A gender stereotype can be described as culturally specific, publicly shared characteristics that are associated with one particular gender [15]. In Chinese traditional culture, the stereotypes of gender are rather pervasive: the role of female is always associated with housekeeping, while that of male is more associated with social production. Although individuals may attempt to prevent gender stereotypes from affecting their behavior by engaging in controlled processes, the intentions to control gender bias are not always successful. Even for the self-avowed egalitarians, the prejudices of gender have often slipped through in their behavior despite their non-prejudiced intentions. Previous studies revealed that stereotyping is an automatic, implicit, cognitive process, and it can be automatically activated by gender features [22]. The stimulus of gender features can automatically start an activated gender stereotype [3]. When the activated stereotypic information conflicts with one's non-prejudiced beliefs, the following neural processes will result.

In cognitive neuroscience, there are two separate neural systems that work in concert to arrive at an intended behavior in the face of conflicting behavioral tendencies. The first one is a conflict-detection system, which monitors ongoing responses and is sensitive to competition between different response tendencies [4,19]. This system is constantly active, requiring few cognitive resources, and has been shown to operate below conscious awareness. When a conflict is detected by this system, it alerts the second, resource-dependent regulatory system designed to implement the intended response while inhibiting the unintended response [9].

The error-related negativity (ERN) wave, a component of the event-related potential (ERP), reflects the response-performance monitoring process [23]. The onset of the ERN coincides with response initiation, as determined by the onset of the electromyogram (EMG) associated with the responding hand, and peaks roughly 80 ms thereafter. Its spatial distribution lies over the frontal–central regions of the scalp [14], and is symmetrical to the midline [23]. The anterior cingulate cortex, which is specifically involved in the brain’s error processing system, seems to be the generator site of the ERN [13].

In past studies, it was demonstrated that the ERN can be elicited by conflicts and errors. Response conflict monitoring theory of ERN suggests that the amplitude of ERN reflects the degree of response conflict. In higher-conflict incongruent trials, larger ERN amplitudes are produced [23]. From reinforcement learning theory, the ERN is sensitive to the degree of error, as larger errors correspond with larger negativity [14,18]. The ERN is elicited whenever it detects a mismatch between the produced response and the correct, or the intended, response [8], such as the wrong hand, the wrong finger, or both the wrong hand and the wrong finger [5]. Recently, some studies correlate the ERN with some modulating factors. Confronted with the same error, the ERN peaks both earlier and larger in the context of pleasant backgrounds than neutral or unpleasant backgrounds [16]. The impulsivity of participant is correlated with the ERN amplitude, which suggests that
the high-impulsive participants have smaller ERN amplitudes on punishment trials than the low-impulsive participants [12]. Social culture, such as racial stereotypes, can also induce larger ERNs [9]. Few studies however, have focused on how the gender stereotypes affect the ERN.

In this study, the participants were asked to identify the target stimulus which was a tool or a piece of kitchenware, and the prime stimulus was a male or female face. We speculated that the prime stimulus would activate the gender stereotypes which would influence the judgment of the target stimulus. The ERN would be recorded on these trials that the participants’ identification of the target stimulus was wrong, and larger ERNs would be observed on the trials of gender stereotype-incongruent.

Twenty-eight right-handed male undergraduates aged between 20 and 35 years (mean = 27.5) were employed in this study. All had normal or corrected-to-normal visual acuity. They did not have any history of neurological or mental diseases. All the participants claimed holding predominantly non-prejudiced attitudes of gender.¹

The stimuli consisted of 144 prime-target pairs. The primes in these pairs included 12 digital photographs of 6 Asian male and 6 Asian female faces, and the targets included 12 digital photographs of 6 kitchen utensils (fork, knife, spatula, rolling pin, ladle, chopsticks) and 6 hand tools (drill, ratchet, wrench, pliers, brush, shovel). These photographs were digitized at 228 × 172 pixels.

Stimuli were presented sequentially in the center of a computer screen, with a visual angle of 2.58° × 2.4°. Each trial began with a pattern mask (1 s), followed by the prime (200 ms), and then the target (200 ms). The next trial starts after the participant responded or after 1500 ms had elapsed since the onset of the target. The prime-target pairs were divided into four conditions: male-kitchenware, female-kitchenware, male-tool, and female-tool. The stimulus pairs of male-tool and female-kitchenware are called stereotype-congruent pairs and those of male-kitchenware and female-tool are called stereotype-incongruent pairs. The stimulus pairs were randomly presented in sequence and had the same probability. Stimuli and recording triggers were presented using STIM 2 software (Neurosoft Labs, Inc., Sterling, USA).

After participants provided informed consent, they were fitted with an electrode cap for electroencephalographic (EEG) recording. Then, the participants were instructed to classify the target picture as a piece of kitchenware or a tool as quickly as possible, and were told that an erroneous “kitchenware” response on a female-tool trial or an erroneous “tool” response on a male-kitchenware trial was indicative of gender bias.

Electroencephalogram was continuously recorded (band pass 0.05–100 Hz, sampling rate 500 Hz) with Neuroscan Synamp2 Amplifier (Scan 4.3.1, Neurosoft Labs, Inc., Sterling, USA), using an electrode cap with 64 Ag/AgCl electrodes mounted according to the extended international 10–20 system and referenced to linked mastoids. Vertical and horizontal electrooculograms were recorded with two pairs of electrodes, one placed above and below the right eye, and the other 10 mm from the lateral canthi. Electrode impedance was maintained below 5 kΩ throughout the experiment. Following electrode application, participants sat in a comfortable sofa located in a shielded room and were asked to fix a point in the center of the computer display. Next, the experimenter gave participants instructions to classify the second picture as a kitchenware or a tool by pressing the corresponding key with their right or left index finger. Responses were to be made within 500 ms of the target presentation. If the responses exceeded this time limit, the participants would receive a red warning symbol “×” to make them respond more quickly. After 30 practice trials, 576 stimulus trials were presented.

An 800 ms epoch of EEG signal, centered on key press, was selected for each artifact-free trial with the first 400 ms of the epoch as a baseline. Electrooculogram artifacts were corrected using the method proposed by Semlitsch et al. [21]. Trials contaminated by amplifier clipping, bursts of electromyographic activity, or peak-to-peak deflection exceeding ±80 μV were excluded from averaging. The remaining trials were baseline corrected. ERPs for correct and incorrect trials were averaged within their respective trial types and the averaged ERPs were digitally filtered with a low pass filter at 30 Hz (24 dB/Octave). The ERNs were scored, in this study, when the peak negative deflection occurred between 50 ms before key press and 150 ms after key press at the frontocentral site.

In each analysis, only participants with valid response on all measures were adopted. To ensure reliable and stable ERP component with relative higher signal-to-noise ratio extracting from digitized EEG signal, it was necessary to exclude the participants with insufficient sample of valid trials. Fabiani et al. [10] suggested that more than 20 valid trials for a condition are needed to obtain the stable ERP. In this study, we excluded the participants with fewer than 20 artifact-free error trials in any of the four conditions (7 participants). Thus, the primary analyses were conducted on the data from the rest (21 participants).

The data of means (Ms) and corresponding standard deviations (S.D.s) of reaction times (RTs) on correct responses and of response error rates in different conditions are given in Table 1. To examine whether the reaction times on stereotype-congruent trials are faster than on stereotype-incongruent trials, we conducted a 2 (gender: male vs. female) × 2 (object: tool vs. kitchenware) within-subjects analysis of variance (ANOVA) for RTs on correct responses. This analysis produced a main effect for gender, F(1,20) = 22.52, p = 0.001; responses to male faces (M = 363.033, S.D. = 77.659), F(1,20) = 10.766, p = 0.004, and were slower to identify tools following female faces (M = 379.72, S.D. = 72.22) compared with male faces (M = 344.887, S.D. = 72.934), F(1,20) = 22.688, p = 0.000. These results demonstrated an automatic association between the male face and tools and between the female face and kitchenware. Additionally, the slower responses to the pairings of male-kitchenware and female-tool suggested that participants adopted a more controlled response strategy on these trials in an effort to avoid gender bias.

In addition, to examine the effect of stereotypic associations on the response error rates, we conducted a 2 (gender) × 2 (object) within-subjects ANOVA for error rates. There was no main effect for gender, F(1,20) = 1.433, p = 0.245, nor object, F(1,20) = 0.234, ¹ All these participants were asked to fulfill a questionnaire that was created by authors following the idea of Burt's sex role stereotyping scales [7].

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>RT for correct response (ms)</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male-T</td>
<td>344.88 (72.93)</td>
<td>0.3084 (0.1365)</td>
</tr>
<tr>
<td>Male-K</td>
<td>363.03 (77.66)</td>
<td>0.4361 (0.1688)</td>
</tr>
<tr>
<td>Female-T</td>
<td>379.72 (72.22)</td>
<td>0.4353 (0.1593)</td>
</tr>
<tr>
<td>Female-K</td>
<td>345.92 (76.01)</td>
<td>0.3373 (0.1590)</td>
</tr>
</tbody>
</table>


Q. Ma et al. / Neuroscience Letters 442 (2008) 186–189 187
The interaction between gender and object was significant, $F(1,20) = 13.313, p = 0.002$. Simple effect analyses expressed that participants made a higher error rate on tool trials when the target was preceded by a female face ($M = 0.435, S.D. = 0.159$) compared with a male face ($M = 0.308, S.D. = 0.136$), $F(1,20) = 14.379, p = 0.001$, and on kitchenware trials they made higher error rates when the target was preceded by a male face ($M = 0.436, S.D. = 0.169$) compared with a female face ($M = 0.337, S.D. = 0.159$), $F(1,20) = 9.120, p = 0.007$. These results indicated a prepotent tendency to favor a “kitchenware” response after viewing a female face and a “tool” response after viewing a male face.

On the basis of previous findings that the ERN is relatively focal over frontocentral locations [11] and the scalp distribution maps of the present data, the electrode site $Fz$ was chosen for analysis. The data of ERN amplitudes at the scalp site of $Fz$ in different conditions are given in Fig. 1. We conducted 2 (gender: male vs. female) $\times$ 2 (object: tool vs. kitchenware) $\times$ 2 (response: correct vs. erroneous) within-subject repeated measure ANOVA for the ERN peak amplitude to examine the effect of the ERN evoked by the activation of gender stereotype. The ANOVA produced a marginally significant main effect for gender of face, $F(1,20) = 3.655, p = 0.070$, the mean difference of amplitudes between cues of male and female faces was $-0.581 \mu V$. A main effect emerged for response, $F(1,20) = 16.735, p = 0.001$; the mean difference of amplitudes between erroneous and correct responses was $-3.112 \mu V$. There was no significant main effect for object, $F(1,20) = 1.030, p = 0.322$. These effects were qualified by a marginally significant interaction of gender $\times$ response, $F(1,20) = 4.217, p = 0.053$, and a significant interaction of gender $\times$ object $\times$ response, $F(1,20) = 5.374, p = 0.031$, whereas both the interaction of gender $\times$ object and of object $\times$ response were not significant, $F(1,20) = 0.005, p = 0.937, F(1,20) = 1.307, p = 0.266$, respectively.

To further understand the significant interaction of gender $\times$ response, a simple effect analysis was conducted, showing that the ERN amplitudes on male error trials ($M = -1.601, S.D. = 3.417$) were significantly larger (negative-polarity) than that on female error trials ($M = -0.508, S.D. = 4.156$), $t(41) = -2.555, p = 0.014$, while the ERN amplitudes on male correct trials were not significantly different from that on female correct trials, $t(41) = -0.193, p = 0.848$.

To further understand the significant interaction of gender $\times$ object $\times$ response, we conducted the analyses of pairwise comparisons, and obtained the following results. It was not significant that the mean difference of the ERN amplitudes of the correct responses on tool trials when the target was preceded by a male face compared with a female face, $t(20) = -1.659, p = 0.113$, and on kitchenware trials when the target was preceded by a male face, compared with a female face, $t(20) = 0.982, p = 0.338$. The ERN amplitudes were significantly larger when kitchen utensils were erroneously classified as tools following male faces ($M = -2.226, S.D. = 3.163$) than when the same errors were made following female faces ($M = -0.467, S.D. = 3.677$), $t(20) = -3.745, p = 0.002$, whereas the ERNs associated with erroneously classifying tools as kitchen utensils following male and female faces did not differ statistically, $t(20) = -0.637, p = 0.532$. The significantly larger ERN observed for errors on male-kitchenware trials suggests that greater conflict is detected when a response tendency is gender biased than when it is not biased.

This study demonstrates that ERNs are obviously elicited in erroneous responses to four types of prime-target conditions: male-tool, female-tool, male-kitchenware, and female-kitchenware. Statistical test shows that the mean amplitude of ERNs of male-kitchenware errors is significantly larger than that of female-kitchenware errors, whereas the mean amplitude of ERNs of male-tool errors does not differ significantly from that of female-tool errors.

The ERN has been linked with response conflict, reinforcement learning or response monitoring (see [14], for a review). As the ERN reflects a response-conflict monitoring process, it can be elicited by negative feedback and by error commission [14], and modulated by modulating factors, such as affectiveness or impulsiveness [12,16]. The amplitude of ERN reflects the degree of response conflict or response error [14,18,23]. In this study, the ERNs that were elicited in the task of forced-alternative identification response have two sources—“pure” operation errors and the conflict between the gender stereotype activation and the non-prejudice beliefs. The numbers of “pure” hand manipulation errors on male-kitchenware and female-kitchenware trials are random variables which should have the same mean for any participant. Therefore, the finding that the mean amplitude of ERNs on erroneous response trials of male-kitchenware is significantly larger than that of female-kitchenware is mainly a result of the activated stereotypes of gender.

Previous studies demonstrated that stereotypes and unconscious prejudices could be automatically activated when one identifies another person’s skin color, facial features, gender characteristics, and the like [17]. Borgida et al. suggested that gender is a fundamental dimension of categorization. Once an individual is categorized as belonging to one gender, the stereotypes of that gender may quickly come to the perceiver’s mind [6]. From the view of dual system theory, there may be two explanations for

\[ A \text{ “pure” operation error by hand means the participant wants to press the correct key but pressed the incorrect key by using a wrong finger. This error is different from the error that participant pressed a key he wants to press.} \]
stereotypic affect. One is that the first system does not appraise stereotype-biased response, and therefore it does not signal the second system. The alternative explanation is that the second system does not implement the control after it received the signal of need of control [9]. Some race bias studies suggested that controlled, belief-based processes are more effectively implemented in deliberate responses [1]. While deliberate responses, like self-reports, do not invariably reflect attitudes and beliefs, particularly when the expression of the belief is perceived to be socially inappropriate (see [20], for a review). In this study, we infer that the second system has no time to implement the control, the unconscious, erroneous response had submitted already when the signal of need of control being delivered from the first system to the second.

The unintended stereotypic activation may be the source of the difference of ERN amplitude. In this identification task, the participants’ gender stereotype is activated, and the activated stereotypes have a strong conflict with the participants’ intended response (the gender-unbiased response). The results supported our hypothesis. Some researches suggest that the ERN represents a conflict monitoring function in regulating race-biased behavior [2,9]. Therefore, gender stereotypes in the task may be associated with response-performance monitoring. It may imply that the gender stereotypes, as a subjective factor, may modulate the ERN. The ERPs are sensitive to violations of gender-based occupational stereotypes even when subjects judged a stereotype violation to be acceptable [20]. The present findings prove that the ERPs might be advantageous tools for studying automatic activation of gender stereotypes. The present study focuses on the gender-role stereotype of males. It will be extended to females and do some comparisons between them in the future.

Taken together, the ERN that is elicited in this experiment may come from an error commission and modulated strongly by a gender stereotypic factor, which suggest that the activation of gender stereotype may be indexed by the ERN.

Conflict of interest

None.

Acknowledgements

We would like to appreciate the advice of two anonymous reviewers. This work was supported by grant No. 08BJL054 from the National Social Science Foundation of China.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neulet.2008.06.080.

References