SCIENCE CHINA
Life Sciences

• RESEARCH PAPER •
Progress of Projects Supported by NSFC

September 2013 Vol.56 No.9: 835–847
doi: 10.1007/s11427-013-4527-8

Single-trial ERP evidence for the three-stage scheme of facial expression processing

ZHANG DanDan[1,2], LUO WenBo[3,4*] & LUO YueJia[1,5*]

[1]*Institute of Affective and Social Neuroscience, Shenzhen University, Shenzhen 518060, China;*
[2]*State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China;*
[3]*School of Psychology, Liaoning Normal University, Dalian 116029, China*
[4] *Laboratory of Cognition and Mental Health, Chongqing University of Arts and Sciences, Chongqing 402160, China;*
[5] *Sichuan Research Center of Applied Psychology, Chengdu Medical College, Chengdu 610500, China*

Received March 26, 2013; accepted June 17, 2013; published online July 30, 2013

Using a rapid serial visual presentation paradigm, we previously showed that the average amplitudes of six event-related potential (ERP) components were affected by different categories of emotional faces. In the current study, we investigated the six discriminating components on a single-trial level to clarify whether the amplitude difference between experimental conditions results from a difference in the real variability of single-trial amplitudes or from latency jitter across trials. It is found that there were consistent amplitude differences in the single-trial P1, N170, VPP, N3, and P3 components, demonstrating that a substantial proportion of the average amplitude differences can be explained by the pure variability in amplitudes on a single-trial basis between experimental conditions. These single-trial results verified the three-stage scheme of facial expression processing beyond multitrial ERP averaging, and showed the three processing stages of "fear popup", "emotional/unemotional discrimination", and "complete separation" based on the single-trial ERP dynamics.

facial expression, single trial, event-related potential (ERP), three stages

Citation: Zhang D D, Luo W B, Luo Y J. Single-trial ERP evidence for the three-stage scheme of facial expression processing. Sci China Life Sci, 2013, 56: 835−847, doi: 10.1007/s11427-013-4527-8

The perception of emotional expression is one of the most highly developed visual skills in humans and plays a critical role in the regulation and facilitation of social interactions. Numerous reports have used brain imaging to elucidate the neural pathways and cognitive models of facial expression [1−3]; however, the detailed and integrated temporal organization of emotional face processing remains relatively unclear [4,5]. An understanding of the temporal sequence of neural activity is essential for a comprehensive understanding of facial expression processing. Given the biological and social significance of emotions, different facial expressions, such as fear and happiness, must be rapidly discerned by the individual for the real-time regulation of behavior [6]. In addition, upon encountering a complex environment involving various emotional stimuli, the brain must evaluate incoming stimuli and allocate more cognitive resources to accelerate and enhance the processing of important emotional events first [7].

It has been shown that attention modulates the neural responses to different facial expressions [5,8–10]. In a previous study, we investigated the average event-related potentials (ERPs) elicited by emotional faces in a rapid serial visual presentation (RSVP) paradigm [11]. During the experiment, participants were asked to detect two target stimuli (T1 and T2) within an interval of less than 500 ms in a rapidly presented stimuli stream. Similar to Flaisch et al.

*Corresponding author (email: wenbo9390@sina.com; luoyj@bnu.edu.cn)

[10], we implemented the RSVP task with facial expressions as the T2 targets. We found that the average amplitudes of six ERP components, which appeared at various time intervals after facial stimuli onset, were affected by different emotional faces (i.e., fearful, happy, and neutral expressions). Based on these results, a three-stage time scheme of facial expression processing was preliminarily proposed [11]. In brief, we defined 'Stage 1' as the discrimination of fearful facial expressions, reflected by increased amplitudes of N1 and P1 components. 'Stage 2' was defined as the process of distinguishing emotional faces from unemotional ones, with larger N170 and vertex positive potential (VPP) amplitudes in response to emotional expressions than to neutral expressions. In 'Stage 3', the brain is able to distinguish among various expression categories, which is reflected by distinct amplitudes of N3 and P3 components for fearful, happy, and neutral faces.

However, the ERP analyses in our previous work focused on the conventional averaging method, i.e., exploring the amplitude differences based on stimulus-locked averaging data. It has long been known that single-trial ERPs typically show significant variability under the same stimulus condition with regard to the amplitude and the latency of certain components, and that stimulus averaging is likely to obscure any information contained in trial-to-trial variations [12–14]. As illustrated in Figure 1, the differing amplitudes of the average ERP between two experimental conditions might be due to single-trial amplitude differences, latency variations (i.e., latency jitter), or both. Specifically, the larger amplitude in average ERP elicited by certain category of facial expression may reflect truly stronger neural activity as compared with another expression (Figure 1A), or it may be due to an ERP response that shows a more consistent latency from trial to trial (Figure 1B). Either or both of these alternatives may lead to average amplitude differences, thus, data analysis on the single-trial level should be carried out to clarify the neural mechanisms underlying ERP results. In fact, several investigations have been applied to the data collected in different cognitive and clinical experiments to explore single-trial ERP characteristics [15,16]. For instance, Rousselet et al. [17] showed that a larger N170 amplitude elicited by faces compared with other objects may be explained by a true increase in single-trial amplitude. Ford et al. [18] explored the decreased average P3 amplitude in schizophrenic patients compared to controls and found that both single-trial amplitude and latency jitter contributed to the average amplitude reduction in the P3 component.

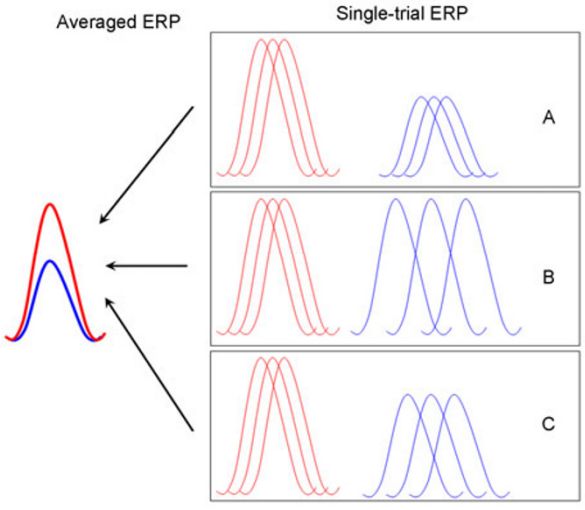

Figure 1 Illustration of three potential mechanisms for amplitude differences in average ERP. The average amplitude differences between two experimental conditions may due to: true amplitude difference in single trials (A), the difference in single-trial latency variability (B), or both mechanisms (C).

As a follow-up study, the current work analyze, on the single-trial level, ERP data acquired in an RSVP paradigm very similar to that of Luo et al. [11]. Assuming that the peak amplitude and peak latency across trials obey normal distributions [19], we compared the mean of single-trial peak amplitude (associated with Figure 1A) and the standard deviation (SD) of single-trial peak latency (associated with Figure 1B) following presentation with different emotional stimuli. We hypothesized that each of the alternatives shown in Figure 1 has different cognitive implications: (i) when the available attention resources are limited in an RSVP experiment, the reduction of peak amplitude in all trials under one emotional condition may suggest the allocation of fewer neural networks for the processing of the associated facial expression [18] (Figure 1A); (ii) a lower peak latency variability likely indicates a smaller fluctuation in neural processing speed (i.e., a relatively consistent speed) associated with certain categories of facial expression (Figure 1B); (iii) the differences in the proportion of involved neural networks, as well as in the processing speeds of the brain, both contribute to average amplitude differences (Figure 1C). Thus, the current study was undertaken to investigate the single-trial characteristics and potential neural mechanisms underlying the average ERP differences revealed in our previous work. Based on these results, we generalized the three-stage scheme of facial expression processing from the framework of traditional average ERP estimates to the single-trial level.

1 Materials and methods

1.1 Participants

Seventeen healthy participants (nine females; age range 19–26 years) were recruited from Southwest University in China as paid volunteers. All participants were right-handed and had normal or corrected-to-normal vision. The experimental protocol was approved by the local ethics committee and in accordance with the Declaration of Helsinki.

1.2 Stimuli and experimental procedure

The experiment was performed as in our previous study,

with slight modifications [11]. In brief, 30 pictures of human faces (12 inverted neutral faces, six upright neutral faces, six upright happy faces, and six upright fearful faces) were selected from the Chinese Facial Affective Picture System (CFAPS), with equal face number between males and females [11]. The 18 upright face pictures were evaluated by additional 60 volunteers and the average scores showed that the pictures differed significantly in valence ($F_{(2,15)}$=338.03, P<0.001; happy: 6.90±0.22, neutral: 4.70±0.31, fearful: 2.75±0.29) while there was no difference in arousal ($F_{(2,15)}$<1).

Experimental procedure was designed based on the RSVP paradigm (Figure 2). In total, 14 pictures (12 inverted neutral faces as distractive stimuli and two upright pictures as target stimuli) were presented in one experimental trial, with no inter-picture interval (119 ms for each picture). The first target stimulus (T1) was one of the three upright house pictures and the second target stimulus (T2) was one of the eighteen upright face pictures. The T1 appeared randomly and equiprobably as the third, fourth, or fifth picture, followed sequentially by one distractive stimulus, the T2, and other distractive stimuli. Participants were required to respond to two questions as accurately as possible regarding the appearance of the house in T1 (Q1: which house did you see) and the category of facial expression in T2 (Q2: which category of facial expression did you see) (Figure 2). The next experiment trial automatically began following the response to Q2.

To remove the superposed electrical activity elicited by the prior- and post-distractive stimuli in order to obtain the ERPs elicited purely by T2, a baseline condition was designed with the face picture at T2 replaced by a blank screen. Four conditions were presented in a random order during the experiment (experimental condition was defined by T2, i.e., neutral face, happy face, fearful face, and blank screen). To facilitate the single-trial analysis of ERP data, we increased the trial number per condition to 108 trials, compared with 60 trials in the previous experiment.

1.3 EEG data recording and preprocessing

Brain electrical activity was recorded referentially against left mastoid and off-line re-referenced to averaged mastoids, by a 64-channel amplifier with a sampling frequency of 500 Hz (Brain Products, Gilching, Germany). Four electrooculogram (EOG) channels were included to pick up eye movement artifacts. Electrode impedances were kept below 5 kΩ. The data analyses in this study were performed using Matlab R2011a (MathWorks, Natick, USA) and SPSS Statistics 17.0 (IBM, Somers, USA).

The continuous EEGs were segmented in association with T2, beginning 200 ms prior to the stimulus onset and lasting for 1200 ms. All epochs were baseline corrected with respect to the mean voltage over the 200 ms preceding stimulus-onset. Trials were accepted only if the two questions were properly answered. Epochs containing obvious EOG were rejected manually. Finally, different trial numbers per condition in each subject were available for analysis (happy face: 96.5±20.0 trials, neutral face: 93.2±22.5 trials, fearful face: 97.8±18.9 trials, and baseline condition: 99.5±19.7 trials).

1.4 Average ERP acquisition and analysis

To comprehensively elucidate the time course of facial expression processes, the conventional average ERP analysis was replicated in a more concise form compared with our previous study [11].

The recorded EEG data were filtered with a 0.5–30 Hz finite impulse response (FIR) filter with zero phase distortion, followed by epoch segmentation. Stimulus-locked average ERPs under happy, neutral, and fearful face conditions were computed separately for each participant as the difference between facial expressions and baseline condition (i.e., the average ERP in facial expression condition subtracts the average ERP in baseline condition).

Six ERP components were analyzed in the present study. Among these, the frontocentral N1 and occipital P1 components reflect the initial visual perception; VPP is the positive counterpart of the face-sensitive N170 in the occipito-temporal cortex; and the later components, i.e., N3 and P3, reflect the more refined processing of facial expression. Component peaks were manually detected from the average

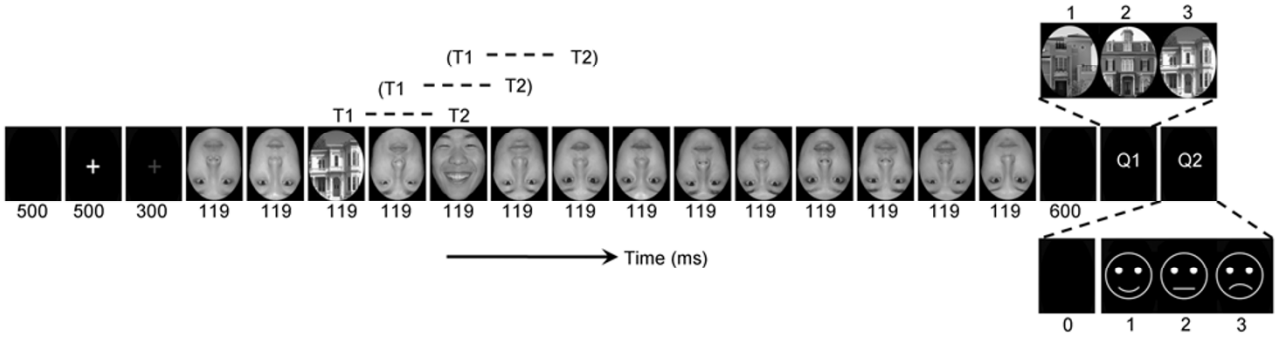

Figure 2 Schematic diagram of one experimental trial. The first and the second questions (Q1 and Q2) test the response accuracy of the first and the second targets (T1 and T2), respectively. Three alternative answers in Q1 indicate three possible house pictures shown at T1. Four alternative answers in Q2 indicate four experimental conditions: 0-blank screen (baseline), 1-happy face, 2-neutral face, 3-fearful face.

ERP on a subject-by-subject basis. The N1 and P1 components were measured as the baseline-to-peak amplitudes, while the peak-to-peak amplitudes (i.e., the amplitude difference between the associated peak and the previous peak) were computed for the other four components [11]. To obtain reliable average ERPs with a high signal-to-noise ratio, each component was measured as the averaged signal based on data from four electrodes where the component was with the highest amplitudes (baseline-to-peak or peak-to-peak amplitudes). The electrode selection was based on visual inspection of the topographies and the previous relevant studies [11,20]. In particular, the N1 and VPP components were analyzed using the average data at the F1, F2, Fz, and FCz electrodes; the P1 component was analyzed using the average data at the PO3, PO4, Pz, and POz electrodes; the N170 component at P7, P8, PO7, and PO8; the N3 component at AF7, AF8, F7, and F8; and the P3 component at FCz, Cz, CPz, and Pz.

A repeated measures single-factor analysis of variance (ANOVA) was performed with facial expression as the within-subjects factor and with peak latency and peak amplitude of the average ERP as dependent variables, followed by pairwise comparisons. For all statistically analyses in this study, Greenhous-Geisser corrections were performed where appropriate. Partial eta-squared (η_P^2) was reported to demonstrate the effect size in ANOVA tests.

1.5 Single-trial peak detection and analysis

Single-trial peak detection was performed on the same ERP data used in the average ERP analysis; therefore the ERPs were averaged from four electrodes. The signal processing of the single-trial analysis is summarized in Figure 3. ERP trials without electrical superposition were acquired by subtracting the average ERP in the baseline condition from each individual trial in three emotional conditions. The Maximum Likelihood Estimation (MLE) technique was then employed to detect the occurrence of the component peak [21]. Compared with traditional peak detection algorithms such as peak-picking and template matching (cross-correlation or cross-covariance), the MLE method performs better on the dataset with a realistic signal-to-noise ratio [21,22]. The MLE assumes that the ERP signal hidden in EEG background activity has an invariant shape but may vary both in its latency and in its amplitude. By maximizing

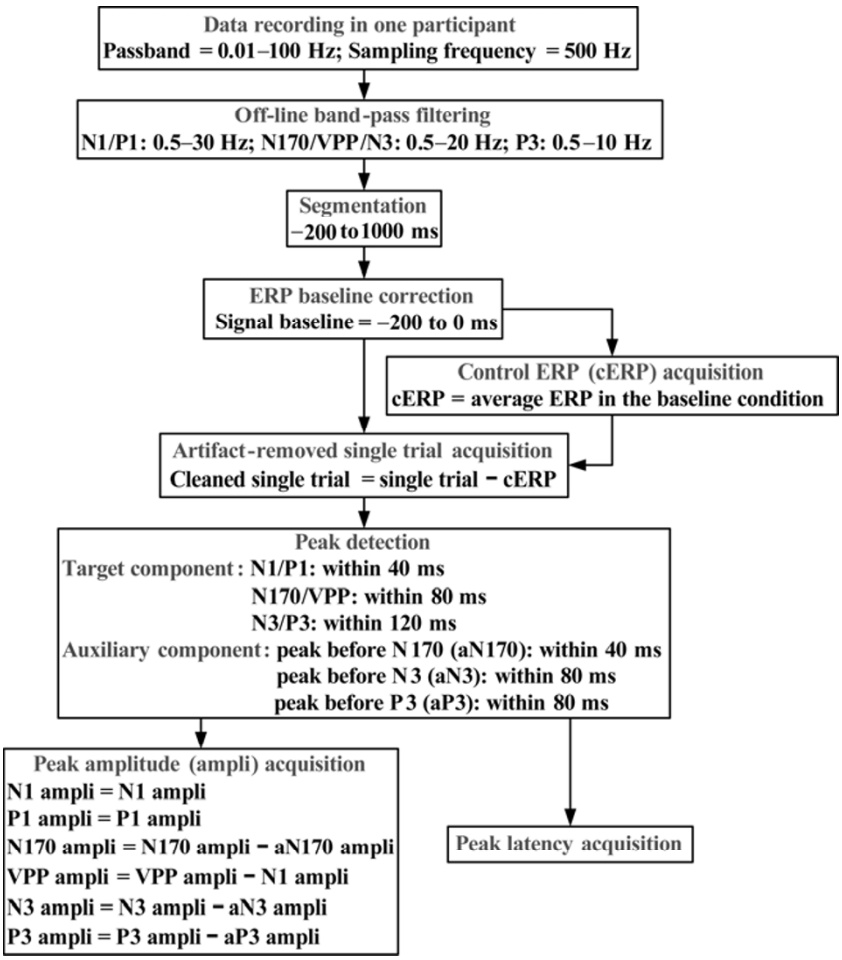

Figure 3 Signal flow graph and calculation parameters for the acquisition of single-trial peak amplitude and peak latency.

the log likelihood function of this model in the frequency domain, the MLE estimates the unknown parameters of signals such as latencies and amplitudes in single trials using iterative Fisher scoring [23,24]. Calculation parameters (10 frequency components and five iterations) were set as the authors recommend [16,21,22]. Finally, the peak amplitude (absolute amplitude for N1 and P1, peak-to-peak amplitude for N170, VPP, N3, and P3) and peak latency of each single trial were acquired for further analyses.

As illustrated in Figure 1, the peak amplitude of the average ERP is primarily influenced by two variables, namely, the mean of the single-trial peak amplitude and the SD of the single-trial peak latency. Repeated measures single-factor ANOVAs were conducted with facial expression as the within-subjects factor and with the mean of single- trial peak amplitude and the SD of single-trial peak latency as two dependent variables, followed by pairwise comparisons. Finally, a multiple linear regression (hierarchical method) was employed to explore the relationships between the criterion variable (i.e., the peak amplitude in average ERP) and the two predictor variables (i.e., the mean of single-trial peak amplitude and the SD of single-trial peak latency) [25].

2 Results

2.1 Traditional average ERP analysis

The behavioral results and the peak latencies of six ERP components were statistically identical to those found in our previous study [11]. We focused on the peak amplitude of average ERPs in this section.

The ANOVA test reveals that N1 and P1 amplitudes were significantly affected by facial expression ($F_{(2,32)}$=8.12, P=0.001; $F_{(2,32)}$=6.48, P<0.01). Fearful faces elicited larger N1 (mean±SD: −2.53±1.34 μV) and P1 amplitudes (4.81± 2.53 μV) than did happy (N1: −1.61±1.46 μV, P<0.01; P1: 3.51±1.80 μV, P<0.05) and neutral faces (N1: −1.91±1.19 μV, P<0.01; P1: 3.34±1.80 μV, P<0.01), while the latter two conditions do not show significant amplitude differences (P=0.21 for N1; P=0.68 for P1).

The N170 and VPP components displayed significant amplitude differences at three facial expressions ($F_{(2,32)}$= 17.2, P<0.001; $F_{(2,32)}$=21.5, P<0.001). Happy faces (N170: −6.24±3.09 μV, P<0.01; VPP: 8.95±3.18 μV, P<0.001) and fearful faces (N170: −6.46±2.72 μV, P<0.001; VPP: 9.29±3.24 μV, P<0.001) elicited larger N170 and VPP amplitudes than did neutral faces (N170: −4.98±2.02 μV; VPP: 6.77±2.55 μV), while the former two emotional conditions do not show significant amplitude differences (P=0.27 for N170; P=0.41 for VPP).

Facial expression had a strong effect on N3 and P3 amplitudes ($F_{(2,32)}$=8.90, P=0.001; $F_{(2,32)}$=25.3, P<0.001). Fearful faces elicited larger N3 (−10.9±6.07 μV) and P3 amplitudes (13.4±3.75 μV) than did happy (N3: −9.64± 5.42 μV, P<0.05; P3: 11.5±3.65 μV, P<0.01) and neutral faces (N3: −8.26±4.16 μV, P<0.01; P3: 8.71±2.49 μV, P<0.001). Happy faces elicited larger N3 (P<0.05) and P3 amplitudes (P=0.001) than did neutral faces. Figure 4 displays the ERP waveforms of the grand-mean ERPs in three emotional conditions.

2.2 Single-trial analysis of peak amplitude and peak latency

2.2.1 Mean of the single-trial peak amplitude

The ANOVA result demonstrates that the mean single-trial N1 peak amplitude was significantly affected by facial expression ($F_{(2,32)}$=3.46, P=0.043, η_P^2=0.143). Fearful faces elicited a larger single-trial peak amplitude (−6.15±1.69 μV) than did happy faces (−5.28±1.93 μV, P<0.01), while there are no significant amplitude differences between neutral (−5.73±1.56 μV, P=0.14) and happy conditions, or between neutral and fearful conditions (P=0.05).

The mean single-trial P1 peak amplitude was significantly affected by facial expression ($F_{(2,32)}$=6.28, P=0.002, η_P^2=0.311). Single-trial peak amplitudes were enhanced specifically for fearful faces (7.31±2.82 μV) relative to happy (6.09±2.26 μV, P<0.05) and neutral faces (5.98± 2.23 μV, P<0.05), while the latter two conditions do not show significant amplitude differences (P=0.76).

The mean single-trial peak amplitudes of the N170 and VPP components showed significant differences at three facial expressions ($F_{(2,32)}$=16.4, P<0.001, η_P^2=0.506; $F_{(2,32)}$=17.0, P<0.001, η_P^2=0.516). Happy faces (N170: −9.56±3.22 μV, P=0.001; VPP: 16.8±4.96 μV, P<0.001) and fearful faces (N170: −9.76±3.09 μV, P<0.001; VPP: 17.1±4.46 μV, P<0.001) elicited larger N170 and VPP single-trial amplitudes than did neutral faces (N170: −8.35± 2.26 μV; VPP: 14.7±3.83 μV), while the former two emotional conditions do not show significant amplitude differences (P=0.30 for N170; P=0.52 for VPP).

Facial expression had a strong effect on the mean single-trial peak amplitudes of the N3 and P3 components ($F_{(2,32)}$=15.3, P<0.001, η_P^2=0.489; $F_{(2,32)}$=21.2, P<0.001, η_P^2=0.570). Fearful faces elicited larger N3 (−18.7±6.64 μV) and P3 amplitudes (19.1±4.74 μV) than did happy faces (N3: −17.6±6.97 μV, P<0.05; P3: 17.4±4.66 μV, P<0.05) and neutral faces (N3: −15.6±5.34 μV, P<0.001; P3: 14.6± 3.12 μV, P<0.001). Happy faces elicited larger N3 (P<0.01) and P3 amplitudes (P<0.01) than did neutral faces. Representative data from one participant, as an example of the single-trial amplitude differences observed with three emotional conditions, are shown in Figure 5.

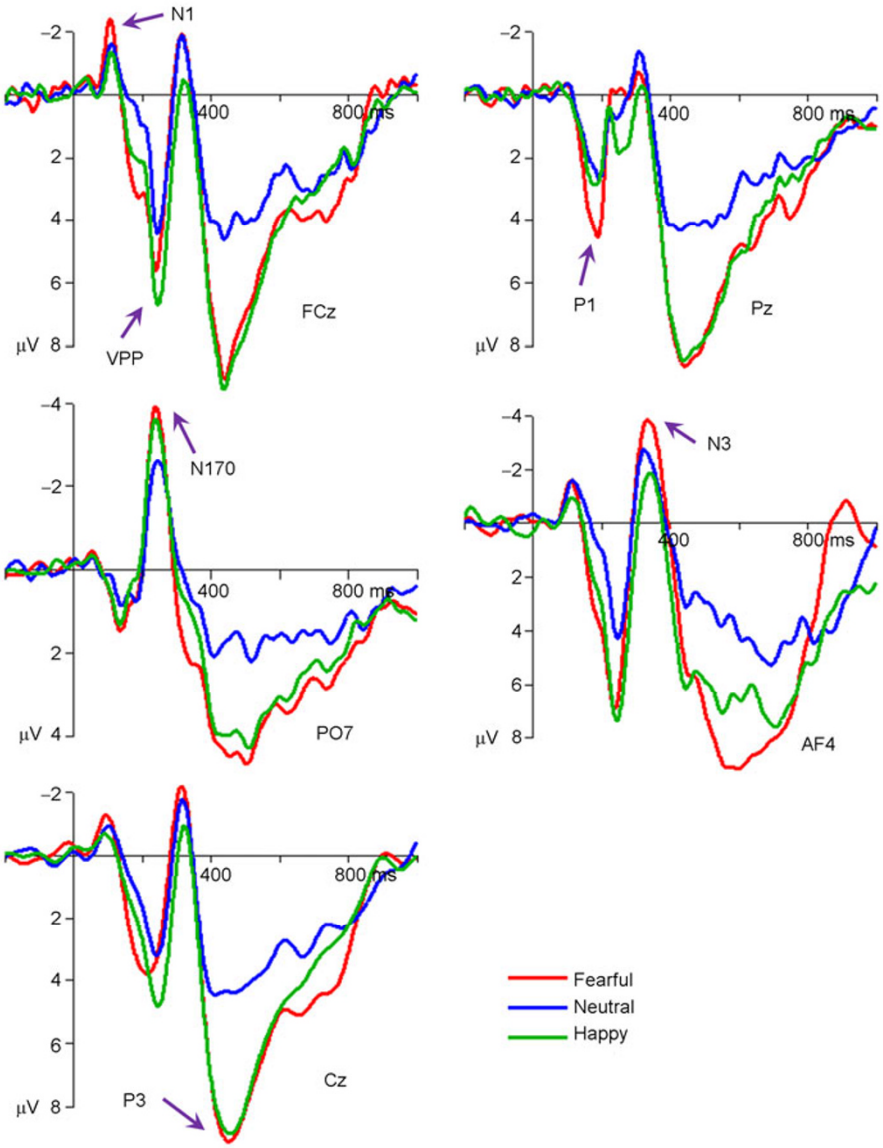

Figure 4 The grand-mean ERP waveforms of 17 participants in three emotional conditions.

2.2.2 *SD of the single-trial peak latency*

Facial expression significantly influenced the SD of the single-trial N1 peak latency ($F_{(2,32)}$=3.86, P=0.038, η_P^2 = 0.144). Neutral faces elicited a larger latency jitter (SD of latency=13.5±0.74 ms) than did fearful faces (SD of latency=13.0±0.71 ms, P<0.05), while there are no significant differences between happy (SD of latency=13.2±0.79 ms, P=0.21) and fearful conditions, or between happy and neutral conditions (P=0.12).

No significant effect of facial expression is found in the SD of single-trial peak latency for the other five components (P1: $F_{(2,32)}$<1, SD of latency=12.2±1.21 ms, 12.3±1.27 ms, and 12.4±1.48 ms for fearful, happy, and neutral faces; N170: $F_{(2,32)}$=1.63, P=0.21, SD of latency=17.6±2.65 ms, 18.0±2.93 ms, and 18.5±2.93 ms for fearful, happy, and neutral faces; VPP: $F_{(2,32)}$<1, SD of latency=20.8±2.77 ms, 20.7±1.38 ms, and 21.2±2.18 ms for fearful, happy, and neutral faces; N3: $F_{(2,32)}$=1.15, P=0.33, SD of latency=28.5±3.10 ms, 29.2±2.20 ms, and 28.1±3.92 ms for fearful, happy, and neutral faces; P3: $F_{(2,32)}$<1, SD of latency=34.5±3.26 ms, 34.9±2.75 ms, and 35.6±3.35 ms for fearful, happy, and neutral faces). A pairwise comparison of the average ERP amplitude, the mean amplitude of the single-trial ERP, and the latency SD for the single-trial ERP are summarized in Table 1.

Finally, for the sake of convenient observation, the average ERPs with and without latency realignment are shown in Figure 6. We also performed a two-tailed Pearson correlation between the SD of single-trial latency and the difference in component amplitude between original and realigned average ERPs (pooling fearful, happy, and neutral

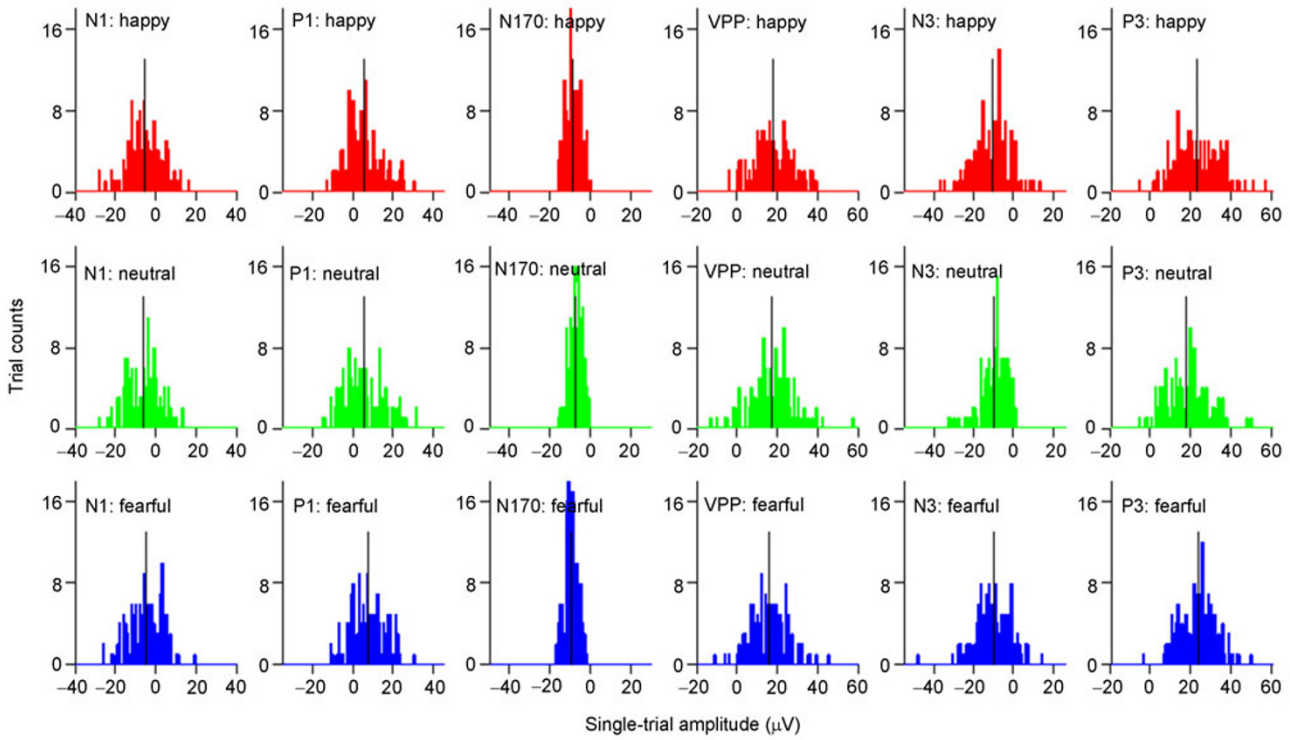

Figure 5 Single-trial amplitude distribution from one representative participant. Six columns display the single-trial peak amplitudes of the six components. The mean of the single-trial peak amplitude in each emotional condition is indicated by a black vertical line.

Table 1 Pairwise comparison results of the peak amplitude in average ERP, the mean single-trial peak amplitude, and the SD of single-trial peak latency[a)]

Component	Average ERP	Single-trial ERP	
	Peak amplitude	Mean of peak amplitude	SD of peak latency
N1	F>H, F>N	F>H	F<N
P1	F>H, F>N	F>H, F>N	No diff
N170	F>N, H>N	F>N, H>N	No diff
VPP	F>N, H>N	F>N, H>N	No diff
N3	F>H>N	F>H>N	No diff
P3	F>H>N	F>H>N	No diff

a) $P<0.05$. F, fearful faces; H, happy faces; N, neutral faces. No diff, no significant difference revealed by the ANOVA (i.e., $F<1$ or $P>0.05$).

data; Figure 7). Results showed that the SD of latency was significantly correlated with the amplitude difference in all the six ERP components (N1: $r=0.57$, $P<0.001$; P1: $r=0.64$, $P<0.001$; N170: $r=0.60$, $P<0.001$; VPP: $r=0.36$, $P=0.009$; N3: $r=0.44$, $P=0.001$; P3: $r=0.32$, $P=0.022$).

2.3 Multiple linear regression

Compared with the SD of the single-trial peak latency, the mean of the single-trial peak amplitude accounts for a much larger proportion of the average amplitude variance (Table 1). Therefore, we implemented multiple linear regressions using a hierarchical method, with the mean single-trial amplitude as the single predictor in Model 1, and with both the single-trial variables as predictors in Model 2 (Table 2). The results revealed that (i) the regression models provide a good fit to the data with highly significant F values ($P<0.001$), (ii) the adjusted R square given by the first model indicates that the mean single-trial amplitude could explained 60%, 91%, 87%, 74%, 85%, and 88% of the variation in the average amplitudes of N1, P1, N170, VPP, N3, and P3, respectively, while the inclusion of an addictive predictor (i.e., the SD of the single-trial latency) resulted in an additional explanation rate of less than 10% (7%, 4%, 7%, 9%, 4%, and 1% associated with the five components), and (iii) in the second model, the standardized regression coefficient of the first predictor was larger than that of the second predictor, which was also reflected by the slope of

Table 2 Multiple linear regression results[a)]

Component	Model 1 (one predictor)			Model 2 (two predictors)			
	$F_{(1,49)}$	R^2	Beta	$F_{(2,48)}$	R^2	Beta1	Beta2
N1	77.5 P<0.001	0.60	0.78 P<0.001	52.7 P<0.001	0.67	0.91 P<0.001	0.30 P=0.001
P1	500 P<0.001	0.91	0.95 P<0.001	493 P<0.001	0.95	0.98 P<0.001	–0.21 P<0.001
N170	338 P<0.001	0.87	0.94 P<0.001	412 P<0.001	0.94	0.88 P<0.001	0.27 P<0.001
VPP	142 P<0.001	0.74	0.86 P<0.001	125 P<0.001	0.83	0.84 P<0.001	–0.31 P<0.001
N3	286 P<0.001	0.85	0.92 P<0.001	205 P<0.001	0.89	0.93 P<0.001	0.20 P<0.001
P3	387 P<0.001	0.88	0.94 P<0.001	205 P<0.001	0.89	0.96 P<0.001	–0.09 P=0.07

a) Using the hierarchical method, the mean of single-trial peak amplitude and the SD of single-trial peak latency were considered as the first and the second predictors, and the average peak amplitudes of six components were considered as criterion variables. R, correlation coefficient of the model; R^2, adjusted R square; Beta, standardized regression coefficient of each predictor (Beta1 for the mean of single-trial peak amplitude and Beta2 for the SD of single-trial peak latency).

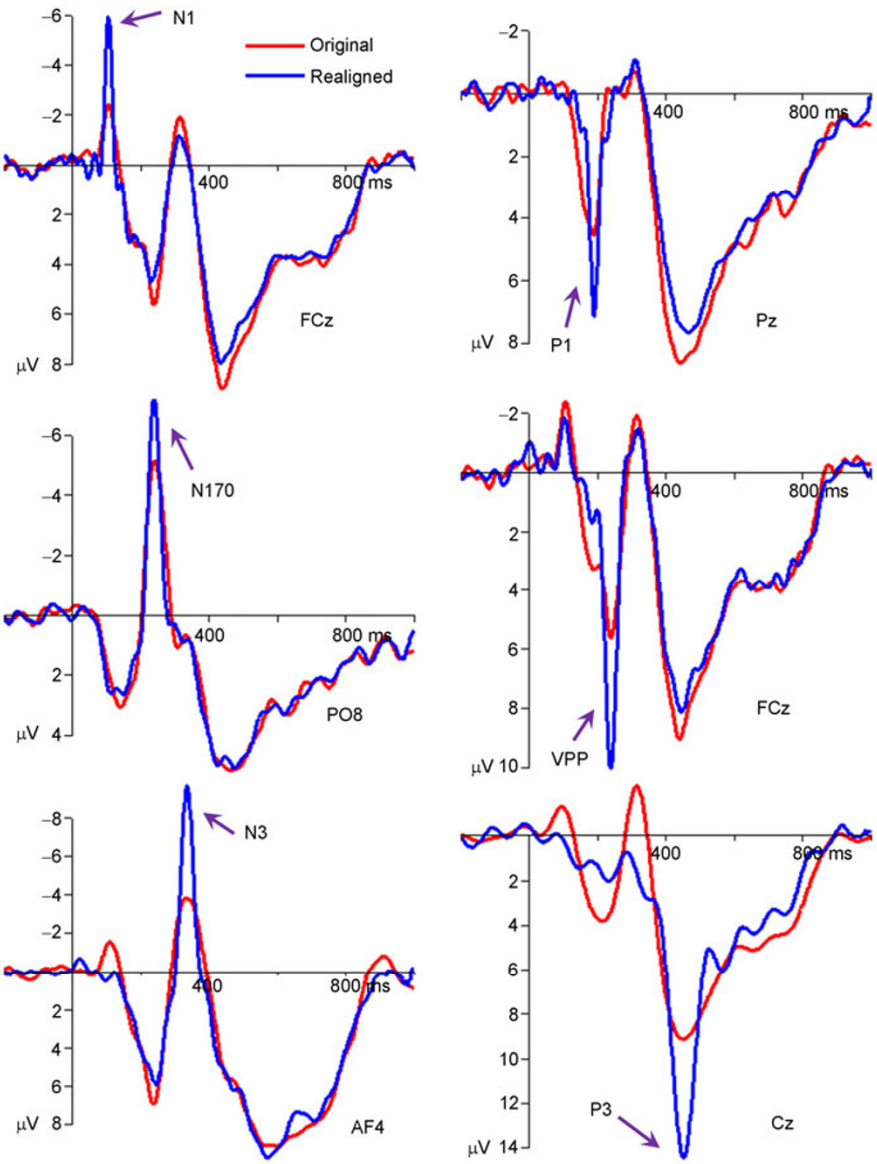

Figure 6 A comparison of the averaged ERP components before and after latency realignment. For the sake of brevity, only the ERP data in fearful condition are shown.

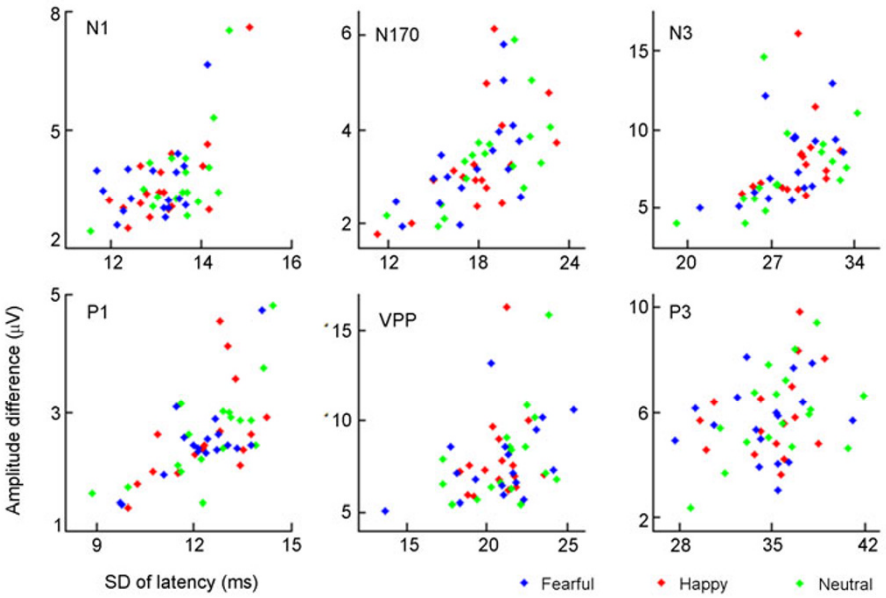

Figure 7 Scatterplots of the relationship between the SD of single-trial latency and the difference in component amplitude between original and realigned average ERPs.

the plane in Figure 8 (the slope in the direction of the mean single-trial amplitude was larger than in the direction of the SD for the single-trial latency).

3 Discussion

3.1 Single-trial characteristics and potential neural mechanisms underlying average amplitude differences

As illustrated in Figure 1, the average ERP is only a gross representation of neural activity, while single-trial measurements reveal more detailed information about dynamic brain function. In the current study, we investigated six discriminating components that has been identified and shown to be associated with facial expressions in our previous work [11]. The analyses of latency and amplitude were carried out on single-trial level to clarify whether the amplitude differences between experimental conditions results from differences in the real variability of single-trial amplitudes or from latency jitter [15,17,18,22].

The ANOVA results of the single-trial peak estimates indicate that nearly all of the average amplitude differences are attributable to true amplitude variations in single trials between experimental conditions (Table 1). We further verified these results by applying separate multiple linear regression analyses to the six components (Table 2). These data demonstrate that the latency jitter, or variation, across trials contributes little to the average amplitude differences.

When detecting facial expressions with deficient attention resources (as in the case of the RSVP experiment), the brain may assign different proportions of neural network function to the processing of different categories of facial expressions according to their emotional importance for survival and social interaction [18]. The data recorded in the time intervals of P1, N170, VPP, N3 and P3 indicate that the processing of fearful faces was particularly emphasized, as the fearful stimuli were always associated with the highest single-trial amplitudes. Given our assumption in Introduction section, these data show that the brain is likely to consider emotional expressions more important than neutral ones, preferentially allocating more neural resources to deal with happy faces, which results in larger single-trial amplitudes in response to happy expression than to neutral ones.

However, it is important to note that, while P1, N170, VPP, N3, and P3 show similar single-trial characteristics, the N1 component was an exception; a significant difference was detected in the latency SD between fearful and neutral conditions (Table 1). In light of these results, we suggest that the N1 component recorded in response to fearful faces may be a combined signal from dual neural pathways [20]. The first pathway is located in the cortex and commits a high level of neural network function to the processing of fearful expressions, reflected by the larger single-trial amplitudes in response to fearful faces. The second neural pathway includes subcortical areas, and the neural activity elicited by fearful faces is amplified as it first passes through the amygdala and reaches prefrontal cortex [26–29]. It has been shown that amygdala activation in response to fearful facial expressions is nearly automatic and needs little attention resources [6]. Therefore, the neural processing speed of fearful faces should be quite stable in the second pathway, resulting in a low peak latency variability compared with happy and neutral faces. Instead of the dual neural pathways of fearful face processing, there may

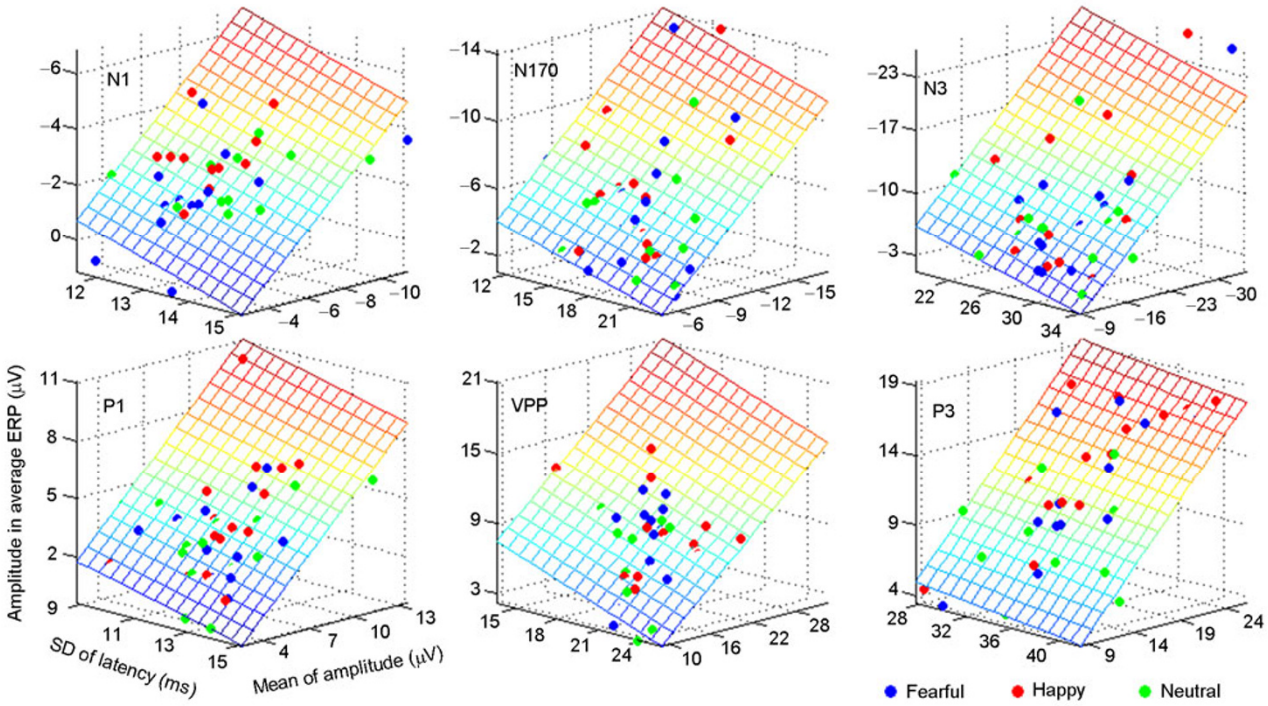

Figure 8 Scatter plot and wireframe mesh plot of average peak amplitudes against the SD of the single-trial peak latency and the mean of the single-trial peak amplitude. The observation count for each regression analysis was 51 (17 participants×3 conditions). Only observations above the wireframe mesh are shown.

be another explanation for the less variability of the N1 latency in fearful condition. Considering that the attentional engagement of participants usually waxes and wanes during EEG recording [18], fearful faces, compared with other kinds of facial expressions, may elicit less attentional fluctuations and produce more consistent single-trial latencies at the early stage of facial expression processing. To decide which explanation is more appropriate, further researches using single-neuron or neural imaging techniques are needed.

In addition, the two linear regression models for N1 component explain 60% and 67%, respectively, of the variation in average amplitude. These percentages are smaller than those found by the regression models associated with the other five components (Table 2). Moreover, the relation between emotional condition and single-trial properties (amplitude and latency) was relatively weak for the N1 compared to other components (η_P^2 =0.143 for N1 amplitude, and η_P^2 >0.3 for the other five components), suggesting that the N1 component may contain less discriminant information on facial expression at a single-trial level. These statistical results indicated that coarse and rapid characteristics of facial expression processing occur during the time interval of the N1 component. The precision may be sacrificed to rapidly obtain a rough categorization of the emotional stimuli [30–32].

Taken together, the findings from the current study have clarified the single-trial characteristics of the six ERP components (Figure 9). These data demonstrate that the amplitude differences of the P1, N170, VPP, N3, and P3 components are robust (i.e., there are consistently amplitude differences in single-trial data) and that both the latency jitter and amplitude differences contribute to the variations of N1 amplitude in average ERPs.

3.2 Converging evidence for the three-stage processing of facial expressions

We have previously drawn a conclusion in Luo et al. [11], which was perhaps premature, that the facial expression processing revealed by an analysis of average ERP could be separated into three temporal stages. The Luo et al. [11] findings, together with the single-trial analysis presented in the current study, provide converging evidence for the following three-stage scheme of facial expression processing.

3.2.1 Stage 1: Fear popup

The process of discriminating between fearful and other facial expressions occurs with the highest priority, with a fast processing speed and a relatively low categorization precision. On the average ERP level, several studies have shown that early components such as N1 and P1 have an increased amplitude in response to negative, particularly fearful faces, as early as 80 ms post-presentation [20,30, 33–36]; thus, the ERP results at this stage reflect a negativ-

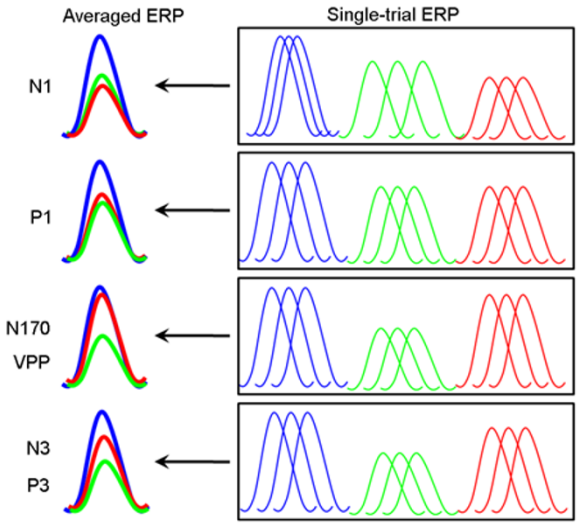

Figure 9 Single-trial characteristics underlying the average amplitude differences in the six ERP components. The average amplitude differences in N1 were due to both the single-trial amplitude differences and the different amounts of latency jitter. The average amplitude differences of P1, N170, VPP, N3, and P3 were simply due to true amplitude differences in single trials among emotional stimuli. ERP components are all shown as upward waveforms for display purpose. Blue, red, and green lines indicate the ERPs in response to fearful, happy, and neutral faces respectively.

ity bias [37,38]. On the single-trial ERP level, we find that an analysis of the fronto-central N1 component indicates that fearful expressions may be processed by parallel pathways located in the cortical and amygdalo-cortical areas. These dual neural pathways are reflected by a larger single-trial N1 amplitude (cortical) and a lower latency variation (amygdalo-cortical) in response to fearful faces. Unlike the N1 component, the P1 component, recorded from the parieto-occipital cortex, is likely elicited purely by cortical function, with larger single-trial amplitudes for fearful expressions.

3.2.2 Stage 2: Emotional/unemotional discrimination

After the brain has processed the fearful faces, it may focus on other emotional facial expressions if attentional resources are limited. During this stage, the perceived detail of facial expressions is only sufficient to distinguish emotional faces from unemotional ones. It has been shown that the average amplitudes of the N170, VPP, and other related ERP components are able to differentiate emotional facial expressions from neutral expressions, with larger amplitudes for emotional stimuli from 150 ms post-presentation [2,5,10,20,36,39–42]. Single-trial ERPs create distinguishing information at this stage by displaying larger amplitudes in response to emotional stimuli, such as fearful and happy faces.

3.2.3 Stage 3: Complete separation

The brain further evaluates the fine-grained information related to the affective valence of a face. In this third stage, the brain is finally able to distinguish among various categories of emotional faces. This elaborate processing of facial emotions is reflected by the separate average amplitudes of N3 and P3 components among different categories of facial expressions from approximately 300 ms post-presentation [20,33,39,40,43]. On the single-trial level, the ERPs elicited by fearful, happy, and neutral expressions show isolated amplitudes in the intervals representing N3 and P3, with similar amounts of latency jitter among emotional conditions.

Many psychophysiological studies using average ERP analyses and single-neuron recording techniques have revealed that brain activity is modulated in particular temporal patterns following the presentation of emotional faces [6,7,41]. In particular, Schyns's group [44,45] investigated the dynamics of the sensitivity of the EEG to facial discriminant features. They found that facial information was integrated from the eyes downward to the expression-specific facial parts (e.g., wide-opened eyes for fear and smiling mouth for happy) [44]. In this study, we observed a sequence from coarse to fine processing of facial expression information, which is consistent with the eye-to-diagnostic-feature scanning dynamics found by Schyns et al. [44]. Moreover, the finding of Schyns et al. [44] that the brain activity had minimal sensitivity to stimulus information after the N170 latency did not contradict our results: Schyns et al. [44] only focused on the brain responses on the occipitotemporal electrodes over the N170 time course while we analyzed the ERP data sensitive to facial expressions on the whole scalp. Finally, although this study demonstrated that the six ERP components were crucial for facial expression recognition, we never meant that the periods of time between component peaks contained no information processing. In contrast, we believe the brain works in a continuous manner, as revealed in Schyns et al. [44].

3.3 Technical and other issues

Interpreting the differential amplitude of the average ERP between two conditions is not straightforward, as the differences can be related to many underlying neural events, such as event-related brain activity evoking and phase reorganization (including phase resetting) [46–49]. To simplify the physiological interpretation, we describe three potential mechanisms leading to average amplitude difference (Figure 1), with the assumption that the ERP components are evoked by the stimulus and superimposed onto background electrophysiological activity unrelated to the stimulation. This "evoked model" is the most prevalent framework for ERP origination, and has been shown to be applicable to face-related ERP components [17,47,48].

Another issue that should be noted is that peak-to-peak amplitudes were used instead of absolute amplitudes to measure the N170, VPP, N3, and P3 ERP components in this study. There is plenty of evidence that ERP components

are not independent and may be correlated with each other. For example, Kuefner et al. [50] found that the amplitudes/topographies of P1 and N170 were correlated across ages (from four years to adulthood). We showed previously that significant amplitude differences between conditions existed in the ERP components prior to the N170, VPP, N3 and P3 [11]. Thus, peak-to-peak amplitudes should be used to isolate the amplitude contribution of the focused component.

In the current study, the peak amplitude and peak latency of single trials were treated as normally distributed variables. Importantly, normal distribution may not be a precise model for peak latency in some cases [13,22]. Fortunately, we only focused on the latency differences of certain ERP component among three emotional conditions, therefore the normal distribution may not affect the results, and can provide a simplified measurement for the readers' understanding [19]. Moreover, in future study it might be worthwhile to denoise the single-trial ERP data using the wavelet method [51] before peak detection so as to enhance the accuracy of detection.

Finally, and less importantly, we showed in Figure 5 that the amplitude distribution of N170 was more concentrated than that of VPP, which may argue for the view that N170 and VPP engage in dissociable neural networks [5,52]. However, since the trial-to-trial variability in amplitude, as well as the absolute value of single-trial peak latency, does not affect the average amplitude of ERP components [15], for the sake of clarity, we did not investigate the SD of single-trial amplitude and the mean of single-trial latency in the current study.

3.4 Concluding remarks

Single-trial analysis was applied to a RSVP experiment in order to examine the temporal characteristics of the six ERP components that had previously been shown to be associated with facial expression. The result is twofold. First, the fronto-central N1 component showed differences in both the latency jitter and single-trial amplitudes, which suggests that the N1 elicited by fearful faces is produced by parallel neural pathways. Secondly, a substantial proportion of the average amplitude differences in P1, N170, VPP, N3, and P3 may be accounted for by the pure amplitude variability on a single-trial basis, which is likely due to the levels of neural resources allocated to different categories of emotional faces. These single-trial findings, together with our previous average ERP analysis, strengthen our confidence in the three-stage scheme of facial expression processing, which was first proposed in Luo et al. [11] and consists of the "fear popup," "emotional/unemotional discrimination", and "complete separation" expression processing stages. Therefore, for the first time in facial expression processing, we provide a neurophysiological correlate, based on average and single-trial ERP dynamics, of the comprehensive temporal evolution of emotional face perception and discrimination.

The authors thank Rolf Verleger and Hannah Rosquist for their generously providing the Matlab code of maximum likelihood estimation for data analysis. This work was supported by the National Natural Science Foundation of China (31170984, 31000503, 91132704) and National Basic Research Program of China (2011CB711000).

1 Haxby J V, Hoffman E A, GobbiniM I. The distributed human neural system for face perception. Trends Cogn Sci, 2000, 4: 223–233
2 Adolphs R. Recognizing emotion from facial expressions: psychological and neurological mechanisms. Behav Cogn Neurosci Rev, 2002: 1, 21–62
3 Calder A J, Young A W. Understanding the recognition of facial identity and facial expression. Nat Rev Neurosc, 2005, 6: 641–651
4 Esslen M, Pascual-Marqui R D, Hell D, et al. Brain areas and time course of emotional processing. Neuroimage, 2004, 21: 1189–1203
5 Eimer M, Holmes A. Event-related brain potential correlates of emotional face processing. Neuropsychologia, 2007, 45: 15–31
6 Palermo R, Rhodes G. Are you always on my mind? A review of how face perception and attention interact. Neuropsychologia, 2007, 45: 75–92
7 Utama N P, Takemoto A, Koike Y, et al. Phased processing of facial emotion: an ERP study. Neurosci Res, 2009, 64: 30–40
8 Pessoa L, McKenna M, Gutierrez E, et al. Neural processing of emotional faces requires attention. Proc Natl Acad Sci USA, 2002, 99: 11458–11463
9 Schupp H T, Stockburger J, Codispoti M, et al. Selective visual attention to emotion. J Neurosci, 2007, 27: 1082–1089
10 Flaisch T, Junghöfer M, Bradley M M, et al. Rapid picture processing: affective primes and targets. Psychophysiology, 2008, 45: 1–10
11 Luo W, Feng W, He W, et al. Three stages of facial expression processing: ERP study with rapid serial visual presentation. Neuroimage, 2010, 49: 1857–1867
12 Jung T P, Makeig S, Westerfield M, et al. Analysis and visualization of single-trial event-related potentials. Hum Brain Mapp, 2001, 14: 166–185
13 Limpiti T, Van Veen B D, Wakai, R T. A spatiotemporal framework for MEG/EEG evoked response amplitude and latency variability estimation. IEEE Trans Biomed Eng, 2010, 57: 616–625
14 Blankertz B, Lemm S, Treder M, et al. Single-trial analysis and classification of ERP components—a tutorial. Neuroimage, 2011, 56: 814–825
15 Unsal A, Segalowitz S J. Sources of P300 attenuation after head injury: single-trial amplitude, latency jitter, and EEG power. Psychophysiology, 1995, 32: 249–256
16 Walhovd K B, Rosquist H, Fjell A M. P300 amplitude age reductions are not caused by latency jitter. Psychophysiology, 2008, 45: 545–553
17 Rousselet G A, Husk J S, Bennett P J, et al. Single-trial EEG dynamics of object and face visual processing. Neuroimage, 2007, 36: 843–862
18 Ford J M, White P, Lim K O, et al. Schizophrenics have fewer and smaller P300s: a single-trial analysis. Biol Psychia, 1994, 35: 96–103
19 Spencer K M. Averaging, detection, and classification of single-trial ERPs. In: Handy T C, ed. Event-Related Potentials: a Method Handbook. London: The MIT Press, 2005. 209–227
20 Williams L M, Palmer D, Liddell B J, et al. The 'when' and 'where' of perceiving signals of threat versus non-threat. Neuroimage, 2006, 31: 458–467
21 Jaśkowski P, Verleger R. Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method. IEEE Trans Biomed Eng, 1999, 46: 987–993
22 Jaśkowski P, Verleger R. An evaluation of methods for single-trial estimation of P3 latency. Psychophysiology, 2000, 37: 153–162

23 Pham D T, Möcks J, Köhler W, et al. Variable latencies of noisy signals: estimation and testing in brain potential data. Biometrika, 1987, 74: 525–533
24 Möcks J, Köhler W, Gasser T, et al. Novel approaches to the problem of latency jitter. Psychophysiology, 1988, 25: 217–226
25 Rousselet G A, Pernet C R, Bennett P J, et al. Parametric study of EEG sensitivity to phase noise during face processing. BMC Neurosci, 2008, 9: 1–22
26 Adolphs R, Tranel D, Damasio H, et al. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature, 1994, 372: 613–614
27 Adolphs R, Tranel D, Damasio H, et al. Fear and the human amygdala. J Neurosci, 1995, 15: 5879–5891
28 LeDoux J. Fear and the brain: where have we been, and where are we going? Biol Psychiatry, 1998, 44: 1229–1238
29 Rotshtein P, Richardson M P, Winston J S, et al. Amygdala damage affects event-related potentials for fearful faces at specific time windows. Hum Brain Mapp, 2010, 31: 1089–1105
30 Kawasaki H, Kaufman O, Damasio H, et al. Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. Nat Neurosci, 2001, 4: 15–16
31 Meeren H K, van Heijnsbergen C C, de Gelder B. Rapid perceptual integration of facial expression and emotional body language. Proc Natl Acad Sci USA, 2005, 102: 16518–16523
32 Vuilleumier P, Pourtois G. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. Neuropsychologia, 2007, 45: 174–194
33 Eimer M, Holmes A. An ERP study on the time course of emotional face processing. Neuroreport, 2002, 13: 427–431
34 Eger E, Jedynak A, Iwaki T, et al. Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. Neuropsychologia, 2003, 41: 808–817
35 Pourtois G, Grandjean D, Sander D, et al. Electrophysiological correlates of rapid spatial orienting towards fearful faces. Cereb Cortex, 2004, 14: 619–633
36 Pourtois G, Dan E S, Grandjean D, et al. Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. Hum Brain Mapp, 2005, 26: 65–79
37 Ito T A, Larsen J T, Smith N K, et al. Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. J Pers Soc Psychol, 1998, 75: 887–900
38 Cacioppo J T, Gardner W L. Emotion. Annu Rev Psychol, 1999, 50: 191–214
39 Krolak-Salmon P, Fischer C, Vighetto A, et al. Processing of facial emotional expression: spatio-temporal data as assessed by scalp event-related potentials. Eur J Neurosci, 2001, 13: 987–994
40 Campanella S, Quinet P, Bruyer R, et al. Categorical perception of happiness and fear facial expressions: an ERP study. J Cogn Neurosci, 2002, 14: 210–227
41 Batty M, Taylor M J. Early processing of the six basic facial emotional expressions. Cognitive Brain Res, 2003, 17: 613–620
42 Jiang Y, Shannon R W, Vizueta N, et al. Dynamics of processing invisible faces in the brain: automatic neural encoding of facial expression information. Neuroimage, 2009, 44: 1171–1177
43 Carretié L, Iglesias J. An ERP study on the specificity of facial expression processing. Int J Psychophysiol, 1995, 19: 183–192
44 Schyns P G, Petro L S, Smith M L. Dynamics of visual information integration in the brain for categorizing facial expressions. Curr Biol, 2007, 17: 1580–1585
45 Smith M L, Fries P, Gosselin F, et al. Inverse mapping the neuronal substrates of face categorizations. Cereb Cortex, 2009, 19: 2428–2438
46 Makeig S, Debener S, Onton J, et al. Mining event-related brain dynamics. Trends Cogn Sci, 2004, 8: 204–210
47 Klimesch W, Sauseng P, Hanslmayr S, et al. Event-related phase reorganization may explain evoked neural dynamics. Neurosci Biobehav R, 2007, 31: 1003–1016
48 Sauseng P, Klimesch W, Gruber W R, et al. Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. Neuroscience, 2007, 146: 1435–1444
49 Rossion B, Jacques C. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. Neuroimage, 2008, 39: 1959–1979
50 Kuefner D, de Heering A, Jacques C, et al. Early visually evoked electrophysiological responses over the human brain (P1, N170) show stable patterns of face-sensitivity from 4 years to adulthood. Front Hum Neurosci, 2010, 3: 67
51 Quian Quiroga R. Obtaining single stimulus evoked potentials with wavelet denoising. Physica D, 2010, 145: 278–292
52 Botzel K, Schulze S, Stodieck S R. Scalp topography and analysis of intracranial sources of face-evoked potentials. Exp Brain Res, 1995, 104: 135–143