



The role of right insula and its functional connectivity in the regulation of negative implicit stereotypes against rural migrant workers

Lei Jia^{1,2} · Billy Sung³ · Jun Wang^{1,2}

Accepted: 31 January 2022 / Published online: 14 February 2022
© The Psychonomic Society, Inc. 2022

Abstract

Previous studies have revealed that stereotyping processes involving negative affective content (e.g., antipathy) can lead to a significant increase in insula activity. However, whether the insula is sensitive to stereotype inconsistency or plays a crucial role in stereotype regulation remains unclear. To help fill this gap, 21 young adults were presented with a modified single-category implicit association test (SC-IAT) that assessed their stereotypes about rural migrant workers. In a within-subjects design, participants completed separate blocks of compatible and incompatible trials, each of which consisted of stereotype-consistent or stereotype-inconsistent trait labels that had a positive or a negative valence. Functional MRI was used to identify specific brain regions associated with negative and positive stereotyping. The behavioral results revealed a typical stereotype regulation effect in which participants responded slower to stereotype-inconsistent condition than stereotype-consistent condition, although such effect was significantly modulated by IAT compatibility, rather than by emotional valence. MRI results revealed that activity in the right insula was significantly sensitive to stereotype regulation processes in negative incompatible tasks, whereas such effect was marginally significant in positive incompatible tasks. Moreover, psychophysiological interaction analyses indicated complex functional connectivity among the right insula and cognitive control regions [e.g., ventrolateral prefrontal cortex (vlPFC) and dorsal anterior cingulate cortex (dACC)], social mentalizing regions [medial prefrontal cortex (mPFC)], and motivation regions (putamen) in the condition where negative stereotypes were violated. Taken together, these findings demonstrate that the right insula serves as a crucial node in regulating implicit stereotyping, particularly in negative stereotyping tasks.

Keywords Stereotyping · Rural migrant workers · Insula · fMRI · Conflict control

Introduction

In the current social psychology literature, stereotypes are typically defined as beliefs or associations regarding a group's attributes that include traits or circumstantial characteristics (Amodio, 2014; Hilton & von Hippel, 1996).

Such stereotyping processes can lead to an affective reaction toward people in other social groups (Allport, 1954; Link and Phelan, 2001; Rosenthal & Overstreet, 2016). Previous models of stereotyping have distinguished between stereotype activation and stereotype application (Brewer, 1988; Fiske & Neuberg, 1990; Gilbert & Hixon, 1991). In these models, *stereotype activation* (i.e., categorization elicited by activated stereotypical beliefs and affects) is regarded as a first and necessary step that is followed by *stereotype application* in which the perceiver decides whether to endorse the activated stereotypical beliefs (*stereotype endorsement*) or regulate them (*stereotype regulation*) through cognitive control (Blair & Banaji, 1996). Based on such models, many well-controlled experimental paradigms, such as the implicit association test (IAT), priming tasks, and go/no-go tasks, have been developed to test implicit stereotyping processes (Blair, 2002; Blair & Banaji, 1996; Hilton & von Hippel, 1996).

✉ Jun Wang
jun.wang@zjnu.edu.cn

¹ Department of Psychology, Zhejiang Normal University, 688 Yingbin Rd, Jinhua, Zhejiang Province 321004, People's Republic of China

² Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, 321004 Jinhua, People's Republic of China

³ Consumer Research Lab, School of Management and Marketing, Curtin University, 208 Kent Street, Bentley WA6102, Australia

Over the past two decades, social neuroscientists have identified a set of brain regions involved in stereotyping processes, including the amygdala, insula, temporoparietal junction (TPJ), medial prefrontal cortex (mPFC), lateral prefrontal cortex (particularly the inferior frontal gyrus or IFG), and anterior and lateral temporal lobe (the ATL and LTL). Previous studies on racial and gender stereotypes have examined the functions of those brain regions in well-controlled experimental paradigms (for reviews, see Amodio, 2014; Quadflieg & Macrae, 2011). In these brain regions, the bilateral amygdala and insula are usually associated with negative emotions, such as threat, antipathy, or disgust, which are related to the subjective experience of negative affect in racial stereotyping (Chekroud et al., 2014; Knutson et al., 2007; Lieberman et al., 2005). Additionally, the mPFC and TPJ of the core social mentalizing system, which are commonly involved in the processing of social beliefs, theory of mind (ToM), trait inference, and perspective taking, are engaged in distinguishing ingroup and outgroup members (Contreras et al., 2011; Delplanque et al., 2019; Quadflieg et al., 2008; Van der Cruyssen et al., 2015; Van Overwalle & Baetens, 2009). The IFG, ATL, and LTL, which are crucial brain regions involved in the semantic storage and retrieval of social knowledge, are usually engaged in stereotype representation and its semantic activation (Amodio, 2014; Gallate et al., 2011; Wong et al., 2011). All of these brain regions work together to form a basic stereotyping network that integrates semantic and affective processes. Furthermore, when stereotypes need to be regulated in response to prosocial or egalitarian norms, the dorsolateral prefrontal cortex (dlPFC) and posterior medial frontal cortex (pmFC; including the dorsal anterior cingulate cortex or dACC) of the domain-general conflict monitoring network activates stereotype regulation and affects the stereotyping network (Amodio, 2014; Cattaneo et al., 2011). Based on this evidence, it is reasonable to predict that the functional connectivity between semantic representation regions (i.e., IFG, ATL, and LTL), core mentalizing regions (i.e., mPFC and TPJ), and emotion/affect regions (i.e., amygdala and insula) support the manifestation of stereotypical beliefs. Such integrated stereotype network is then modulated by the conflict monitoring system (e.g., dlPFC & dACC) (Amodio, 2014; Quadflieg & Macrae, 2011).

Although many previous studies have supported this postulated stereotyping neural network (Amodio, 2014; Quadflieg & Macrae, 2011), three research questions remain unanswered. First, because most of the relevant work has been conducted in Western cultures only, its cultural generalizability has not been established. According to the *stereotype content model* (SCM; Cuddy et al., 2008, 2009), different stereotype contents can lead to different stereotyping processes, and it is known that culture can impact neural responses related to stereotyping (Cuddy et al., 2008;

Hughes et al., 2019). Second, although many previous studies of stereotyping have found conjoint brain activation in these semantic, mentalizing, emotional, and conflict control regions, there is still a lack of direct empirical evidence of brain connectivity to verify the postulated network (Amodio, 2014). Third, it is not clear whether the insula has a modulation role in regulating negative stereotyping, in addition to its role in negative emotion/affect input (Craig, 2009).

The insula is a large cortical region that runs from the medial to the temporal lobes, adjacent to the frontal cortex, and broadly functions to represent antipathy and disgust. Although the insula has rarely been the focus of interest in previous neuroimaging studies of stereotypes and prejudices, its activity is frequently associated with social categorizations of racial ingroup and outgroup members (Knutson et al., 2007; Lieberman et al., 2005). In previous research, insula activity in racial stereotyping tasks has been interpreted as a reflection of a negative visceral reaction, such as antipathy towards racial outgroups (Amodio, 2014; Harris & Fiske, 2006; Knutson et al., 2007). According to this interpretation, the insula contributes to the subjective affect that is experienced as part of a stereotyping response. Hence, representation of the negative affective response in the insula may, through its connections with the ACC and PFC, facilitate the ability to detect and regulate behavior on the basis of a negative affective response (Amodio, 2014). Nevertheless, this interpretation could be challenged by the fact that the insula is associated with consciousness (or awareness) (Craig, 2009). Indeed, previous research on cognitive control and performance monitoring has reported strong activation in the insula that was not only associated with negative antipathy but also linked to the conflict monitoring and control (Cole & Schneider, 2007; Craig, 2009). Given that insula plays a role in conflict monitoring and control, which serve as critical components (or processes) in stereotype regulation (Amodio, 2014), it is reasonable to predict that the insula also might play an important role in stereotype regulation processes. However, few studies have tested this prediction.

To examine the postulated stereotyping-related neural network (Amodio, 2014; Quadflieg & Macrae, 2011) and test the cognitive role of insula in stereotyping, the present research focused on a different kind of stereotype with obvious antipathy emotions in Chinese culture. Specifically, we examined the processing of stereotypes in relation to *rural migrant workers* (RMWs) in a Chinese social context. The role of the insula and its neural network was assessed by fMRI as participants completed a modified *single-category implicit association test* (SC-IAT; Karpinski & Steinman, 2006).

RMWs are a social group in China that emerged from a household-registration system known as the *hukou* system. To manage and control citizens' movements, this system divides citizens into two groups: urban and rural residents. A

citizen's birth location determines which group they fall into and which residency permit (i.e., which *hukou*) they receive. Before the 1990s, rural residents were commonly prevented from living in cities if they did not possess a city *hukou*, which made rural-urban economic and educational inequalities salient (Wu & Treiman, 2004). However, because of rapid industrialisation and urbanisation during the past 40 years in China, an increasing number of rural residents with a low level of education have migrated to metropolitan cities in search of work. Typically, these RMWs tend to be hired in labor-intensive, nonagricultural jobs that city residents are unwilling to do. Because of traditional social hierarchical views that are widely held, RMWs often are perceived to be coarse and unintelligent, and they elicit feelings of antipathy from many people (Kuang & Liu, 2012; Wang et al., 2011). For instance, Wang et al. (2011) found that stereotypes against RMWs are widely held by Chinese urban college students and that these stereotypes can be measured using IAT paradigms.

The present study was conducted following a 2 (stereotype consistency: stereotype-consistent (SC) vs. stereotype-inconsistent (SI)) \times 2 (IAT compatibility: compatible task vs. incompatible task) \times 2 (emotional valence: negative vs. positive) within-subjects experimental design. Using a SC-IAT paradigm, we investigated whether the insula plays a crucial role in regulating stereotypes against RMWs. In our version of the SC-IAT paradigm, we presented *stereotype-consistent and stereotype-inconsistent* (SC and SI) traits associated with RMWs as stimuli. According to prior studies on stereotyping (Brewer, 1988; Fiske & Neuberg, 1990; Gilbert & Hixon, 1991), people tend to show slower responses to SI trials than to SC trials, because SI trials engage stereotype violation and trigger additional stereotype regulation processes (e.g., conflict monitoring and control). Additionally, there is an IAT compatibility effect (Karpinski & Steinman, 2006) in which participants in the SC-IAT usually show faster responses in compatible conditions than in incompatible conditions, because the strength of the semantic association between social status and attitude-consistent traits in compatible conditions leads to less cognitive conflict and does not involve the same level of regulation as in incompatible conditions (Greenwald et al., 1998; Karpinski & Steinman, 2006).

The present study was designed to explore the role of brain activation and functional connectivity in the insula cortex during the processes of stereotype regulation. We tested two hypotheses about the roles of the insula in stereotyping against RMWs. The first hypothesis relates to the cognitive roles of the insula in stereotyping processes. Because the insula is associated with not only the feelings of antipathy but also conflict control (Cole & Schneider, 2007; Craig, 2009), we predict that there would be significant insula activity with reference to stereotypes against RMWs. In particular, we predict that the activity of the insula would be

modulated by both stereotype consistency and IAT compatibility in the SC-IAT when stereotype violation is engaged and stereotype regulation processes (e.g., conflict monitoring and control) are triggered, especially in negative stereotype contexts. If this hypothesis is correct, when stereotype regulation processes are triggered, participants should show lower responses to stereotype-inconsistent (SI) trials than stereotype-consistent (SC) trials. Due to the influences from stereotype regulation and task compatibility, insula activity should produce a stereotype consistency \times IAT compatibility interaction (or a 3-way interaction between stereotype consistency, IAT compatibility, and emotional valence) and make the $SI > SC$ contrast stronger in the incompatible task than in the compatible task under negative valence.

Our second hypothesis relates to the functional connectivity of the insula in stereotyping regulation. If the insula is an important part of the stereotype regulation network, then the functional connectivity of the insula in stereotyping should differ in the compatible and incompatible tasks due to the potential interaction between stereotype consistency and IAT compatibility. Based on the stereotyping neural network previously observed in literature (Amodio, 2014), we also predicted that the insula might have functional connectivity, not only with semantic and mentalizing regions (e.g., MPFC, TPJ, and temporal lobe), but also with conflict control regions (e.g., dIPFC & ACC).

Methods

Subjects

This research was approved by the research ethics committees of Southwest University and Zhejiang Normal University. A total of 21 undergraduates (19–24 years; 9 males) from Southwest University were recruited through printed advertisements. Similar to what was done by Wang et al. (2011), all participants in the present study had possessed a permanent city resident identity (*hukou*) since middle school. They were right-handed and reported having no neurological disorders, significant physical illness, head injury, or alcohol/drug abuse. After the experiment, they were debriefed and given appropriate financial compensation for participating. We have disclosed all measures, manipulations, and exclusions. In particular, we completed the data collection before we began conducting any analyses.

Materials

The stimuli for the SC-IAT included 10 neutral Chinese social status nouns that describe rural migrant workers (e.g., male migrant worker, female migrant worker, and migrant manual labourer), as well as 20 stereotype-consistent (10

positive and 10 negative) and 20 stereotype-inconsistent (10 positive and 10 negative) traits. All status nouns and stereotype-consistent traits were chosen from previous research on negative stereotypes against rural migrant workers in China (Wang et al., 2011; Zhen & Liu, 2012). The stereotype-inconsistent traits were simply the antonyms of the stereotype-consistent traits. All traits are shown in Table 1. In the experiment, they were presented in black, 48-point Song font on a black screen.

Procedure

Before the experiment began, written consent was obtained from all of the participants. Participants were informed that they would do a word-categorization task in which they would see some status nouns that describe rural migrant workers, and also some trait words. They were informed that their task was to categorize these words as quickly as possible.

An illustration of timing of each trial is presented in Fig. 1. Each trial began with either a status noun that refers to RMWs or a trait and ended with a blank screen. The status noun or trait was presented for 2 seconds, and then a blank screen appeared for either 2, 4, or 6 seconds.

The formal SC-IAT experiment consisted of two sessions: one for the compatible task, and one for the incompatible task. Each session had 150 trials, including 30 trials with status nouns that refer to rural migrant workers, and 30×4 [2 (emotional valence: positive vs. negative) \times 2 (stereotype consistency: SC vs. SI)] trials with trait words. Each status noun or trait was repeated three times in both the compatible task and the incompatible task. During the compatible task, participants were told to press the “1” key for status nouns of rural migrant workers and negative traits and the “2” key for positive traits. In the incompatible tasks, they pressed the “1” key for status nouns and positive traits and the “2” key for negative traits. The sequence of compatible and incompatible tasks was reversed for half of the participants. Participants were allowed to have a short rest (less than 5 minutes) between these tasks.

On each trial, participants provided a response to the target (i.e., the status noun or trait) within the given duration. After a response was recorded, the target remained on the screen for the entire 2 seconds duration of the trial. Before

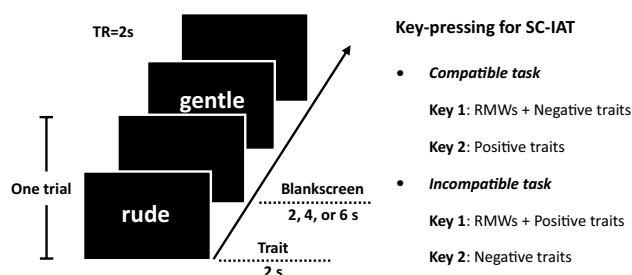


Fig. 1 Illustration of the SC-IAT, including what was displayed (left) and the key-pressing instructions key-for the compatible and incompatible tasks (right). For both compatible task and incompatible task, there were four types of trials (or four conditions): the positive SC trials, the positive SI trials, the negative SC trials, and the negative SI trials. However, only stereotype-inconsistent (SI) trials (i.e., positive SI trials and negative SI trials) would engage stereotype-violation, which should trigger the stereotype regulation processes (e.g., conflict monitoring and control)

the experiment began, each participant was allowed to complete a series of 12 practice trials. After that, they were taken to the MRI scanner and prepared for scanning.

Image acquisition and pre-processing

Participants were scanned on a Siemens Allegra 3.0 Tesla scanner (Siemens Medical Systems, Erlangen, Germany) at the brain-imaging centre of Southwest University. For each participant, functional images (including 336 time points in SC-IAT) were acquired using a spin-echo EPI sequence with the following parameters: 32 axial slices, repetition time (TR) = 2,000 ms, echo time (TE) = 30 ms, flip angle (FA) = 90°, field of view (FOV) = 220 \times 220 mm².

Image pre-processing was conducted using the SPM12 software package (Wellcome Department of Cognitive Neurology, London, UK), which was run on the Matlab software platform (MathWorks Ltd.). Functional data were time-corrected for differences in acquisition time between slices and realigned within and across runs to correct for inter-scan head motions. They were then spatially normalised in a standard EPI template and resampled to 3 \times 3 \times 3 mm³. Afterwards, normalized data were spatially smoothed using a Gaussian kernel (8-mm full width, half maximum).

Table 1 Stereotype consistent and inconsistent traits used in the present study

| | |
|---|--|
| Negative stereotype consistent traits | Dirty, conservative, rude, shortsighted, grubby, stingy, inferior, clumsy, stupid, barbaric |
| Positive stereotype consistent traits | Sincere, modesty, friendly, hard-working, honest, cordial, persevering, straightforward, kindhearted, diligent |
| Negative stereotype inconsistent traits | Impassive, cunning, unfriendly, lazy, insincere, haughty, cowardly, sinister, contemptible, indolent |
| Positive stereotype inconsistent traits | Tidy, stylish, polite, wise, gentle, elegant, noble, smart, intelligent, civilized |

Whole-brain and ROI analysis

To identify all the significant brain regions activated in the processing of stereotypes, whole-brain analysis was employed. Each participant's (1st level analysis) pre-processed images were statistically analysed using the general linear model (GLM), in which the two-run event-related design was modelled using the canonical haemodynamic response function and its temporal derivative. A total of ten conditions, including eight experimental conditions [i.e., 2 (stereotype consistency: SC vs. SI) \times 2 (IAT compatibility: compatible vs. incompatible) \times 2 (emotional valence: negative vs. positive)], one condition for blank screen, and one special condition for error/failed trials] were included in the GLM. In addition, six scan-to-scan motion parameters that were produced during realignment were also entered into the model as regressors, to account for the residual effects of movement. Afterwards, a 128-s cutoff, high-pass filter was used to reduce low frequency noise and slow drifts.

For the whole-brain analyses (2nd level analysis), comparisons of interest were analysed in a flexible factorial design model in which participants were treated as a random effect. Statistical comparisons between conditions were conducted using analysis of variance (ANOVA) procedures on the parameter estimates associated with each condition. Given our prediction, we focused our attention on contrasts relating to predicted stereotype regulation processes (i.e., main effects and interactions relating to the $SI > SC$ contrasts involved). To detect all significant relevant brain regions areas, a voxel-based statistical threshold of $p < 0.05$ (corrected for small volumes using FDR) was used for all comparisons with a minimum cluster extent of 10 voxels ($K > 10$, without any mask). However, to make sure that somewhat less strong activation would not be missed, statistical results under a less stringent cluster-wise correction ($p < 0.005$ uncorrected, $K > 30$ voxels, without any mask) also were recorded (Forman et al., 1995; Lieberman & Cunningham, 2009; Ma et al., 2012).

In addition, regions of interest (ROIs) were determined for the brain regions identified in previous studies. Based on both previous findings and our research aims, we selected the following the ROIs:

- (1) Social mentalizing regions: mPFC (dmPFC: MNI $x-y-z$: 0 50 35; vmPFC: 0 50 5) and TPJ ($\pm 50 -55 45$; Ma et al., 2012; Van Overwalle, 2009);
- (2) Conflict control regions: posterior medial frontal cortex (pmFC; 0 20 45) and lateral PFC (lPFC; $\pm 40 25 20$), with all ROI masks for these regions based on a sphere with a radius of 15 mm (Ma et al., 2012; Van Overwalle, 2009); and
- (3) Emotion and arousal related regions: for the bilateral insula and amygdala, ROIs were created based on their

bilateral automated anatomical labelling (AAL) structure using the WFU_PickAtlas toolbox (ANSIR Laboratory, Wake Forest University School of Medicine; <http://www.ansir.wfubmc.edu>). Mean percentage signal changes in each ROI under each experimental condition were extrapolated using the MarsBar toolbox (<http://marsbar.sourceforge.net>) and analysed using a 2 (stereotype consistency: SC vs. SI) \times 2 (IAT compatibility: compatible vs. incompatible) \times 2 (emotional valence: negative vs. positive) repeated-measures ANOVA with a threshold of $p < 0.05$.

Psychophysiological interaction analysis

To identify and contrast the functional connectivity of the insula in the compatible and incompatible tasks, a generalized form of context-dependent psychophysiological interaction (gPPI) analysis was employed. The gPPI analysis tests the hypothesis that activity in one brain area can be explained by an interaction between the presence of a cognitive process, and activity in another part of the brain (Cisler et al., 2014; McLaren et al., 2012). In the present study, if the brain activation in the insula was significant in whole brain analysis and ROI analysis, the anatomical labelling (AAL) structure of the insula would be used as a seed region to examine how the functional connectivity of the insula in the negative (and/or positive) stereotype regulation processes (e.g., $SI > SC$ contrast) differed in the compatible and incompatible tasks. If so, the individual contrast images in the compatible and incompatible tasks of the SC-IAT were collected and entered into a group-level analysis random effects analysis, in which brain activations associated with stereotype consistency were investigated using one-sample t -tests. The cluster-wise correction methods that were used in the group-level analysis were the same as in the whole-brain analysis.

Results

Behavioral results

We conducted a 2 (stereotype consistency: SC vs. SI) \times 2 (IAT compatibility: compatible vs. incompatible) \times 2 (emotional valence: negative vs. positive) repeated-measure ANOVA on participants' behavioral data, and the results are shown in Fig. 2. An ANOVA for response accuracy showed a significant stereotype consistency \times IAT compatibility interaction ($F_{(1,20)} = 5.70$, $p = 0.027$, $\eta_p^2 = 0.22$), but there were no significant main effects of stereotype consistency ($F_{(1,20)} = 1.62$, $p = 0.022$, $\eta_p^2 = 0.08$), IAT compatibility ($F_{(1,20)} = 3.00$, $p = 0.10$, $\eta_p^2 = 0.13$), or emotional valence ($F_{(1,20)} = 2.78$, $p = 0.11$, $\eta_p^2 = 0.12$). Post-hoc analysis of the

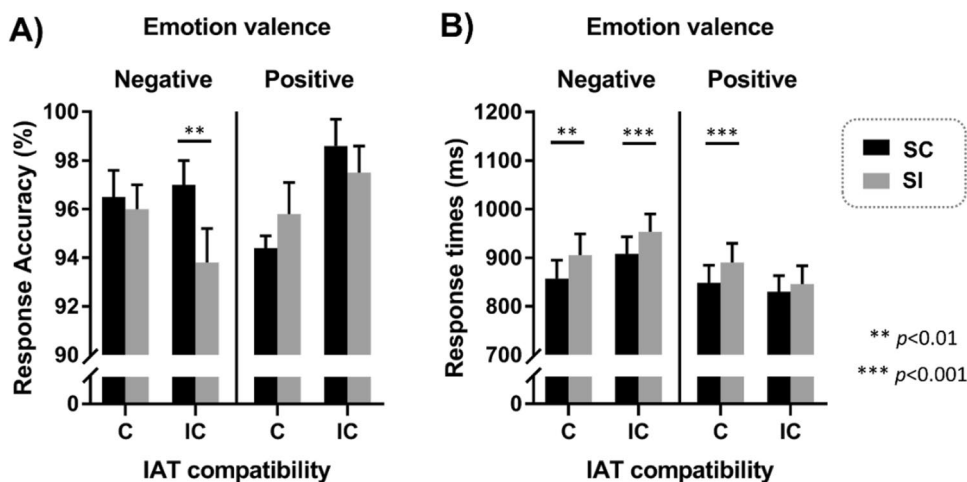


Fig. 2 Descriptive results ($M \pm SE$) of participants' behavioral responses in the SC-IAT. Significant effects in contrasts between SC and SI conditions are indicated by asterisks ($*p < 0.05$; $**p < 0.01$;

$***p < 0.001$) based on t -tests corrected by multiple comparisons using the Bonferroni method

interaction showed that participants performed with higher response accuracy in SC trials than in SI trials in the incompatible task ($M_{SC} = 97.90\%$, $SE = 0.65\%$; $M_{SI} = 95.64\%$, $SE = 1.17\%$; $F_{(1,20)} = 4.69$, $p = 0.043$, $\eta_p^2 = 0.19$), but this effect disappeared in the compatible task ($M_{SC} = 95.52\%$, $SE = 0.84\%$; $M_{SI} = 95.93\%$, $SE = 0.89\%$; $F_{(1,20)} = 0.27$, $p = 0.61$, $\eta_p^2 = 0.01$).

The ANOVA for response times (RTs) showed significant main effects of stereotype consistency ($F_{(1,20)} = 21.25$, $p < 0.001$, $\eta_p^2 = 0.52$) and emotional valence ($F_{(1,20)} = 26.39$, $p < 0.001$, $\eta_p^2 = 0.57$), but the main effect of IAT compatibility was not significant ($F_{(1,20)} = 0.15$, $p = 0.70$, $\eta_p^2 < 0.01$). Post-hoc analysis showed that participants had slower responses to the SI trials ($M_{SI} = 898.74$ ms, $SE = 32.95$ ms) than the SC trials ($M_{SC} = 859.76$ ms, $SE = 36.99$ ms) and also had slower responses to negative traits ($M_{\text{negative}} = 906.12$ ms, $SE = 35.72$ ms) than to positive traits ($M_{\text{positive}} = 852.49$ ms, $SE = 34.60$ ms). We also found a significant IAT compatibility \times emotional valence interaction ($F_{(1,20)} = 21.74$, $p < 0.001$, $\eta_p^2 = 0.53$). Simple-effect analysis on this interaction showed that participants had slower responses to the negative traits than to the positive traits in the incompatible task ($M_{\text{Negative}} = 930.78$ ms, $SE = 35.37$ ms; $M_{\text{Positive}} = 835.60$ ms, $SE = 34.16$ ms; $F_{(1,20)} = 75.61$, $p < 0.001$, $\eta_p^2 = 0.79$), but this effect disappeared in the compatible task ($M_{\text{Negative}} = 881.25$ ms, $SE = 40.06$ ms; $M_{\text{Positive}} = 869.37$ ms, $SE = 37.49$ ms; $F_{(1,20)} = 0.57$, $p = 0.46$, $\eta_p^2 = 0.03$).

To test the predicted regulation effect for negative and positive stereotyping (i.e., in stereotype-inconsistent or SI trials) directly, in the post-hoc test we employed a t -test on RTs and compare the SC and SI conditions at each IAT compatibility and emotional valence level. The results showed

that participants had slower responses in SI trials than in SC trials respectively in the negative compatible tasks, the negative incompatible tasks, and the positive compatible tasks. Such effect was nullified in the positive compatible tasks (Fig. 2).

Whole-brain results

The results of the whole-brain analysis are shown in Fig. 3. For the conjunction analysis of stereotype regulation (i.e., the main effect of stereotype consistency), the $SC > SI$ contrast showed no significant brain activity at an FDR corrected statistical threshold of $p < 0.05$. However, the $SI > SC$ contrast showed significant activation in a set of brain regions, mainly including the thalamus (extending to putamen and caudate), bilateral cingulate cortex (anterior and posterior cingulate gyrus, ACC & PCC), bilateral insula (i.e., left anterior insula and right posterior insula), left TPJ, and bilateral prefrontal cortex areas, such as the dorsolateral, ventrolateral, and medial prefrontal cortex (dlPFC, vlPFC, and mPFC) (Fig. 3; see more details in *Supplementary Materials*).

We also observed brain activations which were sensitive to the *stereotype consistency* \times *IAT compatibility* interaction and *stereotype consistency* \times *emotional valence* interaction, respectively. Specifically, the contrast of *incompatible task > compatible task: SI > SC* revealed activations in a set of brain regions, mainly including the right dlPFC, and bilateral precuneus at an FDR corrected statistical threshold of $p < 0.05$, as well as the right putamen (extending to the right insula), MOG (extending to fusiform gyrus) and left cerebellum at a less stringent cluster-wise correction at $p < 0.005$ (uncorrected) with $K > 30$. In addition, the contrast

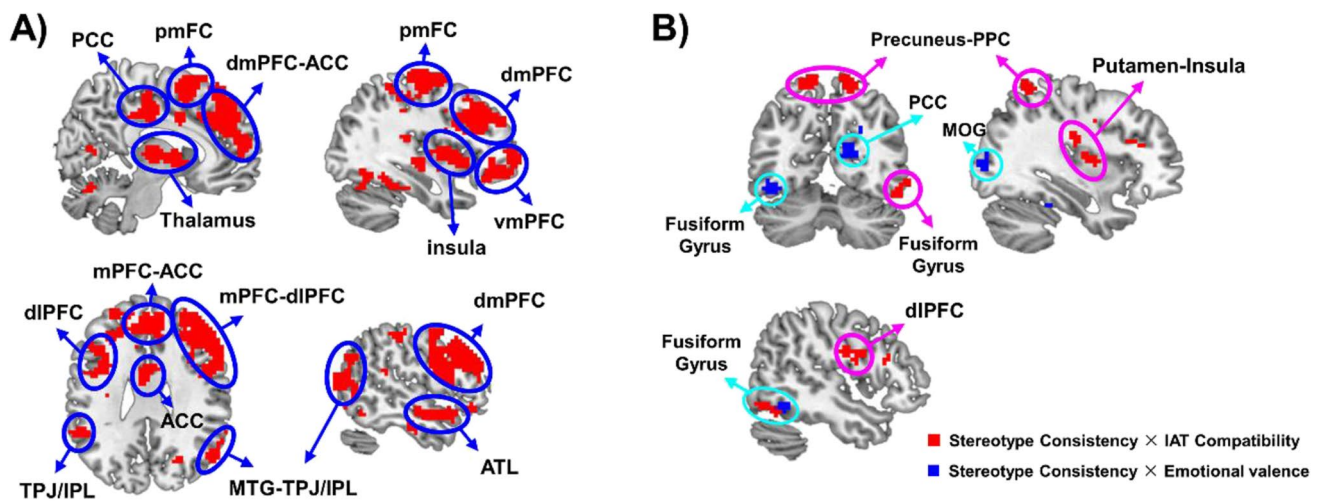


Fig. 3 Brain activity evoked by the main effect of stereotype consistency (i.e., conjunction analysis: $SI > SC$) (red regions in A), interaction between stereotype consistency and IAT compatibility (red regions in B), and interaction between stereotype consistency and

emotional valence (blue regions in B), respectively. Brain regions shown in this figure were identified using a less stringent cluster-wise correction at $p < 0.005$ (uncorrected) with $K > 30$.

of *positive > negative*: $SI > SC$ also revealed activations in a set of brain regions, mainly including the right fusiform gyrus and left cerebellum at an FDR corrected statistical threshold of $p < 0.05$, as well as the right PCC at a less stringent cluster-wise correction at $p < 0.005$ (uncorrected) with $K > 30$. However, we did not observe any brain activity that was sensitive to the three-way interaction between stereotype consistency \times IAT compatibility interaction \times emotional valence.

Previous research has indicated that there are different neural bases between the negative and positive stereotyping processes (Wraga et al., 2007). Although no brain activity was observed to the three-way interaction, we analyzed the brain activations associated to stereotype consistency \times IAT compatibility interaction respectively at positive and negative valences to test our first prediction and explore the stereotype regulations which might be impacted by IAT compatibility at different emotional valence.

As shown in Fig. 4, different brain regions were activated by stereotype consistency, IAT compatibility, and their interactions respectively in negative and positive valence conditions. For the conjunction analysis of stereotype consistency (i.e., the main effect of stereotype consistency) in positive valence condition, $SI > SC$ contrast showed significant brain activation mainly in the bilateral dIPFC, bilateral ACC (extending to pmFC), bilateral putamen, bilateral Cerebellum, right MTG, and left TPJ. However, no brain activation was found in the contrast of $SC > SI$, even at a less stringent cluster-wise correction. For conjunction analysis of the IAT compatibility (i.e., the main effect of IAT compatibility) in positive valence condition, *Incompatible task > Incompatible task* contrast showed significant brain activation in the

left dIPFC and right vmPFC. However, no brain activation was found in the *compatible task > incompatible task* contrast, even at a less stringent cluster-wise correction. In addition, there was a significant stereotype consistency \times IAT compatibility interaction in positive valence condition. This significant activation was revealed in the right ACC for the contrast of *incompatible task > compatible task*: $SI > SC$ only. No other significant effects were found (Fig. 4; Table 2).

Note: (1) SC = stereotype-consistent, SI = stereotype-inconsistent. (2) Brain activation reported in this table was set at a threshold of $p < 0.005$ (uncorrected) with extent voxel threshold > 30 ; (3) *Results thresholded at $p < 0.05$ corrected for multiple comparisons as determined by small-volume using FDR

Note: (1) SC = stereotype-consistent, SI = stereotype-inconsistent. (2) Brain activation reported here was set at a threshold of $p < 0.005$ (uncorrected) with extent voxel threshold > 30 . (3) *Results thresholded at $p < 0.05$ corrected for multiple comparisons, as determined by small-volume using FDR

For the conjunction analysis of stereotype consistency in negative valence condition, in the compatible task, the $SI > SC$ contrast showed no significant brain activation. However, in the incompatible task, the $SI > SC$ contrast revealed significant activation in the insula (mainly right posterior insula), TPJ, vlPFC, and mPFC. In addition, the conjunction analysis of the IAT compatibility in negative valence condition showed no significant brain activation at an FDR corrected statistical threshold of $p < 0.05$, although the left dIPFC were activated in the *Incompatible task > Compatible*

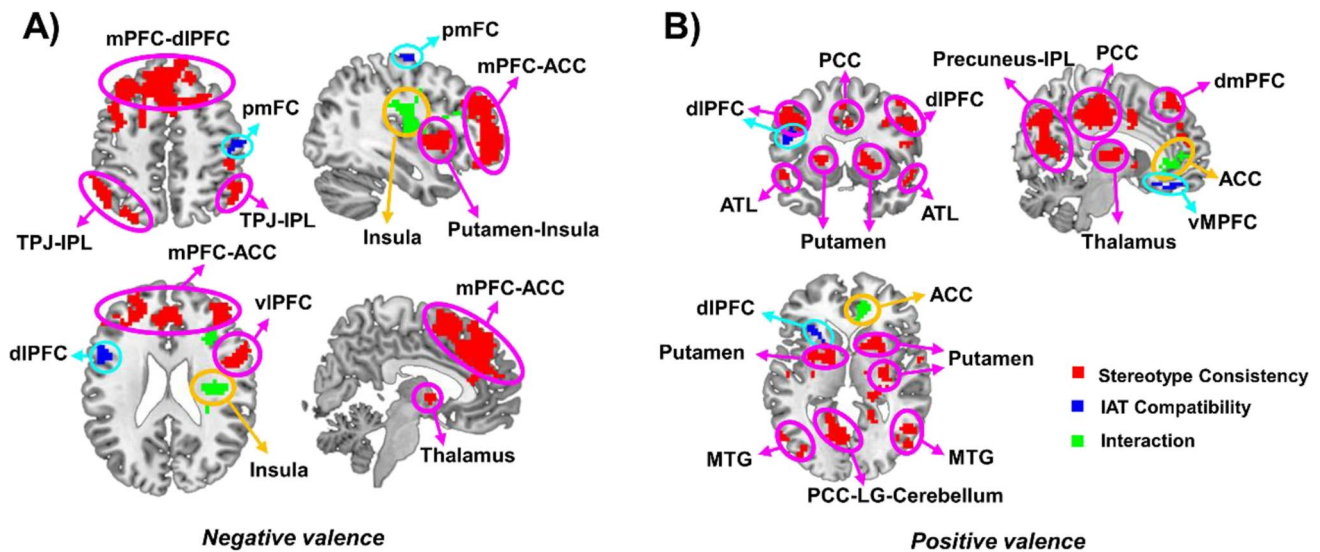


Fig. 4 Brain activity evoked by the main effects of stereotype consistency (i.e., conjunction analysis: $SI > SC$; marked in red color), main effects of the task compatibility of the SC-IAT (i.e., conjunction analysis: incompatible task > compatible task; marked in blue color),

task, at a less stringent cluster-wise correction (i.e., $p < 0.005$ uncorrected, $K > 30$). Nonetheless, we did find a significant stereotype consistency \times IAT compatibility interaction in negative valence condition. This significant activation was revealed in the right posterior insula for the contrast of *incompatible task > compatible task*: $SI > SC$ only. No other significant effects were found (Fig. 4; Table 3).

Signal change estimations on ROIs

Percentage signal change estimations on ROIs are shown in Fig. 5. For each ROI, we conducted a repeated-measures ANOVA with stereotype consistency (SC vs. SI), IAT compatibility (compatible vs. incompatible), and emotional valence (positive vs. negative) as factors. In line with the whole-brain results, the ROI analysis showed only a stereotype consistency \times IAT compatibility interaction on the right insula, even though it also revealed a main effect of stereotype consistency on other social mentalizing regions (i.e., left TPJ and dmPFC), conflict control regions (pmPC and bilateral IPFC), and emotion-related regions (bilateral amygdala and bilateral insula). Besides that, the ROI analysis also showed a stereotype consistency \times emotional valence interaction on right IPFC. All simple effects related to the stereotype regulation (i.e., $SI > SC$) in these main effects of stereotype consistency, stereotype consistency \times IAT compatibility interaction, and stereotype consistency \times emotional valence interaction are shown in Fig. 5. For the negative valence condition, we observed the $SI > SC$ effect on the bilateral TPJ and pmFC in the compatible task, but

and their interaction (marked in green color), respectively in negative valence (a) and positive valence (b). Brain regions shown in this figure were identified using a less stringent cluster-wise correction at $p < 0.005$ (uncorrected) with $K > 30$

observed the $SI > SC$ effect on the right insula, dmPFC, and right IPFC in the incompatible task. For the positive valence condition, we observed the $SI > SC$ effect on the left IPFC in the compatible task, but observed the $SI > SC$ effect on the bilateral amygdala in the incompatible task.

Functional connectivity of the insula

In the whole-brain and ROI results, only the right posterior insula was sensitive to the interaction between stereotype consistency and IAT compatibility in the negative valence condition. For that reason, in the gPPI analysis we only explored the functional connectivity of the right insula in the negative valence condition. These results are shown in Table 4 and Fig. 6.

Note: (1) Table shows the functional connectivity of the right insula, which was identified by the gPPI analysis. (2) Brain activation reported in this table was set at a threshold of $p < 0.005$ (uncorrected) with extent voxel threshold $K > 30$. (3) *Results thresholded at $p < 0.05$ corrected for multiple comparisons as determined by small volume using FDR

As predicted, the functional connectivity of the insula in the negative valence condition did indeed differ in the compatible and incompatible tasks. The gPPI analysis identified a complex functional connectivity between the insula, mPFC, dACC, vIPFC, MTG, and putamen in the incompatible task, but not in the compatible task.

Table 2 Peak voxel, *t*-value, and number of voxels (*k*) of the stereotype inconsistency > consistency contrasts in positive valence

| Regions activated | Hem | Peak MNI coordinate | | | <i>t</i> | <i>k</i> |
|--|------|---------------------|------------|------------|-------------|-------------|
| | | x | y | z | | |
| Compatible task: SI > SC | | | | | | |
| * rACC | R | 12 | 45 | 9 | 4.36 | 106 |
| Precuneus | R | 18 | −57 | 24 | 4.77 | 39 |
| Incompatible task: SI > SC | | | | | | |
| ATL | R | 54 | 9 | −15 | 4.00 | 76 |
| | L | −51 | 12 | −9 | 3.39 | 31 |
| | L | −39 | −48 | 15 | 3.95 | 51 |
| MTG | L | −69 | −27 | −3 | 5.22 | 39 |
| * Putamen, Caudate | R | 21 | 9 | 9 | 5.79 | 504 |
| | L | −21 | 12 | 3 | 5.24 | 336 |
| *dlPFC | R | 45 | 3 | 24 | 5.54 | 313 |
| | L | −39 | 6 | 30 | 5.95 | 236 |
| * PCC-IPL-Precuneus-dACC-pmFC | L | −9 | 0 | 36 | 6.70 | 1557 |
| * Cerebellum-Occipital-Fusiform | R | 42 | −51 | −18 | 7.99 | 828 |
| | L | −42 | −51 | −27 | 8.40 | 1393 |
| Conjunction analysis: SI > SC | | | | | | |
| *dlPFC | R | 48 | 30 | 21 | 4.99 | 413 |
| | L | −51 | 12 | 39 | 4.25 | 114 |
| *dACC-pmFC | L,R | −9 | −18 | 51 | 5.66 | 1203 |
| | R | 33 | −12 | 63 | 4.85 | 386 |
| *MTG | R | 48 | −9 | −15 | 4.10 | 81 |
| | L | 51 | 18 | −18 | 4.26 | 30 |
| * Putamen-Insula-Thalamus | R | 27 | −12 | 15 | 5.17 | 374 |
| * Putamen-Insula | L | −15 | 0 | 0 | 5.45 | 150 |
| Hippocampus | R | 24 | −27 | −6 | 3.93 | 31 |
| * TPJ/IPL | L | −45 | −45 | 15 | 5.15 | 108 |
| | R | 57 | −42 | 21 | 3.94 | 32 |
| * Cerebellum-Occipital-Fusiform | L, R | −39 | −45 | −27 | 5.73 | 1941 |
| Compatible task > Incompatible task: SI > SC | | | | | | |
| None | | | | | | |
| Incompatible task > Compatible task: SI > SC | | | | | | |
| * ACC | R | 12 | 42 | 0 | 4.63 | 73 |

Discussion

The present study investigated the essential role of the insula in the regulation of stereotypes against rural migrant workers (RMWs) in a Chinese cultural context. Using a SC-IAT paradigm, we aimed to elicit stereotype regulation processes in stereotype-inconsistent conditions (relative to stereotype-consistent conditions) in order to examine two hypotheses pertaining to the activity of the insula. As predicted, our behavioral results indicated that we had indeed elicited the stereotype regulation processes in stereotype-inconsistent (SI) trials relative to in stereotype-consistent (SC) trials, in both the negative and positive valence conditions. Specifically, the analysis of participants' response accuracy showed a significant stereotype consistency × IAT compatibility

interaction, which indicated that participants had more accurate responses on SC trials than on SI trials in the incompatible task. In contrast, RT analyses showed significant main effects of stereotype consistency and emotional valence, in which participants had slower responses to SI trials than to SC trials and slower responses to negative traits than to positive traits. For the stereotype regulation effects (i.e., SC and SI contrast), we found such effects in the negative compatible trials, the negative incompatible trials, and the positive compatible trials, but not in the positive incompatible trials. Taken together, these behavioral results provided clear evidence of stereotype regulation processes. Based on these results, we used fMRI to measure insula activity during the negative stereotype regulation task and observed four interesting findings.

Table 3 Peak voxel, *t*-value, and number of voxels (*k*) of the stereotype inconsistent > consistent (*SI>SC*) contrasts of the whole-brain analysis in negative valence

| Regions activated | Hem | Peak MNI coordinate | | | <i>t</i> | <i>k</i> |
|--|------|---------------------|-----|-----|----------|----------|
| | | x | y | z | | |
| Compatible task: <i>SI > SC</i> | | | | | | |
| dIPFC, dACC | R, L | −6 | 6 | 66 | 3.87 | 152 |
| Incompatible task: <i>SI > SC</i> | | | | | | |
| dIPFC | R | 45 | 54 | 6 | 4.89 | 31 |
| | L | −39 | 24 | 36 | 4.61 | 112 |
| rACC | L | −6 | 21 | −9 | 3.90 | 43 |
| *PCC, precuneus | R, L | 15 | −18 | 36 | 5.33 | 1271 |
| Amygdala | L | −24 | −6 | −12 | 4.44 | 43 |
| *vIPFC | R | 33 | 54 | 24 | 4.90 | 239 |
| *mPFC, dACC | L | −12 | 36 | 33 | 4.70 | 253 |
| Postcentral gyrus | R | 27 | −39 | 63 | 4.95 | 84 |
| *Insula | R | 36 | −15 | 21 | 5.34 | 386 |
| | R | 42 | 12 | −3 | 3.52 | 59 |
| Conjunction analysis: <i>SI > SC</i> | | | | | | |
| *dIPFC, mPFC, vIPFC | R, L | −39 | 15 | 42 | 5.14 | 2507 |
| *dIPFC, insula | R | 51 | 15 | 15 | 3.74 | 199 |
| *Insula | L | −42 | 12 | 6 | 3.76 | 51 |
| *TPJ/IPL (supramarginal, angular) | R | 54 | −51 | 48 | 3.64 | 85 |
| | L | −48 | −54 | 45 | 3.58 | 199 |
| Compatible task > Incompatible task: <i>SI > SC</i> | | | | | | |
| None | | | | | | |
| Incompatible task > Compatible task: <i>SI > SC</i> | | | | | | |
| * Insula | R | 36 | −6 | 12 | 4.98 | 443 |
| PCC | L | −24 | −15 | 33 | 3.91 | 130 |

First, we found that the right insula plays a conflict control role in regulating stereotyping processes, although our results indicated such a role was significant in the negative stereotype regulation context and maintained marginally significant in the positive stereotype regulation context. In addition to the emotion input role found in implicit stereotype within the Western culture (Knutson et al., 2007; Lieberman et al., 2005), the present study provided further empirical evidence from the Chinese culture that the right insula is sensitive to stereotype inconsistency though we didn't directly perform the cross-culture comparison. Specifically, in the whole-brain results we observed that significant bilateral insula (e.g., left anterior insula and right posterior insula) activity was sensitive to stereotype inconsistency (i.e., *conjunction analysis: SI>SC*), but only the right insula was sensitive to the stereotype consistency × IAT compatibility interaction. This interaction was further confirmed by analyses in which the AAL structures of the bilateral insula were used as ROIs. In whole-brain and ROI results, we observed that the right posterior insula was significant for the *SI>SC* contrast in the negative incompatible task, but not in negative compatible task, indicating that the right insula

was more sensitive to negative stereotypes in a stereotype-inconsistent (in other words, stereotype violation) condition than in a consistent condition. This finding supports our first hypothesis by showing that the right posterior insula is an important neural node in negative stereotyping regulation networks. Because similar finding also were observed for the right insula in positive tasks (at a marginally significant level), the right insula's association with stereotype regulation should not be specific for negative stereotypes, but also be suitable for positive stereotypes.

Traditionally, the anterior insula is thought to serve as an integrative center for the representation and evaluation of interoceptive information, while the posterior insula is thought to serve a major role in general awareness of the body, including visceral and somatosensory sensations (Amodio, 2014; Craig, 2009). However, recent studies have shown that the right posterior insula is also responsive to emotional stimuli (e.g., threat-related pictures such as snakes), reflecting an engagement self-awareness (i.e., arousal of consciousness) in emotion evaluation (Straube & Miltner, 2011). Indeed, our results concerning the right posterior insula are consistent with this finding because the

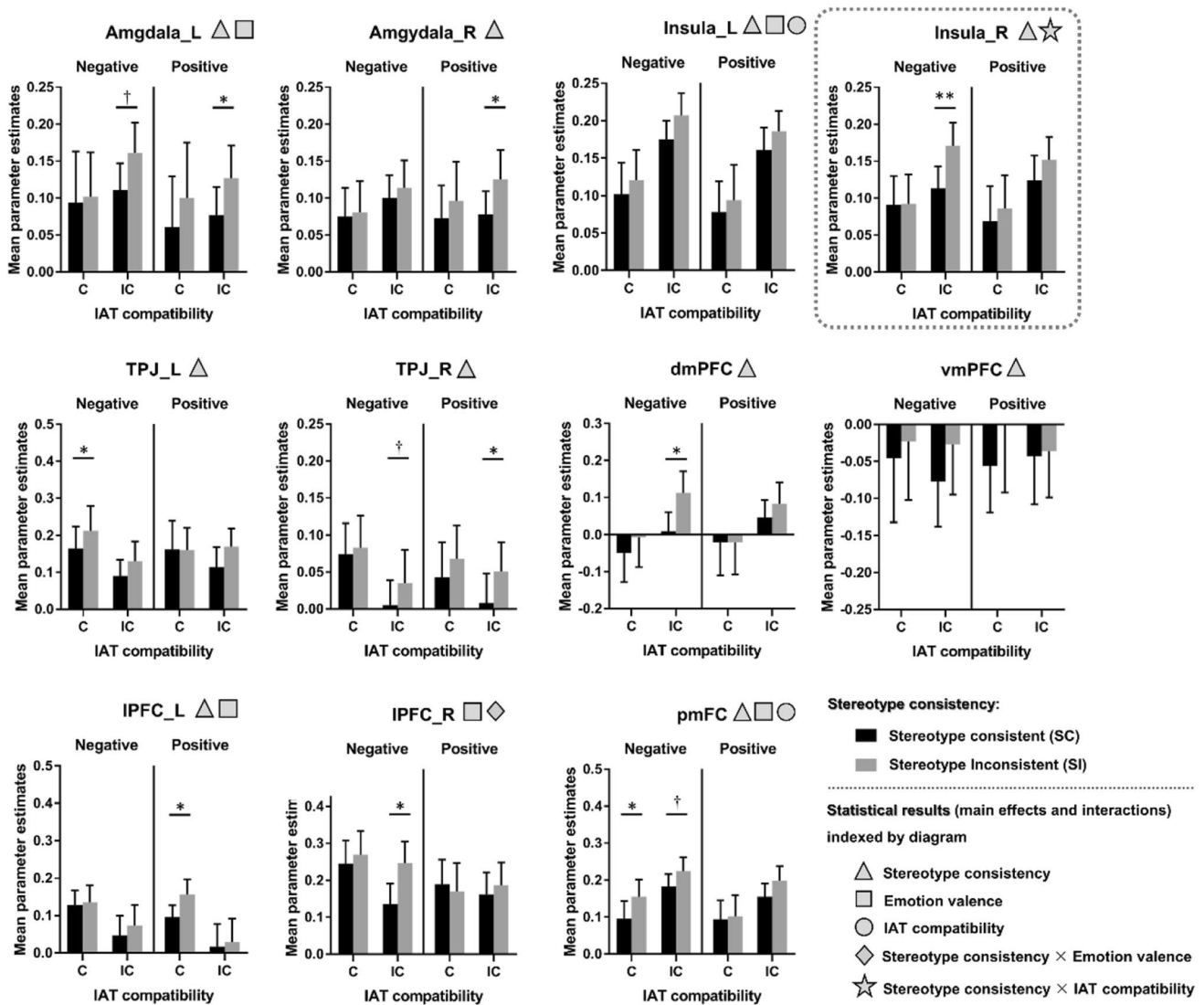


Fig. 5 Results of the signal change estimations on ROIs ($M \pm SE$). Repeated-measures ANOVA with stereotype consistency (SC vs. SI), emotional valence (positive vs. negative), and IAT compatibility (compatible vs. incompatible) was conducted to test for effects of percentage signal change on each ROI. For the ANOVA results, triangles indicate a main effect of stereotype consistency; squares indicate a main effect of emotional valence; circles indicate a main effect of IAT compatibility; rhombuses indicate a two-way interaction between stereotype consistency and emotional valence; and stars indicate a two-

way interaction between stereotype consistency and IAT compatibility. Three-way interaction between stereotype consistency, emotional valence, and IAT compatibility was not observed. Significant effects in the contrasts between SC and SI conditions on signal change estimations are marked by asterisks ($*p < 0.05$) based on paired-samples *t*-tested corrected for multiple comparisons using the Bonferroni method. Marginally significant effects in the contrasts between SC and SI conditions on signal change estimations are marked by cross symbols ($^{\dagger}0.05 < p < 0.10$, uncorrected).

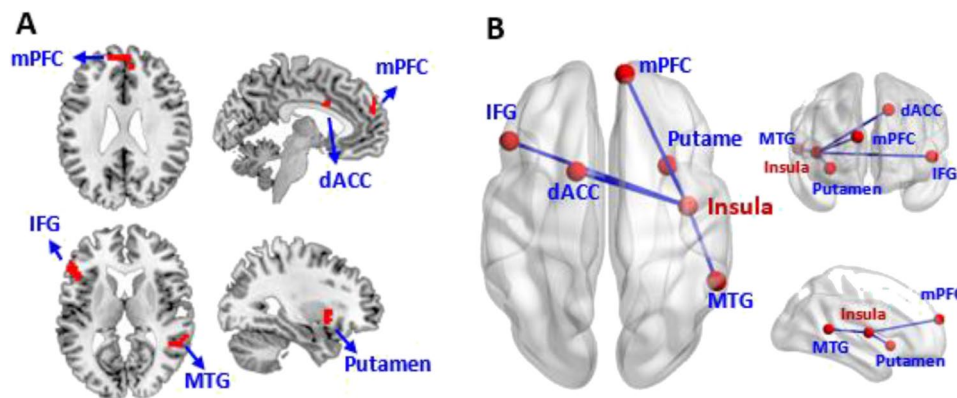
regulation of implicit stereotype commonly requires more attention and self-awareness than does stereotype activation (Bartholow et al., 2006; Correll et al., 2006). Consequently, this could explain why the right posterior insula, but not the anterior insula, was activated in the stereotype consistency \times IAT compatibility interaction and sensitive to stereotyping regulation (especially in negative stereotyping contexts).

Second, in addition to activity in the insula, we found that core mentalizing regions, namely the mPFC and TPJ, were involved in the stereotyping regulation network. In

previous research on racial and gender stereotypes, TPJ and dmPFC have been shown to be involved in stereotype activation and regulation tasks (Amodio, 2014; Contreras et al., 2011; Delplanque et al., 2019; Knutson et al., 2007). However, stronger brain activity in TPJ and dmPFC is commonly observed when processing stereotype-inconsistent information as compared to stereotype-consistent information, indicating that more social mentalizing resources are required to process stereotype violations (Cloutier et al., 2011). Hence, the TPJ and dmPFC activity in this study might indicate

Table 4 Brain regions identified by the generalized psychophysiological interaction (gPPI) analysis

| Regions activated | Hem | Peak MNI coordinate | | | <i>t</i> | <i>k</i> |
|----------------------------|------|---------------------|-----|----|----------|----------|
| | | x | y | z | | |
| Compatible task: SI > SC | | | | | | |
| None | | | | | | |
| Incompatible task: SI > SC | | | | | | |
| * MPFC | R, L | 9 | 57 | 21 | 3.94 | 73 |
| * dACC | L | -15 | 9 | 42 | 4.39 | 71 |
| * vIPFC | L | -48 | 24 | 6 | 4.32 | 66 |
| * MTG | R | 54 | -45 | 12 | 4.51 | 42 |
| * putamen | R | 30 | 12 | -3 | 5.33 | 30 |
| Postcentral Gyrus | R | 42 | -9 | 51 | 3.46 | 39 |

**Fig. 6** The functional connectivity of the right insula identified in the PPI analysis. **a** Clusters of brain regions were identified in this PPI analysis. Brain regions shown in this figure were identified using a

less stringent cluster-wise correction at $p < 0.005$ (uncorrected) with $K > 30$. **b** Illustration of the right insula's functional connectivity

similar mechanisms of stereotyping. It should be noted that the left TPJ, right TPJ, and the dmPFC that were recruited in negative stereotyping regulation in this study showed different processing patterns. The right TPJ and the dmPFC and showed a similar pattern to that of the right insula, with its activity being associated with the contrast of $SI > SC$ in the negative incompatible task and disappearing in the negative compatible task. The right TPJ also was associated with the $SI > SC$ in the positive incompatible task. However, the left TPJ activity showed a different pattern, and it was just associated with the $SI > SC$ in the negative compatible task. The different processing patterns of the left TPJ, right TPJ, and dmPFC might indicate different negative stereotyping regulation processes between the compatible task and the incompatible task. Indeed, compared to the right TPJ which is sensitive to social inconsistency, the left TPJ is usually recruited for the stereotype activation processes (Quadflieg et al., 2008). In addition, although the right TPJ and dmPFC are both core mentalizing regions involved in processing social inconsistency, The right TPJ is usually recruited to resolve social inconsistency in implicit and automatic

processes, whereas the dmPFC is commonly recruited to resolve social inconsistency in intentional and elaborate evaluation processes (Lieberman, 2007; Ma et al., 2012; Van Duynslaeger et al., 2007; Van Overwalle & Baetens, 2009). In our compatible task, the pattern of key-presses for SC and SI traits was consistent with implicit stereotypes against RMWs (either negative or positive valence). It appears that the stereotype inconsistency could be regulated with less intentional social mentalizing input, which is compatible with the right TPJ. However, because of the reversed key-press component of the incompatible task, more intentional controlled processes were involved, including trait coding, semantic conflict control, and stereotypical belief suppression, and these processes might engage dmPFC (Lieberman, 2007; Van Overwalle, 2009).

Third, extending work made by Wraga et al. (2007), the present study provides new evidence for the neural distinctions between negative and positive stereotyping processes. In the present study, our fMRI results showed a stereotype consistency \times emotional valence interaction in the brain activations of the right fusiform gyrus, left cerebellum, and

right PCC (the ROI analysis also showed that the right IPFC were sensitive to stereotype consistency \times emotional valence interaction, but it didn't pass the less stringent cluster-wise correction in the whole brain analysis). Among these brain regions, fusiform gyrus are usually engaged in visual recognition, while right PCC and right IPFC are usually involved in cognitive control (Amodio, 2014; Hayden et al., 2010; Knutson et al., 2007). Hence, our finding indicates that positive affect, compared with negative affect, might facilitate visual recognition and strengthen the cognitive control in stereotype regulation tasks.

Finally, our findings also demonstrate different insula functional connectivity between the compatible task and the incompatible task in negative stereotyping. In the whole-brain results, the right insula was sensitive to the interaction between stereotype consistency and IAT compatibility. For that reason, we tested its functional connectivity in negative valence conditions only. Consequently, the gPPI analysis revealed complex functional connectivity in the negative incompatible task, which included not only general cognitive control (i.e., vIPFC & dACC) and memory regions (MTG), but also social mentalizing (mPFC) and motivation/affective regions (putamen). However, such functional connectivity disappeared in the negative compatible task, indicating that this complex functional connectivity was generated in a more intense stereotype regulation context that involved high cognitive conflict control (or high strength of insula activity).

To the best of our knowledge, this study provides the first evidence of insula functional connectivity in stereotyping processes. This evidence supports the previous brain model of stereotyping (Amodio, 2014) observed in studies from the Western culture, and it further confirms our prediction that that insula has functional connectivity with the semantic, mentalizing, and conflict control regions. Therefore, the right posterior insula should be considered to be a crucial node in neural circuits for stereotype regulation. Specifically, when stereotypes against RMWs were activated and subsequently regulated, insula activity was linked to mentalizing and cognitive control regions in the network. This might bring about two possible processing pathways in stereotyping regulation. First, the enhanced right insula activity could raise the self-awareness (or arousal) level to alter the conflict monitoring and control processes reflected by brain activity in dACC and vIPFC (Cole & Schneider, 2007). Second, the enhanced right insula activity could send a strengthened conflict inhibition signal to the social mentalizing and memory retrieval regions in the frontal and temporal cortexes (i.e., mPFC & MTG), so that interference from the automatic stereotype belief can be controlled.

More evidence will be needed to understand the role of the insula in stereotype regulation. Nevertheless, our research indicates that insula has a special conflict control role in stereotyping regulation that is different from

the emotion/affective input function that has been revealed in previous studies (Amodio, 2014; Knutson et al., 2007; Lieberman et al., 2005).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13415-022-00988-z>.

Acknowledgments This research was supported by the National Natural Science Foundation of Zhejiang Province of China [LY20C090003] and was conducted with the approval of the ethics committee of Southwest University and Zhejiang Normal University. There are no potential or actual conflicts of interest associated with this research. The authors thank two anonymous reviewers for their constructive suggestions on an earlier draft. They also thank Brian Compton from University of California, San Diego for English language editing.

Author contributions L. J.: experiment design, data acquisition, data analysis, manuscript writing, and revision. B.S.: manuscript writing and revision. J. W.: experiment design, manuscript writing and revision.

References

- Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature reviews. Neuroscience*, 15(10), 670–682. <https://doi.org/10.1038/nrn3800>
- Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley
- Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, 90(2), 272–287. <https://doi.org/10.1037/0022-3514.90.2.272>
- Blair, I. V. (2002). The Malleability of Automatic Stereotypes and Prejudice. *Personality and Social Psychology Review*, 6(3), 242–261. https://doi.org/10.1207/s15327957pspr0603_8
- Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, 70(6), 1142–1163. <https://doi.org/10.1037/0022-3514.70.6.1142>
- Brewer, M. (1988). A dual process model of impression formation. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Erlbaum.
- Cattaneo, Z., Mattavelli, G., Platania, E., & Papagno, C. (2011). The role of the prefrontal cortex in controlling gender-stereotypical associations: A TMS investigation. *NeuroImage*, 56(3), 1839–1846. <https://doi.org/10.1016/j.neuroimage.2011.02.037>
- Chekroud, A., Everett, J., Bridge, H., & Hewstone, M. (2014). A review of neuroimaging studies of race-related prejudice: does amygdala response reflect threat? *Frontiers in Human Neuroscience*, 8(179). <https://doi.org/10.3389/fnhum.2014.00179>
- Cisler, J. M., Bush, K., & Steele, J. S. (2014). A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI. *Neuroimage*, 84(1), 1042–1052. <https://doi.org/10.1016/j.neuroimage.2013.09.018>
- Cloutier, J., Gabrieli, J. D. E., O'Young, D., & Ambady, N. (2011). An fMRI study of violations of social expectations: When people are not who we expect them to be. *NeuroImage*, 57, 583–588. <https://doi.org/10.1016/j.neuroimage.2011.04.051>
- Cole, M. W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, 37(1), 343–360. <https://doi.org/10.1016/j.neuroimage.2007.03.071>

- Contreras, J. M., Banaji, M. R., & Mitchell, J. P. (2011). Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Social Cognitive and Affective Neuroscience*, 7, 764–770. <https://doi.org/10.1093/scan/nsr053>
- Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology*, 42(1), 120–128. <https://doi.org/10.1016/j.jesp.2005.02.006>
- Craig, A. D. (2009). How do you feel now? The anterior insula and human awareness. *Nature reviews. Neuroscience*, 10(1), 59–70. <https://doi.org/10.1038/nrn2555>
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149): Academic Press.
- Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., ... Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, 48(1), 1–33. <https://doi.org/10.1348/014466608X314935>
- Delplanque, J., Heleven, E., & Van Overwalle, F. (2019). Neural representations of Groups and Stereotypes using fMRI repetition suppression. *Scientific Reports*, 9(1), 3190. <https://doi.org/10.1038/s41598-019-39859-y>
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1–74). Academic Press. [https://doi.org/10.1016/S0065-601\(08\)60317-2](https://doi.org/10.1016/S0065-601(08)60317-2)
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold. *Magnetic Resonance in Medicine*, 33(5), 636–647. <https://doi.org/10.1002/mrm.1910330508>
- Gallate, J., Wong, C., Ellwood, S., Chi, R., & Snyder, A. (2011). Noninvasive Brain Stimulation Reduces Prejudice Scores on an Implicit Association Test. *Neuropsychology*, 25, 185–192. <https://doi.org/10.1037/a0021102>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60(4), 509–517. <https://doi.org/10.1037/0022-3514.60.4.509>
- Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups. *Psychological Science*, 17(10), 847–853. <https://doi.org/10.1111/j.1467-9280.2006.01793.x>
- Hayden, B., Smith, D., & Platt, M. (2010). Cognitive Control Signals in Posterior Cingulate Cortex [Original Research]. *Frontiers in Human Neuroscience*, 4(223). <https://doi.org/10.3389/fnhum.2010.00223>
- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47, 237–271. <https://doi.org/10.1146/annurev.psych.47.1.237>
- Hughes, C., Babbitt, L. G., & Krendl, A. C. (2019). Culture Impacts the Neural Response to Perceiving Outgroups Among Black and White Faces. *Frontiers in Human Neuroscience*, 13(143). <https://doi.org/10.3389/fnhum.2019.00143>
- Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, 91(1), 16–32. <https://doi.org/10.1037/0022-3514.91.1.16>
- Knutson, K. M., Mah, L., Manly, C. F., & Grafman, J. (2007). Neural correlates of automatic beliefs about gender and race. *Human Brain Mapping*, 28(10), 915–930. <https://doi.org/10.1002/hbm.20320>
- Kuang, L., & Liu, L. (2012). Discrimination against rural-to-urban migrants: The role of the hukou system in China. *PLoS ONE*, 7, e46932. <https://doi.org/10.1371/journal.pone.0046932>
- Lieberman, M. D. (2007). Social Cognitive Neuroscience: A Review of Core Processes. *Annual Review of Psychology*, 58, 259–289. <https://doi.org/10.1146/annurev.psych.58.110405.085654>
- Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social Cognitive and Affective Neuroscience*, 4(4), 423–428. <https://doi.org/10.1093/scan/nsp052>
- Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, 8(6), 720–722. <https://doi.org/10.1038/nn1465>
- Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology*, 27, 363–385. <https://doi.org/10.1146/annurev.soc.27.1.363>
- Ma, N., Vandekerckhove, M., Baetens, K., Van Overwalle, F., Seurinck, R., & Fias, W. (2012). Inconsistencies in spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, 7(8), 937–950. <https://doi.org/10.1093/scan/nsr064>
- McLaren, D. G., Ries, M. L., Xu, G., Johnson, S.C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *NeuroImage*, 61(4), 1277–1286. <https://doi.org/10.1016/j.neuroimage.2012.03.068>
- Quadflieg, S., & Macrae, C. N. (2011). Stereotypes and stereotyping: What's the brain got to do with it? *European Review of Social Psychology*, 22(1), 215–273. <https://doi.org/10.1080/10463283.2011.627998>
- Quadflieg, S., Turk, D. J., Waiter, G. D., Mitchell, J. P., Jenkins, A. C., & Macrae, C. N. (2008). Exploring the Neural Correlates of Social Stereotyping. *Journal of Cognitive Neuroscience*, 21(8), 1560–1570. <https://doi.org/10.1162/jocn.2009.21091>
- Rosenthal, L., & Overstreet, N.M. (2016). Stereotyping. In H. Friedman (Ed.), *Encyclopedia of Mental Health* (2nd edition Vol. 4, pp. 225–229), Academic Press.
- Straube, T., & Miltner, W. H. R. (2011). Attention to aversive emotion and specific activation of the right insula and right somatosensory cortex. *NeuroImage*, 54(3), 2534–2538. <https://doi.org/10.1016/j.neuroimage.2010.10.010>
- Van der Cruyssen, L., Heleven, E., Ma, N., Vandekerckhove, M., & Van Overwalle, F. (2015). Distinct neural correlates of social categories and personality traits. *NeuroImage*, 104, 336–346. <https://doi.org/10.1016/j.neuroimage.2014.09.022>
- Van Duynslaeger, M., Van Overwalle, F., & Verstraeten, E. (2007). Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, 2(3), 174–188. <https://doi.org/10.1093/scan/nsm016>
- Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, 30(3), 829–858. <https://doi.org/10.1002/hbm.20547>
- Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, 48(3), 564–584. <https://doi.org/10.1016/j.neuroimage.2009.06.009>
- Wang, L., Ma, Q., Song, Z., Shi, Y., Wang, Y., & Pfotenhauer, L. (2011). N400 and the activation of prejudice against rural migrant workers in China. *Brain Research*, 1375, 103–110. <https://doi.org/10.1016/j.brainres.2010.12.021>

- Wong, C. L., Harris, J. A., & Gallate, J. E. (2011). Evidence for a social function of the anterior temporal lobes: Low-frequency rTMS reduces implicit gender stereotypes. *Social Neuroscience*, 7(1), 90–104. <https://doi.org/10.1080/17470919.2011.582145>
- Wraga, M., Helt, M., Jacobs, E., & Sullivan, K. (2007). Neural basis of stereotype-induced shifts in women's mental rotation performance. *Social Cognitive and Affective Neuroscience*, 2(1), 12–19. <https://doi.org/10.1093/scan/nsl041>
- Wu, X., & Treiman, D. J. (2004). The household registration system and social stratification in China: 1955–1996. *Demography*, 41(2), 363–384. <https://doi.org/10.1353/dem.2004.0010>
- Zhen, J., & Liu, L. (2012). Undergraduates' stereotype on migrant workers: Content and structure. *Youth Studies*, 4, 35–44. (Published in Chinese) CNKI:SUN:QNYJ.0.2012-04-005

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.