion decoding.

One main findings of the current study is that the P2 amplitudes elicited by bimodal condition were larger than the sum of two unimodal conditions. P2 component has been repeatedly reported to be the neural marker of bimodal emotional integration during time range. For instance, Jessen and Kotz (2011) reported a larger P2 for bimodal condition than for vocal condition. Balconi and Carrera (2011) found that congruent audiovisual emotional stimuli elicited larger P2 compared to incongruent ones. However, several other studies reported a reduced P2 for congruent audiovisual emotional stimuli than for incongruent ones (Pourtois et al., 2002; Kokinous et al., 2015) and a suppression P2 components in the bimodal compared with the unimodal emotions (Pourtois et al., 2002; Paulmann et al., 2009; Kokinous et al., 2015). While the former line of studies interpret the larger P2 amplitude for congruent bimodal emotional expressions as a symbol of augment of emotional salience, the latter line of studies attribute the suppression P2 components in the bimodal vs. unimodal expressions to redundancy of multimodal stimuli. Obviously, the present findings are in line with the first line of studies (Jessen and Kotz, 2011; Schweinberger et al., 2011). Based on the fact that the visual P200 has been functionally linked to the emotional salience of a stimulus (Schutter et al., 2004; Paulmann et al., 2009), the present result suggested that the increase of P2 amplitudes for bimodal stimuli results from that fact that bimodal emotional cues increase the emotional salience embedded in the stimuli. Moreover, the current findings allow us speculate that the current differentiation in P2 components suggests the bimodal emotional integration take place during the stage of emotional information derivation.

More critically, the integration effect only existed in low emotional intensity stimuli, but not for high intensity stimuli, suggesting a modulation of emotional intensity on the facial-vocal emotion integration. Consistent with the previous studies (Leppänen et al., 2007; Wang et al., 2015), the behavioral and brain responses linearly varied as a function of intensity

levels regardless of sensory modalities. However, when test the integration effect as a function of emotional intensity, only the bimodal integration for stimuli with low emotional intensity reached the statistical significant level, suggesting the integration of bimodal emotional cues follows the IE principle. This finding is consistent with the previous findings that multimodal emotions elicited earlier N1 than unimodal emotions under high noisy background but not low noisy background (Jessen et al., 2012) and bimodal cues affect each other when the target modality was less reliable (Collignon et al., 2008). However, the present study extended the previous studies by showing that emotional intensity modulates bimodal emotional integration.

Inconsistent with the previous studies (Jessen et al., 2012; Kokinous et al., 2015, 2017), we did not find significant integration effect on behavioral responses. Moreover, although the N1 amplitude appeared to be smaller for the bimodal condition than for the sum of the unimodal condition, it failed to reach statistically significant level. This result is inconsistent with previous finding that the N1 amplitude for the bimodal condition is smaller than for vocal condition (Jessen and Kotz, 2011; Jessen et al., 2012; Kokinous et al., 2015). However, this result is consistent with the study by Paulmann et al. (2009) and Balconi and Carrera (2011), which only reported significant integration effect during P2 time range. One possible explanation for this discrepancy might be that the current study tested the integration effect with superadditivity, which is stricter than comparing brain response for bimodal condition with unimodal condition (Jessen and Kotz, 2011; Jessen et al., 2012; Kokinous et al., 2015). Another possibility is that we presented facial expression with full emotionality simultaneously with the vocal emotions. As the previous studies reporting integration effect in N1 was locked to the onset of the auditory stimuli, whether there is a discrepancy on the visual and auditory worth further exploration.

CONCLUSION

In sum, the present study manipulated the emotional salience of the stimulus itself and tested the facial-vocal emotional integration with the criterion of superadditivity. The rates of response anger and reaction speed increased as emotional intensity across modalities. More importantly, the P2 amplitudes were larger for bimodal expressions than for the sum of facial and vocal expressions for low emotional intensity stimuli, but not for middle and high emotional intensity stimuli. These findings indicated that bimodal emotional integration appears to be affected by emotional intensity and integration of facial-vocal angry expressions follows the principle of IE.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Key Laboratory of Behavior and Cognitive

Psychology in Shaanxi Province, Shaanxi Normal University with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Shaanxi Normal University.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: ZP and XC. Performed the experiments: ZP and XL. Contributed materials/analysis tools: ZP and XL. YL helped perform the analysis with constructive discussions. ZP and XC performed the data analyses and wrote the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnins.2017.00349/full#supplementary-material

REFERENCES

Balconi, M., and Carrera, A. (2011). Cross-modal integration of emotional face and voice in congruous and incongruous pairs: the P2 ERP effect. *J. Cogn. Psychol.* 23, 132–139. doi: 10.1080/20445911.2011.473560

Besle, J., Fort, A., and Giard, M. -H. (2004). Interest and validity of the additive model in electrophysiological studies of multisensory interactions. *Cogn. Process.* 5, 189–192. doi: 10.1007/s10339-004-0026-y

Brefczynski-Lewis, J., Lowitszch, S., Parsons, M., Lemieux, S., and Puce, A. (2009). Audiovisual non-verbal dynamic faces elicit converging fMRI and ERP responses. *Brain Topogr.* 21, 193–206. doi: 10.1007/s10548-009-0093-6

Chen, X., Han, L., Pan, Z., Luo, Y., and Wang, P. (2016a). Influence of attention on bimodal integration during emotional change decoding: ERP evidence. *Int. J. Psychophysiol.* 106, 14–20. doi: 10.1016/j.ijpsycho.2016.05.009

Chen, X., Pan, Z., Wang, P., Yang, X., Liu, P., You, X., et al. (2016b). The integration of facial and vocal cues during emotional change perception: EEG markers. *Soc. Cogn. Affect. Neurosci.* 11, 1152–1161. doi: 10.1093/scan/nsv083

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023

De Gelder, B., Pourtois, G., and Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proc. Natl. Acad. Sci. U.S.A.* 99, 4121–4126. doi: 10.1073/pnas.062018499

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Dolan, R., Morris, J., and de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10006–10010. doi: 10.1073/pnas.171288598

Dunning, J. P., Auriemmo, A., Castille, C., and Hajcak, G. (2010). In the face of anger: startle modulation to graded facial expressions. *Psychophysiology* 47, 874–878. doi: 10.1111/j.1469-8986.2010.01007.x

Federmeier, K. D., and Kutas, M. (2002). Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia* 40, 730–747. doi: 10.1016/S0028-3932(01)00193-2

Hagan, C. C., Woods, W., Johnson, S., Calder, A. J., Green, G. G., and Young, A. W. (2009). MEG demonstrates a supra-additive response to facial and vocal emotion in the right superior temporal sulcus. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20010–20015. doi: 10.1073/pnas.0905792106

Herba, C. M., Heining, M., Young, A. W., Browning, M., Benson, P. J., Phillips, M. L., et al. (2007). Conscious and nonconscious discrimination of facial expressions. *Vis. Cogn.* 15, 36–47. doi: 10.1080/135062806008 20013

Ho, H. T., Schröger, E., and Kotz, S. A. (2015). Selective attention modulates early human evoked potentials during emotional face–voice processing. *J. Cogn. Neurosci.* 27, 798–818. doi: 10.1162/jocn_a_00734

Jessen, S., and Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage* 58, 665–674. doi: 10.1016/j.neuroimage.2011.06.035

Jessen, S., Obleser, J., and Kotz, S. A. (2012). How bodies and voices interact in early emotion perception. *PLoS ONE* 7:e36070. doi: 10.1371/journal.pone.0036070

Juslin, P. N., and Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion* 1, 381–412. doi: 10.1037/1528-3542.1.4.381

Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., and Banno, H. (2008). "TANDEM-STRAIGHT: a temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation," in *Paper Presented at the Acoustics, Speech and Signal Processing, IEEE International Conference on ICASSP 2008* (Faculty of Systems Engineering Wakayama University).

Klasen, M., Chen, Y. H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040

Kokinous, J., Kotz, S. A., Tavano, A., and Schroger, E. (2015). The role of emotion in dynamic audiovisual integration of faces and voices. *Soc. Cogn. Affect. Neurosci.* 10, 713–720. doi: 10.1093/scan/nsu105

Kokinous, J., Tavano, A., Kotz, S. A., and Schröger, E. (2017). Perceptual integration of faces and voices depends on the interaction of emotional content and spatial frequency. *Biol. Psychol.* 123, 155–165. doi: 10.1016/j.biopsycho.2016.12.007

Leppänen, J. M., Kauppinen, P., Peltola, M. J., and Hietanen, J. K. (2007). Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. *Brain Res.* 1166, 103–109. doi: 10.1016/j.brainres.2007.06.060

Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci.* 8, 204–210. doi: 10.1016/j.tics.2004.03.008

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279. doi: 10.1016/0010-0285(82)90010-X

Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M. (2011). ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology* 48, 229–240. doi: 10.1111/j.1469-8986.2010.01061.x

Paulmann, S., Jessen, S., and Kotz, S. A. (2009). Investigating the multimodal nature of human communication. *J. Psychophysiol.* 23, 63–76. doi: 10.1027/0269-8803.23.2.63

Pourtois, G., Debatisse, D., Despland, P. -A., and de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Brain Res. Cogn. Brain Res.* 14, 99–105. doi: 10.1016/S0926-6410(02)00064-2

Sabbagh, M. A., Moulson, M. C., and Harkness, K. L. (2004). Neural correlates of mental state decoding in human adults: an event-related potential study. *J. Cogn. Neurosci.* 16, 415–426. doi: 10.1162/089892904322926755

Schelenz, P. D., Klasen, M., Reese, B., Regenbogen, C., Wolf, D., Kato, Y., et al. (2013). Multisensory integration of dynamic emotional faces and voices: method for simultaneous EEG-fMRI measurements. *Front. Hum. Neurosci.* 7:729. doi: 10.3389/fnhum.2013.00729

Schutter, D. J., de Haan, E. H., and van Honk, J. (2004). Functionally dissociated aspects in anterior and posterior electrocortical processing of facial threat. *Int. J. Psychophysiol.* 53, 29–36. doi: 10.1016/j.ijpsycho.2004.01.003

Schweinberger, S. R., Kloth, N., and Robertson, D. M. (2011). Hearing facial identities: brain correlates of face-voice integration in person identification. *Cortex* 47, 1026–1037. doi: 10.1016/j.cortex.2010.11.011

Sprengelmeyer, R., and Jentzsch, I. (2006). Event related potentials and the perception of intensity in facial expressions. *Neuropsychologia* 44, 2899–2906. doi: 10.1016/j.neuropsychologia.2006.06.020

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266. doi: 10.1038/nrn2331

Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J. Jr., and Rowland, B. A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Exp. Brain Res.* 198, 113–126. doi: 10.1007/s00221-009-1880-8

Stevenson, R. A., Bushmakin, M., Kim, S., Wallace, M. T., Puce, A., and James, T. W. (2012). Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr.* 25, 308–326. doi: 10.1007/s10548-012-0220-7

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., et al. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr.* 27, 707–730. doi: 10.1007/s10548-014-0365-7

Ulrich, R., Miller, J., and Schröter, H. (2007). Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods* 39, 291–302. doi: 10.3758/BF03193160

Wang, P., Pan, Z., Liu, X., and Chen, X. (2015). Emotional intensity modulates vocal emotion decoding in a late stage of processing: an event-related potential study. *Neuroreport* 26, 1051–1055. doi: 10.1097/WNR.0000000000000466

Wang, Y., Liu, F., Li, R., Yang, Y., Liu, T., and Chen, H. (2013). Two-stage processing in automatic detection of emotional intensity: a scalp event-related potential study. *Neuroreport* 24, 818–821. doi: 10.1097/WNR.0b013e328364d59d

World Medical Association Declaration of Helsinki. (1997). Recommendations guiding physicians in biomedical research involving human subjects. *Cardiovasc. Res.* 35, 2–3. doi: 10.1016/S0008-6363(97)00109-0

Yuan, J. J., Yang, J. M., Meng, X. X., Yu, F. Q., and Li, H. (2008). The valence strength of negative stimuli modulates visual novelty processing: electrophysiological evidence from an event-related potential study. *Neuroscience* 157, 524–531. doi: 10.1016/j.neuroscience.2008.09.023

Yuan, J., Zhang, Q., Chen, A., Li, H., Wang, Q., Zhuang, Z., et al. (2007). Are we sensitive to valence differences in emotionally negative stimuli? Electrophysiological evidence from an ERP study. *Neuropsychologia* 45, 2764–2771. doi: 10.1016/j.neuropsychologia.2007.04.018