The hierarchical sensitivity to social misalignment during decision-making under uncertainty

Yongling Lin,[1] Ruolei Gu,[2,3] Shenghua Luan,[2,3] Li Hu,[2,4] Shaozheng Qin,[1] and Yue-jia Luo[1,5,6,7]

[1]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China, [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China, [3]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing 100101, China, [4]CAS Key Laboratory of Mental Health, Institute of Psychology, Beijing 100101, China, [5]Center for Brain Disorder and Cognitive Science, Shenzhen University, Shenzhen, Guangdong 518061, China, [6]College of Teacher Education, Qilu Normal University, Jinan, Shandong 250200, China, and [7]The Research Center of Brain Science and Visual Cognition, Kunming University of Science and Technology, Kunming, Yunnan 650504, China

Yongling Lin and Ruolei Gu contributed equally to this study.

Correspondence should be addressed to Yue-jia Luo, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China. E-mail: luoyj@bnu.edu.cn.

Abstract

Social misalignment occurs when a person's attitudes and opinions deviate from those of others. We investigated how individuals react to social misalignment in risky (outcome probabilities are known) or ambiguous (outcome probabilities are unknown) decision contexts. During each trial, participants played a forced-choice gamble, and they observed the decisions of four other players after they made a tentative decision, followed by an opportunity to keep or change their initial decision. Behavioral and event-related potential data were collected. Behaviorally, the stronger the participants' initial preference, the less likely they were to switch their decisions, whereas the more their decisions were misaligned with the majority, the more likely they were to switch. Electrophysiological results showed a hierarchical processing pattern of social misalignment. Misalignment was first detected binarily (i.e. match/mismatch) at an early stage, as indexed by the N1 component. During the second stage, participants became sensitive to low levels of misalignment, which were indexed by the feedback-related negativity. The degree of social misalignment was processed in greater detail, as indexed by the P3 component. Moreover, such hierarchical neural sensitivity is generalizable across different decision contexts (i.e. risky and ambiguous). These findings demonstrate a fine-grained neural sensitivity to social misalignment during decision-making under uncertainty.

Key words: social misalignment; social conformity; decision-making; uncertainty; event-related potential

Introduction

The power of social influence such as the attitudes and judgment of others on individual decision-making has been well documented (Cialdini and Goldstein, 2004; Raafat *et al.*, 2009; Zhang and Gläscher, 2020). People are preferentially aligned with the majority (i.e. social alignment) when faced with uncertainties, which may be an evolutionarily adaptive strategy (Nakahashi, 2007; Nakahashi *et al.*, 2012; Morgan *et al.*, 2015).

In a social environment, people may observe the decisions of others and monitor the degree of misalignment between those and their thoughts to inform their own decisions (Tump *et al.*, 2020; Zhang and Gläscher, 2020). Social misalignment pertains to situations where individual opinions and decisions deviate from those of others (Shamay-Tsoory *et al.*, 2019). Reactions to social misalignment are essential for follow-up behavioral adjustments and social conformity (Klucharev *et al.*, 2009; Shamay-Tsoory *et al.*, 2009). Previous studies have demonstrated that social alignment during a short interaction could have long-lasting effects on memory (Edelson *et al.*, 2011) and individual preference (Heerdink *et al.*, 2013; Izuma and Adolphs, 2013).

Neuroscience research can deepen the understanding of the underlying mechanisms of social misalignment. Brain imaging studies have found that social misalignment, compared with social alignment, evokes a stronger activation of the dorsal posterior medial frontal cortex (dorsal pMFC) (Izuma, 2013; Izuma and Adolphs, 2013) and anterior insula (Berns *et al.*, 2005) but weaker activations in the ventral striatum (VS) and orbitofrontal cortex (Campbell-Meiklejohn *et al.*, 2010; Zaki *et al.*, 2011; Charpentier *et al.*, 2014). Further, a transcranial magnetic stimulation study shows that the downregulation of the pMFC reduced social conformity (Klucharev *et al.*, 2011). These findings suggest that conflict with the group's opinion triggers cognitive imbalance and aversive feelings (Klucharev *et al.*, 2009; H. Wu *et al.*, 2016). Consistent with this idea, event-related potential (ERP) studies have revealed that being misaligned with majority opinions evokes ERP components associated with conflict detection and expectancy violation, including the feedback-related negativity (FRN) (Kim *et al.*, 2012; Yu *et al.*, 2018) and N400 (Huang *et al.*, 2014; Feng *et al.*, 2018). For instance, disagreeing with group opinion evokes a larger FRN than agreeing with them (Chen *et al.*, 2012; Shestakova *et al.*, 2013). These frontocentral electrophysiological deflections (i.e. the FRN and N400) are attributed to an error-monitoring system related to a perceived violation of social norms (Mu *et al.*, 2015). Here, the involvement of the error-monitoring system is understandable as individuals tend to believe that others are more likely to make the same decisions as theirs, which is known as the 'false consensus effect' (Ross and Sicoly, 1979; Dawes, 1989). In the same vein, electroencephalography (EEG) studies have found that social misalignment affects unconscious perceptual processing (Trautmann-Lengsfeld and Herrmann, 2013; Zhu *et al.*, 2019) and social pressure enhances early attention to stimulus discrimination, indicated by a higher N1 amplitude (Germar *et al.*, 2016). Social misalignment may capture more attentional resources, which may manifest as a higher N1 (Santamaría-García *et al.*, 2014). In contrast, an augmented ERP P3 component has been observed when individual opinions are aligned with group decisions, suggesting that being aligned with the group is more emotionally rewarding (Yu and Sun, 2013; Xie *et al.*, 2016).

While findings from the aforementioned studies are insightful for understanding the behavioral patterns and neural correlates when individuals are misaligned with the majority, most used a binary categorization for the responses of participants: agreeing or disagreeing with the majority (e.g. Kim *et al.*, 2012; Yu and Sun, 2013; Schnuerch and Gibbons, 2015; Schnuerch *et al.*, 2016). This dichotomy may be too simple to unravel the complex dynamics of social misalignment. One of our studies has shown that human brain activity encodes the social comparison between self and others in a binary manner (i.e. match/mismatch) during its early stages, but the level of social mismatch is evaluated hierarchically during later stages (Luo *et al.*, 2015). In our opinion, the influence of different levels of social misalignment on neural activation needs to be

elaborated to improve the understanding of how the brain encodes social misalignment. Further, dissociating underlying cognitive processes (e.g. attention bias and cognitive conflict) of social misalignment may improve behavioral prediction with neural indexes. Some studies have shown that the activation of the pMFC (a brain area related to social misalignment) can predict human behaviors more accurately than self-reported measures (Falk *et al.*, 2011, 2012).

This study employed the ERP technique because its high time resolution facilitates the identification of the temporal characteristics of the cognitive processes involved in social misalignment during decision-making (Ibanez *et al.*, 2012; Scheepers and Derks, 2016). Based on previous findings, we focused on three ERP components: N1, FRN and P3. The N1 component is associated with perceptual discrimination (Luck *et al.*, 2000; Wang and Suemitsu, 2007) and may reflect the discrimination process of social misalignment—the detection of social misalignment in the current context (Trautmann-Lengsfeld and Herrmann, 2013; Germar *et al.*, 2016). Meanwhile, FRN indicates the monitoring of unfavorable outcomes that violate prior expectations (Gehring and Willoughby, 2002; Masaki *et al.*, 2006; Chen *et al.*, 2012). In light of the 'false consensus effect', we predicted that the FRN amplitude would increase as a function of social misalignment. Finally, the P3 amplitude is sensitive to the emotional and motivational significance of an ongoing event (Polich, 2007; Wu and Zhou, 2009); thus, P3 may be enhanced when individuals are aligned with the majority (Yu and Sun, 2013; Xie *et al.*, 2016). Moreover, this component is considered to reflect an elaborate process of information evaluation (Gu *et al.*, 2011; Zhang *et al.*, 2013) and, therefore, may be capable of ascertaining the degree of social misalignment. Hence, we developed the following ERP hypotheses:

H_1. The N1 may reflect the early discrimination process of social misalignment;

H_2. The FRN increases as a function of the degree of social misalignment;

H_3. Being congruent with the majority enhances the P3 amplitude.

Reducing uncertainties is a major incentive for aligning with others (Rendell *et al.*, 2010; Toelch and Dolan, 2015). This study also determines whether the neural processing of social misalignment can be generalized across different forms of decision-making with uncertainties (i.e. risky *vs* ambiguous; see Xiong *et al.*, 2020). Researchers have disproportionately focused on the social influence on risky decision-making associated with well-defined probabilities (e.g. Suzuki *et al.*, 2016; Reiter *et al.*, 2019; Chung *et al.*, 2020), while largely ignoring the ambiguous decision-making related to unknown probabilities. Since both forms of uncertain decision-making play important roles in daily social interactions (Hsu *et al.*, 2005), it is important to ascertain whether the neural underpinnings of social misalignment can be generalized from risky to ambiguous decision-making. We did not make a priori hypothesis for this specific topic due to the lack of previous studies.

Method

Participants

Thirty-four healthy participants [16 females; age range: 19–24 years, mean ± standard deviation (s.d.) = 21.2 ± 1.4 years] were recruited from several universities in Beijing, China.

We determined the sample size based on two principles: (i) the number of participants usually involved in previous ERP studies in the relevant field (e.g. Chen *et al.*, 2012; Kim *et al.*, 2012; Yu and Sun, 2013) and (ii) the number of participants needed to ensure 80% statistical power for detecting a small-to-medium effect size (Vazire, 2016). All participants were right-handed and had normal or corrected-to-normal vision. None of them reported a history of neurological or psychiatric disorders. The participants were encouraged to maximize their reward during the experiment. They were informed that the basic payment for their participation would be 100 Yuan and their earnings would be the basic payment plus the cumulative income for four random trials during the task. This payment rule was set to ensure that the best strategy for the participants was to treat the trials equally (Knutson *et al.*, 2007). Written informed consent was obtained before the experiment, which was approved by the Institutional Review Board of the CAS Institute of Psychology. This study was performed strictly in accordance with the approved guidelines.

Task design

Experimental stimuli. Two bottles, each of which contained 20 balls, were shown to participants during every trial of the task (Figure 1A). Three types of colored balls (i.e. red, blue and yellow) were used, and they were counterbalanced across the participants in the experiment. There was only one type of color balls in one of the bottles ('safe bottle') and two types in the other bottle ('uncertain bottle'). The magnitude of the reward corresponding to each ball color was shown beneath the two bottles. For the uncertain bottle, one color represented a large reward (18, 19, 20, 21 or 22 Chinese Yuan; approximately 3–3.5 US dollars), whereas the other color represented no reward. The color of the ball in the safe bottle represented a small magnitude of reward (varied from 5.4 to 15.4 Yuan), which was set following the rule that the expected values (EVs) of the two bottles were equal for the risky context. All reward stimuli were applied to both the risky and ambiguous contexts to ensure their physical comparability, although the EV could not be calculated for the ambiguous context (because of the absence of probability information). The participants were not provided with the rules for designing the experimental stimuli (i.e. bottles).

Regarding the uncertain bottle, if all the color balls inside could be seen clearly, the situation was considered as a 'risky context' (i.e. the probability of drawing each kind of color ball was well defined); if the color balls were obscure, the situation was called an 'ambiguous context'. These definitions are consistent with the classic concepts of risk and ambiguity (Rothschild and Stiglitz, 1970). The probability of drawing a rewarding ball from an uncertain bottle varied from 30% to 70% across the trials (step length = 10%). Here, we kept the winning probability within this range to diminish the effect of the weight function, which indicates that humans have anomalous behavioral and neural responses when event probability is near impossible or certain (Tversky and Fox, 1995; Hsu *et al.*, 2009). The same manipulation has been widely applied in previous studies (Levy *et al.*, 2010; Tymula *et al.*, 2012; Pushkarskaya *et al.*, 2015). The main purpose of varying the outcome and winning probability across the trials was to keep participants focusing on the task. Given that there were five different outcome values and five winning probabilities, there were 25 (i.e. 5 × 5) trials for both the risky and ambiguous contexts.

The feedback stimuli were presented as five ticks. A black tick indicated the participant's decision, and four gray ticks indicated the decisions made by four other players. These ticks appeared on the left or right side of a fixation point depending on the participant's choice. For example, when three gray ticks appeared on the left side while one black and one gray tick appeared on the right side (Figure 1A), they indicated that three anonymous players had selected the left option while the real participant and another player had selected the right option.

Procedure. The experimental procedure was programmed and performed using E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA). Before the experiment, the participants were instructed about the rules and the meanings of the symbols for the task. They were told that during each trial of the task, they needed to choose between an uncertain bottle and a safe bottle, and the computer would randomly draw a color ball from the selected bottle to determine their reward. However, the draw outcome of each trial would not be revealed until the participants had completed all the required tasks for the experiment to avoid the potential influence of the prior outcome
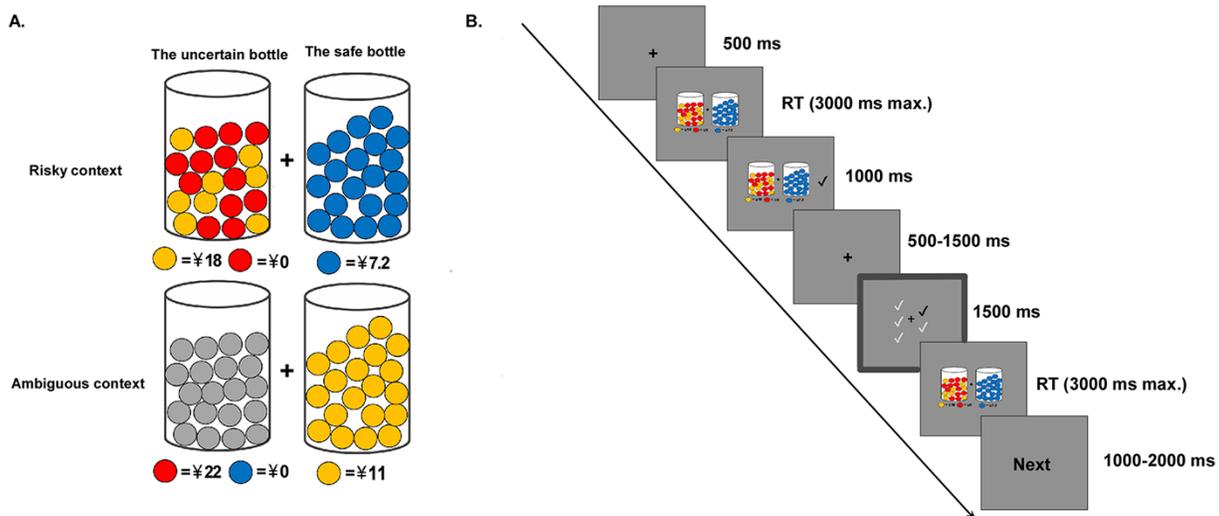

Fig. 1. Experimental design. (A) The risky context and the ambiguous context. (B) Structure of a trial. The target event for the ERP data analysis was the presentation of feedback, which included the decisions of a participant (bold rectangle) and the other four players.

on subsequent decision-making (Zhang *et al.*, 2013; Lin *et al.*, 2019).

During the task, each participant sat in an electrically shielded room approximately 75 cm from a 19-inch LCD computer monitor. The stimuli were presented on the screen at a visual angle of $3.0° \times 3.5°$. As illustrated in Figure 1B, each trial began with a central fixation point. After 500 ms, the two bottles and the corresponding reward information of each ball color were presented on either side of the fixation point. The positions (left/right) of the uncertain bottle and the safe bottle were counterbalanced across the trials. Each participant made a decision by pressing the 'F' key to choose the left bottle or the 'J' key to choose the right bottle within 3000 ms. The selected bottle was highlighted by a black tick for 1000 ms. After this confirmation stage, only the fixation point remained on the screen between 500 and 1200 ms. The participants were provided with the decisions of the other players (four gray ticks), as well as their decisions (a black tick), for 1500 ms. The participants were told that this feedback was retrieved from the decision history of other participants that had completed the task before. In reality, the feedback was pseudo-randomly determined, and its level of misalignment was controlled. After receiving this feedback, the two bottles were presented again for 3000 ms, giving the participant a final opportunity to maintain the initial selection or switch to the other bottle. At last, the word 'Next' was presented at the center of the screen to replace the fixation point for 1000–2000 ms and remind the participant that the current trial had ended.

Overall, the task used a 2 ('uncertain context': risky/ambiguous) $\times$ 5 ['social misalignment': 0-Inconsistency (Incon)/1-Incon/2-Incon/3-Incon/4-Incon] within-subject design. Here, the level of social misalignment was defined according to the number of players who made a different decision from that of the real participant. For example, for the 4-Incon condition, the participant found that his/her initial decision was inconsistent with those of all the four other players. The 2×5 task design permitted 10 conditions, each of which comprised 25 trials (i.e. 250 in total). The entire task lasted for approximately 50 min for each participant, with a short break after every 80 trials. At the end of the task, the participants were paid 100–180 Yuan depending on their task performance and were debriefed.

Behavioral measures

We analyzed three behavioral indices in this study: gamble ratio, initial preference strength and switching rate. Gamble ratio was defined as the participants' initial preference for choosing the uncertain (i.e. risky and ambiguous) bottle, which was calculated by dividing the number of trials during which the uncertain bottle was chosen as the initial decision by the total number of trials. The initial preference strength was defined as the behavioral deviation from preference neutrality, which was calculated as the absolute difference between the gamble ratio and 0.5 (i.e. preference neutrality). Finally, the switching rate was defined as the behavioral adjustment between participants' initial and final decisions, and it was calculated by dividing the number of trials in which participants switched to a different bottle as their second choice by the total number of trials for each condition.

Electrophysiological recording and pre-processing

EEG was recorded from 64 scalp sites using tin electrodes mounted on an elastic cap (Compumedics Neuroscan Inc., El Paso, TX) with an online reference to the left mastoid and off-line algebraic re-reference to the average of the left and right mastoids. To monitor ocular movements and eye blinks, electrooculogram (EOG) recordings were recorded from four electrodes placed lateral to each eye (i.e. horizontal EOG) and above and below the right eye (i.e. vertical EOG). Electrode impedance was maintained below 5 kΩ. The EEG and EOG signals were amplified with a 0.05–100 Hz online band-pass filter and continuously sampled at 1000 Hz/channel.

The EEG data were preprocessed using EEGLAB (Delorme and Makeig, 2004), an open-source toolbox running in the MATLAB environment (MathWorks, Natick, MA). Continuous EEG data were resampled to 500 Hz and low-pass filtered at 30 Hz. The segmented EEG data started 200 ms before feedback presentation through 800 ms with the first 200 ms pre-stimulus as baseline (-200 ms to 0 ms). The trials contaminated by the eye blinks and movements were corrected using an independent component analysis algorithm, such that any trials during which the EEG voltage exceeded a threshold of ±100 μV during the recording epoch were excluded from further analysis. In addition, the trials during which the participants did not make a decision on time (that is, within 3000 ms) were excluded. After the data preprocessing described above, the trials that survived were determined as artifact-free [overall mean value: 240.2 ± 10.6 (96.1%) trials; see Supplementary Table S1 for the remaining number of trials in each condition]. According to the literature, the remaining trials for each condition were sufficient to produce reliable ERP waveforms with a high signal-to-noise ratio (Cohen and Polich, 1997; Marco-Pallares *et al.*, 2011; Leue *et al.*, 2013).

ERP analysis

This study used ERPs elicited by feedback presentation (i.e. decisions made by four other players). We analyzed the following: the mean amplitude of the occipital N1 component (time window: 150–200 ms) across the electrodes PO3, PO4, PO5, PO6, O1 and O2; the frontal FRN (time window: 270–320 ms) across F1, Fz, F2, FC1, FCz and FC2; the parietal P3 component (time window: 300–500 ms) across P1, Pz, P2, PO3, POz and PO4. The time window for each ERP component was determined based on previous research (Luck *et al.*, 2000; Luck, 2014) and the visual detection of grand averaged waveforms. The average data were calculated from multiple electrodes to increase the stability of the ERP results (Luck and Gaspelin, 2017). The electrodes for each ERP component were selected according to the visual detection of the scalp distribution.

Statistics

Descriptive data are presented as arithmetic mean $\pm$ s.d. Greenhouse–Geisser corrections were used whenever appropriate unless otherwise stated. The data were analyzed using SPSS version 22.0 (SPSS, Inc., Chicago, IL) and MATLAB R2014b.

Behavioral statistics. Paired sample *t*-tests were conducted to investigate the differences between the gamble ratios and the initial preference strengths of the risk and ambiguous contexts. A repeated-measure analysis of variance (ANOVA) [2 (*uncertain context*: risky/ambiguous) $\times$ 5 (*social misalignment*: 0-Incon/1-Incon/2-Incon/3-Incon/4-Incon)] was used to analyze the switching rate. Finally, a two-tailed Pearson's *r* correlation was performed to determine the correlation between the initial preference strength and the switching rate and examine the relationship between the decision preference and behavioral adjustment.

ERP statistics. A repeated-measure ANOVA [2 ('uncertain context': risky/ambiguous) × 5 ('social misalignment': 0-Incon/1-Incon/2-Incon/3-Incon/4-Incon)] was used to analyze the amplitudes of the N1, FRN and P3 components. Further, mediation model analyses were performed to test the influence of the ERP components on the relationship between the switching rate and social misalignment. Bootstrapping was applied to test the significance of the amplitude of each ERP component as a mediator, with 1000 resamples and a confidence interval of 95% (Hayes and Rockwood, 2017).

Results

Behavioral results

The paired sample *t*-tests showed that the gamble ratio was higher for the risky context (0.475 ± 0.286) than for the ambiguous context (0.416 ± 0.268; $t = 3.580$, $P = 0.001$), and there was no significant difference between initial preference strengths for the risky (0.248 ± 0.127) and ambiguous (0.238 ± 0.152; $t = 0.515$, $P = 0.610$) contexts.

A repeated-measure ANOVA showed that the main effect of 'social misalignment' was significant [$F(4,33) = 17.437$, $P < 0.001$, $\eta^2_p = 0.346$], and the *post hoc* comparisons showed that the switching rate of the 0-Incon condition (0.093 ± 0.082) was not significantly different from that of the 1-Incon (0.087 ± 0.064, $P = 0.533$) and 2-Incon conditions (0.103 ± 0.076, $P = 0.304$), but it was lower than that of the 3-Incon (0.196 ± 0.198, $P = 0.001$) and 4-Incon conditions (0.258 ± 0.222, $P < 0.001$). Meanwhile, the switching rate of the 1-Incon condition was not significantly different from that of the 2-Incon condition ($P = 0.066$), but it was lower than that of the 3-Incon ($P = 0.002$) and 4-Incon conditions ($P < 0.001$). Finally, there were significant differences between the remaining three conditions (P-values < 0.002). The main effect of the 'uncertain context' on the switching rate was significant [$F(1,33) = 22.037$, $P < 0.001$, $\eta^2_p = 0.400$]: participants showed a higher switching rate for the risky context (0.186 ± 0.128) than for the ambiguous context (0.109 ± 0.111). The interaction of the 'uncertain context' and 'social misalignment' was insignificant [$F(4,33) = 2.389$, $P = 0.079$, $\eta^2_p = 0.068$] (Figure 2A).

Correlational analysis showed that the switching rate was negatively correlated with individual initial preference strength ($r = -0.514$, $P = 0.002$) (Figure 2B).

ERP results

N1 component. The main effect of 'social misalignment' was significant [$F(4,33) = 15.963$, $P < 0.001$, $\eta^2_p = 0.326$]; the N1 amplitude of the 0-Incon (0.219 ± 3.335 μV) condition was lower (i.e. less negative-going) than those of the other conditions (P-values < 0.001), whereas there were no significant differences between the remaining conditions (P-values > 0.5). The main effect of the 'uncertain context' was significant [$F(1,33) = 6.331$, $P = 0.017$, $\eta^2_p = 0.161$]; the N1 amplitude for the ambiguous context was higher than that of the risky context (-1.395 ± 3.825 μV *vs* -0.776 ± 3.540 μV). The interaction between 'uncertain context' and 'social misalignment' was not significant [$F(4,33) = 0.649$, $P = 0.629$, $\eta^2_p = 0.019$] (Figure 3).

FRN. The main effect of 'social misalignment' was significant [$F(4,33) = 9.467$, $P < 0.001$, $\eta^2_p = 0.223$]; the FRN amplitude significantly increased as a function of the level of 'social misalignment' for the 0-Incon (-1.024 ± 3.943 μV), 1-Incon (-2.043 ± 4.324 μV) and 2-Incon (-3.152 ± 4.412 μV) conditions, but there was no significant difference among the 2-Incon, 3-Incon (-3.194 ± 4.673 μV) and 4-Incon (-3.291 ± 4.423 μV) conditions. The FRN amplitude of the 0-Incon condition was lower than those of the other conditions (P-values < 0.016); it was also lower for the 1-Incon condition than for the remaining conditions (P-values < 0.03). The main effect of the 'uncertain context' was significant [$F(1,33) = 6.657$, $P = 0.015$, $\eta^2_p = 0.168$]; the risky context showed a higher (i.e. more negative-going) FRN than the ambiguous context (-2.940 ± 4.246 μV *vs* -2.130 ± 4.014 μV). Finally, the interaction between 'uncertain context' and 'social misalignment' was not significant [$F(4,33) < 1$, $P = 0.572$, $\eta^2_p = 0.020$] (Figure 4).

P3 component. The main effect of 'social misalignment' was significant [$F(4,33) = 16.203$, $P < 0.001$, $\eta^2_p = 0.329$]; the P3
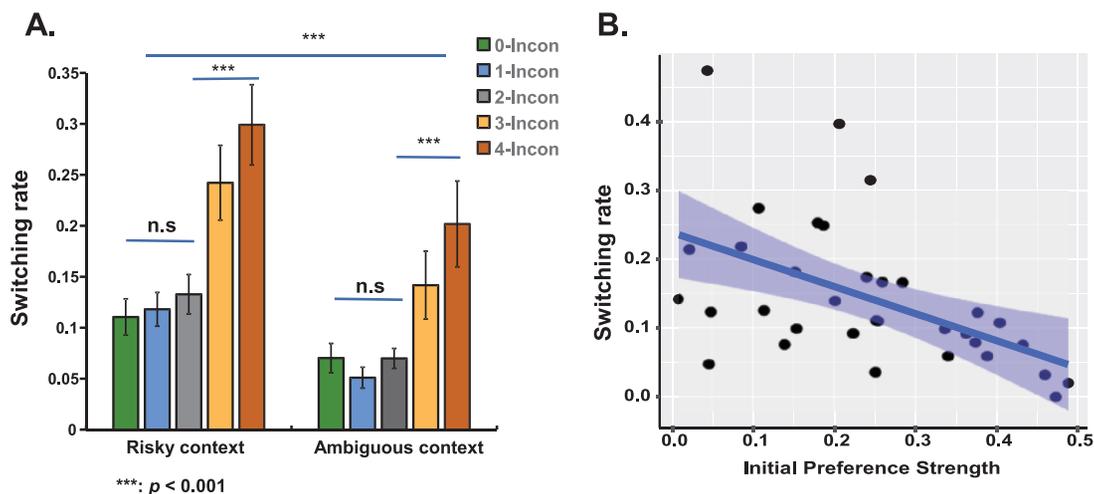

Fig. 2. (A) Behavioral results. The switching rate for each condition. (B) Results of correlation. The correlation between the initial preference strength and the switching rate.
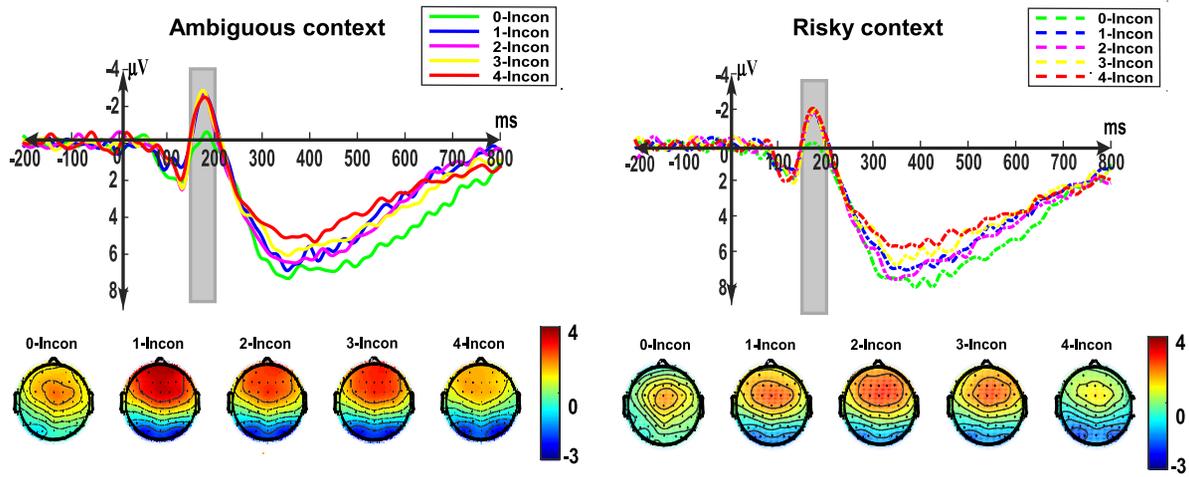
Fig. 3. Grand average ERP of the N1 component. Left to right, the occipital N1 waveforms in the ambiguous and risky contexts. The waveforms represent the mean values of the data at the electrode sites PO3, PO4, PO5, PO6, O1 and O2. The corresponding scalp topography for each level of social misalignment is provided below.
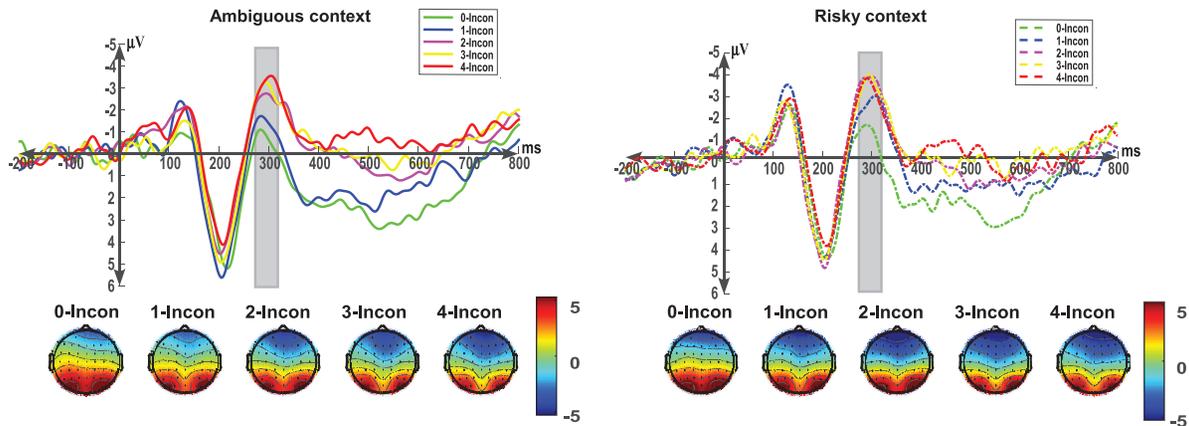

Fig. 4. Grand average ERP of the FRN. Left to right, the frontal FRN waveforms in the ambiguous and risky contexts. The waveforms represent the mean values of the data at the sites F1, Fz, F2, FC1, FCz and FC2. The corresponding scalp topography for each level of social misalignment is provided below.
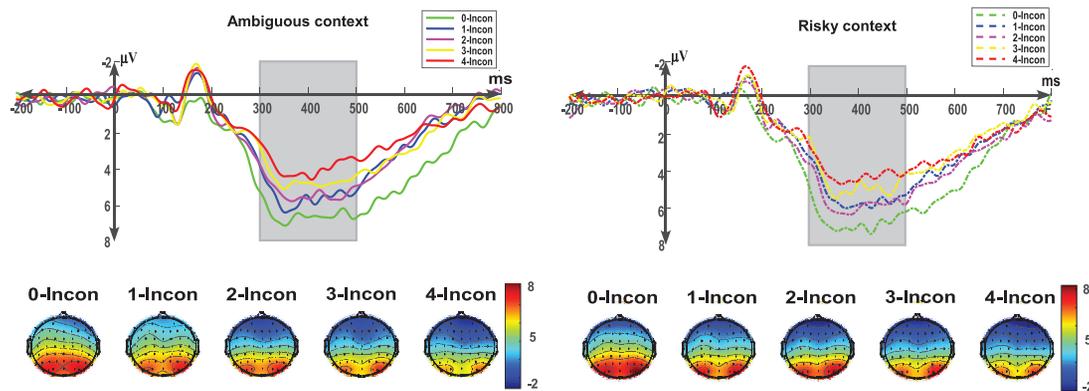

Fig. 5. Grand average ERP of the P3 component. Left to right, the parietal P3 waveforms in the ambiguous and risky contexts. The waveforms represent the mean values of the data at the sites P1, Pz, P2, PO3, POz and PO4. The corresponding scalp topography for each level of social misalignment is provided below.

amplitude gradually decreased as a function of the level of social misalignment for the 0-Incon (7.066 ± 3.446 μV), 1-Incon (5.967 ± 3.621 μV), 2-Incon (6.145 ± 2.991 μV), 3-Incon (5.050 ± 3.003 μV) and 4-Incon (4.551 ± 3.260 μV) conditions. The differences among these conditions were significant,

except that between the 1-Incon and the 2-Incon conditions ($P = 0.584$). However, neither the main effect of 'uncertain context' [$F(1,33) < 1$, $P = 0.383$, $\eta^2_p = 0.023$] nor the 'uncertain context' × 'social misalignment' interaction was significant [$F(4,33) < 1$, $P = 0.197$, $\eta^2_p = 0.006$] (Figure 5).
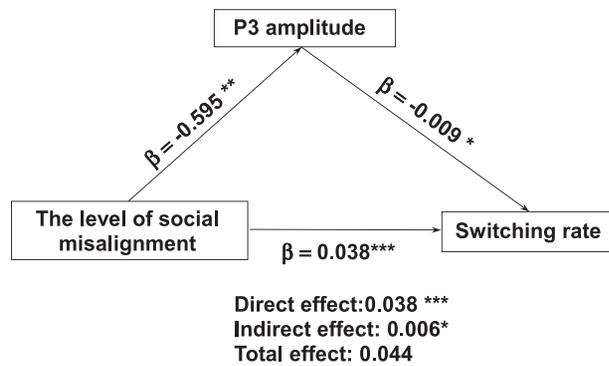
Fig. 6. Results of the mediation analysis. A significant partial mediation effect is observed on the P3 but not the N1 and FRN.

The mediation role of the P3 component. To further explore the mechanism underlying the modulation of behavioral switching by social misalignment, mediation models were analyzed using a bootstrapping procedure (MacKinnon *et al.*, 2007). As shown in Figure 6, the direct effect of social misalignment [β = 0.038, standard error (SE) = 0.008, *t* = 4.723, *P* < 0.001], the effect of social misalignment on the P3 amplitude (β = −0.615, *t* = −3.347, *P* = 0.001), and the relationship between the P3 amplitude and switching rate [β = −0.010, SE = 0.003, 95% CI (−0.016, −0.004)] were all statistically significant. In addition, the indirect effect of social misalignment on the switching rate was significant [β = 0.006, SE = 0.003, 95% CI (0.002, 0.015)]. The total effect was 0.044, which included indirect (0.006) and direct effects (0.038). However, no mediation effect was found on the N1 component and FRN (Supplementary Figures S1 and S2 for more details). In short, only the P3 amplitude in response to feedback presentation partially mediated the relationship between social misalignment and the switching rate.

Discussion

The way individuals react to social misalignment plays an important role in social interactions, mainly because fitting in with the majority is a helpful strategy to reduce uncertainties in various environments. Combining a modified social conformity task with the ERP technique, the current study aimed to investigate the behavioral patterns and neural correlates associated with social misalignment during decision-making under uncertainty. Aided by the high time resolution of the ERPs, our study found a hierarchical processing pattern related to social misalignment. Any mismatch between self and others was detected at an early stage of information processing (indexed by a higher N1). After this stage, the FRN amplitude enlarged gradually as a function of the level of misalignment, and it reached its maximum when the individual decision deviated from that of half (or more) of the players. Finally, different misalignment levels were further differentiated at a late deliberate stage (labeled by the P3); a higher P3 amplitude was found when the decisions of the participants were aligned with those of more players. Moreover, the aforementioned hierarchical processing pattern of social misalignment can be generalized from risky to ambiguous decision-making.

Our behavioral results have replicated the classical behavioral pattern in social conformity studies; participants switched more frequently when they were misaligned with the majority (Chen *et al.*, 2012; Schnuerch *et al.*, 2014). Furthermore, we found that the initial preference strength was negatively correlated with behavioral adjustment during misalignment (indexed by switching rate), which indicates that people with a stronger decision preference are less affected by social misalignment and eventually make fewer behavioral adjustments (Chung *et al.*, 2015).

For the ERP analysis, the N1, FRN and P3 components were selected and used to investigate the temporal processing of social misalignment across risky and ambiguous contexts. Our results showed that the N1 amplitude was sensitive to misalignment. Specifically, N1 was attenuated when individual decisions were aligned with those of all other players, but there was no N1 difference for the remaining misaligned conditions. The N1 is considered to reflect an attentional discrimination process; attended stimuli evoke a greater N1 than unattended ones (Vogel and Luck, 2000; Hopf *et al.*, 2002). Researchers have also observed an N1 effect during social decision-making, which indicates the top-down influence of social factors on the early processes of stimulus differentiation (Schindler *et al.*, 2014; Baess and Prinz, 2015). In our opinion, a higher N1 evoked under misaligned conditions indicates an early attention bias in situations of mismatch with others. This idea is in line with recent findings that attentional resources for stimulus identification and discrimination increase under social pressure (Trautmann-Lengsfeld and Herrmann, 2013; Germar *et al.*, 2016; Zanesco *et al.*, 2019). In short, the N1 component, which reflects an early attentional discrimination process, shows a strong sensitivity for detecting any misalignment with others.

This study also found that the FRN amplitude increased as a function of the misalignment level until half (i.e. two) of the players showed disagreement with a participant, but there was no significant difference in situations where their decisions deviated from those of more than two players. Considering that the FRN is widely regarded as a neural index for the error-monitoring system associated with expectation violation (Montague and Lohrenz, 2007; Harris and Fiske, 2010; Ferdinand *et al.*, 2012), we suggest that a higher FRN evoked by social misalignment reflects a stronger violation of prior expectation since people generally tend to expect that they would fit in with the majority (Nakahashi, 2007). Our FRN finding is also consistent with the findings from previous ERP studies (Chen *et al.*, 2012; Kim *et al.*, 2012). Notably, the incremental FRN amplitude during low misaligned conditions (i.e. disagreeing with less than half of other players) indicates that a small violation of social expectation could be differentiated by brain activity but not by behavioral responses (Figure 2A). This finding supports the idea that ERPs can uncover brain activities even when there are no behavioral changes (Amodio *et al.*, 2014; L. Wu *et al.*, 2016). Meanwhile, there was no significant FRN difference across conditions where individual decisions mismatched with those of more than half of the other players, indicating that a disagreement with the majority is a complete deviation from their expectations (i.e. a ceiling effect). The interpretation here is also in line with the false consensus effect mentioned previously (Bauman and Geher, 2002; Lapinski and Rimal, 2005). In addition, we observed the N1 and FRN effects in different uncertain contexts: a higher N1 was elicited in the ambiguous context than in the risky context, whereas the FRN showed the opposite pattern. In our opinion, these results may have implications for how the brain resolves uncertainties. Compared with the risky context, the human brain may be more alert in the ambiguous context because the absence of probability information leads to more uncertain consequences (Hsu *et al.*, 2005) that manifest as a higher N1. Meanwhile, the confidence of participants

concerning whether their behaviors would be aligned with others may have decreased in the ambiguous context, of which the level of uncertainty was higher; thus, the same level of social misalignment elicited a lower FRN (indicating a weaker signal of expectation violation) (Park *et al.*, 2017).

Following the FRN, an enhanced P3 was found as the level of social alignment increased gradually, while no P3 difference was detected in the risky and ambiguous contexts. The P3 component is sensitive to reward magnitude, and a higher P3 evoked by a greater reward indicates a stronger emotional significance of that reward (Yeung and Sanfey, 2004; Polich, 2007; Polezzi *et al.*, 2010). The current P3 results suggest that being congruent with the majority carries a rewarding message that leads to positive feelings and encourages individuals to maintain their initial selection. Corresponding to this interpretation, we found that the P3 amplitude partially mediated the relationship between social misalignment and subsequent behavioral adjustments; a higher P3 elicited by social alignment led to a lower likelihood for participants changing their initial decision. Our interpretation is not only supported by previous ERP results that social conformity is emotionally positive and informative (Wu *et al.*, 2011; Yu and Sun, 2013), it is also consistent with neuroimaging studies in which an individual selection endorsed by others evokes stronger activations of the brain's reward network (e.g. the VS; see Zaki *et al.*, 2011; Wu *et al.*, 2016). The current P3 finding also indicates that people show a sense of satisfaction when they experience uniformity and connectedness with the majority (Shamay-Tsoory *et al.*, 2019). Taking the mediation effect of P3 amplitude into account, we believe that the P3 component is a promising neural index when combined with behavioral prediction during social misalignment.

It is worth noting that the P3 amplitude differentiated between almost all the levels of social misalignment, but this neural index was insensitive to uncertainty. The distinguishable P3 effect on social misalignment was consistent with previous findings indicating that P3 is associated with the late-stage elaborate processing of affective evaluation (Wu and Zhou, 2009; San Martín, 2012). Meanwhile, we should interpret this insensitivity to uncertainty with caution. In our opinion, when social misalignment occurs, people are prone to rely more on social feedback (i.e. others' choices) rather than explicit probability information to make their own decisions. Therefore, the processing pattern for social misalignment does not change in the risk and ambiguous contexts (i.e. regardless of whether probability information is available) during the late and elaborate stages.

As mentioned above, our ERP results show a hierarchical processing pattern of social misalignment in both risky and ambiguous decision-making contexts. To our knowledge, most previous studies focused on how social factors affect risky decision-making while ignoring ambiguous decision-making (e.g. Chung *et al.*, 2015; Knoll *et al.*, 2015; Reiter *et al.*, 2019). It is not clear whether previous findings for the risky context can be generalized to the ambiguous context since there is a well-documented inherent distinction between these two forms of uncertainty (Hsu *et al.*, 2005; Krain *et al.*, 2006; Blankenstein *et al.*, 2017). Given this background knowledge, our findings not only highlight the robustness of the hierarchical neural sensitivity to social misalignment but also shed light on the relationship between social influence and different forms of uncertain decision-making.

The following limitations should be noted. First, this study did not find an interaction effect between social misalignment and uncertain contexts (risky/ambiguous) in time domain. Second, our study did not test the mediation of the neural representation of social misalignment by personality traits (e.g. narcissism) and the mechanism underlying it, although personality may modulate individual responses to social influence (e.g. Yang *et al.*, 2018). Third, while the group size was fixed (i.e. five persons in total) in the current study, it is necessary to vary it to test the reliability of our current findings. Follow-up studies are needed to investigate these issues.

In conclusion, using the ERP technique with high time resolution, we demonstrated how the human brain distinguishes between different levels of social misalignment across uncertain contexts. In light of these findings, we demonstrate the hierarchical sensitivity to social misalignment across uncertain contexts; during the first stage of information processing, humans focus on detecting any mismatch between self and others (indexed by the N1); during the second stage, they become sensitive to low levels of misalignment (indexed by the FRN); finally, the specific levels of social misalignment are processed elaborately during the third stage (indexed by P3). Moreover, the hierarchical processing pattern reflected by the neural activity can be generalized from risky to ambiguous decision-making. Overall, these findings facilitate a better understanding of the neural underpinnings of social misalignment during decision-making under uncertainty.

Conflict of interest

The authors have declared that there is no conflict of interest in relation to the subject of this study.

Authors' contributions

Y.L., R.G. and S.L. conceived the research. Y.L. and R.G. designed and performed the experiment. Y.L. analyzed the data. Y.L., R.G., S.L., S.Q., L.H. and Y.J.L. wrote the manuscript.

Declaration of ethics

All procedures performed in this study were in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Statement of data processing

In response to the Standard Reviewer Disclosure Request (2014) endorsed by the Center for Open Science (http://osf.io/hadz3), the authors confirm that in this manuscript, they have reported all measures, conditions, data exclusions and how they determined their sample sizes.

Data and code availability

The data and code of this study would be available upon request and with approvals of Beijing Normal University. More information on making this request can be obtained from the first author Yongling Lin (linyl@mail.bnu.edu.cn).

Supplementary data

Supplementary data are available at *SCAN* online.

References

Amodio, D.M., Bartholow, B.D., Ito, T.A. (2014). Tracking the dynamics of the social brain: ERP approaches for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, **9**(3), 385–93.

Baess, P., Prinz, W. (2015). My partner is also on my mind: social context modulates the N1 response. *Experimental Brain Research*, **233**(1), 105–13.

Bauman, K.P., Geher, G. (2002). We think you agree: the detrimental impact of the false consensus effect on behavior. *Current Psychology*, **21**(4), 293–318.

Berns, G.S., Chappelow, J., Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, **58**(3), 245–53.

Blankenstein, N.E., Peper, J.S., Crone, E.A., van Duijvenvoorde, A.C. (2017). Neural mechanisms underlying risk and ambiguity attitudes. *Journal of Cognitive Neuroscience*, **29**(11), 1845–59.

Campbell-Meiklejohn, D.K., Bach, D.R., Roepstorff, A., Dolan, R.J., Frith, C.D. (2010). How the opinion of others affects our valuation of objects. *Current Biology*, **20**(13), 1165–70.

Charpentier, C.J., Moutsiana, C., Garrett, N., Sharot, T. (2014). The brain's temporal dynamics from a collective decision to individual action. *Journal of Neuroscience*, **34**(17), 5816–23.

Chen, J., Wu, Y., Tong, G., Guan, X., Zhou, X. (2012). ERP correlates of social conformity in a line judgment task. *BMC Neuroscience*, **13**, 43. http://www.ncbi.nlm.nih.gov/pubmed/22554347.

Chung, D., Christopoulos, G.I., King-Casas, B., Ball, S.B., Chiu, P.H. (2015). Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. *Nature Neuroscience*, **18**(6), 912–6.

Chung, D., Orloff, M.A., Lauharatanahirun, N., Chiu, P.H., King-Casas, B. (2020). Valuation of peers' safe choices is associated with substance-naïveté in adolescents. *Proceedings of the National Academy of Sciences of the United States of America*, **117**(50), 31729–37.

Cialdini, R.B., Goldstein, N.J. (2004). Social influence: compliance and conformity. *Annual Review of Psychology*, **55**, 591–621.

Cohen, J., Polich, J. (1997). On the number of trials needed for P300. *International Journal of Psychophysiology*, **25**(3), 249–55.

Dawes, R.M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology*, **25**(1), 1–17.

Delorme, A., Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, **134**(1), 9–21.

Edelson, M., Sharot, T., Dolan, R.J., Dudai, Y. (2011). Following the crowd: brain substrates of long-term memory conformity. *Science*, **333**(6038), 108–11.

Falk, E.B., Berkman, E.T., Whalen, D., Lieberman, M.D. (2011). Neural activity during health messaging predicts reductions in smoking above and beyond self-report. *Health Psychology*, **30**(2), 177.

Falk, E.B., Berkman, E.T., Lieberman, M.D. (2012). From neural responses to population behavior: neural focus group predicts population-level media effects. *Psychological Science*, **23**(5), 439–45.

Feng, C., Cao, J., Li, Y., Wu, H., Mobbs, D. (2018). The pursuit of social acceptance: aberrant conformity in social anxiety disorder. *Social Cognitive and Affective Neuroscience*, **13**(8), 809–17.

Ferdinand, N.K., Mecklinger, A., Kray, J., Gehring, W.J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *Journal of Neuroscience*, **32**(35), 12087–92.

Gehring, W.J., Willoughby, A.R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, **295**(5563), 2279–82.

Germar, M., Albrecht, T., Voss, A., Mojzisch, A. (2016). Social conformity is due to biased stimulus processing: electrophysiological and diffusion analyses. *Social Cognitive and Affective Neuroscience*, **11**(9), 1449–59.

Gu, R., Lei, Z., Broster, L., Wu, T., Jiang, Y., Luo, Y.J. (2011). Beyond valence and magnitude: a flexible evaluative coding system in the brain. *Neuropsychologia*, **49**(14), 3891–7.

Harris, L.T., Fiske, S.T. (2010). Neural regions that underlie reinforcement learning are also active for social expectancy violations. *Social Neuroscience*, **5**(1), 76–91.

Hayes, A.F., Rockwood, N.J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: observations, recommendations, and implementation. *Behaviour Research and Therapy*, **98**, 39–57.

Heerdink, M.W., Van Kleef, G.A., Homan, A.C., Fischer, A.H. (2013). On the social influence of emotions in groups: interpersonal effects of anger and happiness on conformity versus deviance. *Journal of Personality and Social Psychology*, **105**(2), 262–84.

Hopf, J.M., Vogel, E., Woodman, G., Heinze, H.J., Luck, S.J. (2002). Localizing visual discrimination processes in time and space. *Journal of Neurophysiology*, **88**(4), 2088–95.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, **310**(5754), 1680–3.

Hsu, M., Krajbich, I., Zhao, C., Camerer, C.F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *Journal of Neuroscience*, **29**(7), 2231–7.

Huang, Y., Kendrick, K.M., Yu, R. (2014). Social conflicts elicit an N400-like component. *Neuropsychologia*, **65**, 211–20.

Ibanez, A., Melloni, M., Huepe, D., *et al.* (2012). What event-related potentials (ERPs) bring to social neuroscience? *Social Neuroscience*, **7**(6), 632–49.

Izuma, K. (2013). The neural basis of social influence and attitude change. *Current Opinion in Neurobiology*, **23**(3), 456–62.

Izuma, K., Adolphs, R. (2013). Social manipulation of preference in the human brain. *Neuron*, **78**(3), 563–73.

Kim, B.R., Liss, A., Rao, M., Singer, Z., Compton, R.J. (2012). Social deviance activates the brain's error-monitoring system. *Cognitive, Affective & Behavioral Neuroscience*, **12**(1), 65–73.

Klucharev, V., Hytonen, K., Rijpkema, M., Smidts, A., Fernandez, G. (2009). Reinforcement learning signal predicts social conformity [Research Support, Non-U.S. Gov't]. *Neuron*, **61**(1), 140–51.

Klucharev, V., Munneke, M.A., Smidts, A., Fernandez, G. (2011). Downregulation of the posterior medial frontal cortex prevents social conformity. *Journal of Neuroscience*, **31**(33), 11934–40.

Knoll, L.J., Magis-Weinberg, L., Speekenbrink, M., Blakemore, S.-J. (2015). Social influence on risk perception during adolescence. *Psychological Science*, **26**(5), 583–92.

Knutson, B., Rick, S., Wirmmer, G.E., Prelec, D., Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, **53**(1), 147–56.

Krain, A.L., Wilson, A.M., Arbuckle, R., Castellanos, F.X., Milham, M.P. (2006). Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *Neuroimage*, **32**(1), 477–84.

Lapinski, M.K., Rimal, R.N. (2005). An explication of social norms. *Communication Theory*, **15**(2), 127–47.

Leue, A., Klein, C., Lange, S., Beauducel, A. (2013). Inter-individual and intra-individual variability of the N2 component: on reliability and signal-to-noise ratio. *Brain and Cognition*, **83**(1), 61–71.

Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, **103**(2), 1036–47.

Lin, Y., Duan, L., Xu, P., Li, X., Gu, R., Luo, Y. (2019). Electrophysiological indexes of option characteristic processing. *Psychophysiology*, **56**(10), e13403.

Luck, S.J., Woodman, G.F., Vogel, E.K. (2000). Event-related potential studies of attention. *Trends in Cognitive Science*, **4**(11), 432–40.

Luck, S.J. (2014). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Luck, S.J., Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, **54**(1), 146–57.

Luo, Y., Feng, C., Wu, T., *et al.* (2015). Social comparison manifests in event-related potentials. *Scientific Report*, **5**(2), 1–9.

MacKinnon, D.P., Fritz, M.S., Williams, J., Lockwood, C.M. (2007). Distribution of the product confidence limits for the indirect effect: program PRODCLIN. *Behavior Research Methods*, **39**(3), 384–9.

Marco-Pallares, J., Cucurell, D., Munte, T.F., Strien, N., Rodriguez-Fornells, A. (2011). On the number of trials needed for a stable feedback-related negativity. *Psychophysiology*, **48**(6), 852–60.

Masaki, H., Takeuchi, S., Gehring, W.J., Takasawa, N., Yamazaki, K. (2006). Affective-motivational influences on feedback-related ERPs in a gambling task. *Brain Research*, **1105**, 110–21.

Montague, P.R., Lohrenz, T. (2007). To detect and correct: norm violations and their enforcement. *Neuron*, **56**(1), 14–8.

Morgan, T.J., Laland, K.N., Harris, P.L. (2015). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental Science*, **18**(4), 511–24.

Mu, Y., Kitayama, S., Han, S., Gelfand, M.J. (2015). How culture gets embrained: cultural differences in event-related potentials of social norm violations. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(50), 15348–53.

Nakahashi, W. (2007). The evolution of conformist transmission in social learning when the environment changes periodically. *Theoretical Population Biology*, **72**(1), 52–66.

Nakahashi, W., Wakano, J.Y., Henrich, J. (2012). Adaptive social learning strategies in temporally and spatially varying environments: how temporal vs. spatial variation, number of cultural traits, and costs of learning influence the evolution of conformist-biased transmission, payoff-biased transmission, and individual learning. *Human Nature*, **23**(4), 386–418.

Nosek, B.A., Simonsohn, U., Moore, D.A., *et al.* (2014). Standard reviewer statement for disclosure of sample, conditions, measures, and exclusions. Open Science Framework. Available: https://osf.io/hadz3/ [February 4, 2014, Date accessed].

Park, S.A., Goiame, S., O'Connor, D.A., Dreher, J.C. (2017). Integration of individual and social information for decision-making in groups of different sizes. *PLoS Biology*, **15**(6), e2001958.

Polezzi, D., Sartori, G., Rumiati, R., Vidotto, G., Daum, I. (2010). Brain correlates of risky decision-making. *Neuroimage*, **49**(2), 1886–94.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, **118**(10), 2128–48.

Pushkarskaya, H., Tolin, D., Ruderman, L., *et al.* (2015). Decision-making under uncertainty in obsessive–compulsive disorder. *Journal of Psychiatric Research*, **69**, 166–73.

Raafat, R.M., Chater, N., Frith, C. (2009). Herding in humans. *Trends in Cognitive Sciences*, **13**(10), 420–8.

Reiter, A.M., Suzuki, S., O'Doherty, J.P., Li, S.-C., Eppinger, B. (2019). Risk contagion by peers affects learning and decision-making in adolescents. *Journal of Experimental Psychology: General*, **148**(9), 1494.

Rendell, L., Boyd, R., Cownden, D., *et al.* (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, **328**(5975), 208–13.

Ross, M., Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, **37**(3), 322–36.

Rothschild, M., Stiglitz, J.E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, **2**(3), 225–43.

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, **6**, 304.

Santamaría-García, H., Pannunzi, M., Ayneto, A., Deco, G., Sebastián-Gallés, N. (2014). 'If you are good, I get better': the role of social hierarchy in perceptual decision-making. *Social Cognitive and Affective Neuroscience*, **9**(10), 1489–97.

Scheepers, D., Derks, B. (2016). Revisiting social identity theory from a neuroscience perspective. *Current Opinion in Psychology*, **11**, 74–8.

Schindler, S., Wegrzyn, M., Steppacher, I., Kissler, J. (2014). It's all in your head - how anticipating evaluation affects the processing of emotional trait adjectives. *Frontiers in Psychology*, **5**, 1292.

Schnuerch, R., Trautmann-Lengsfeld, S.A., Bertram, M., Gibbons, H. (2014). Neural sensitivity to social deviance predicts attentive processing of peer-group judgment. *Social Neuroscience*, **9**(6), 650–60.

Schnuerch, R., Richter, J., Koppehele-Gossel, J., Gibbons, H. (2016). Multiple neural signatures of social proof and deviance during the observation of other people's preferences. *Psychophysiology*, **53**(6), 823–36.

Schnuerch, R., Gibbons, H. (2015). Social proof in the human brain: electrophysiological signatures of agreement and disagreement with the majority. *Psychophysiology*, **52**(10), 1328–42.

Shamay-Tsoory, S.G., Aharon-Peretz, J., Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, **132**(Pt 3), 617–27.

Shamay-Tsoory, S.G., Saporta, N., Marton-Alper, I.Z., Gvirts, H.Z. (2019). Herding brains: a core neural mechanism for social alignment. *Trends in Cognitive Sciences*, **23**(3), 174–86.

Shestakova, A., Rieskamp, J., Tugin, S., Ossadtchi, A., Krutitskaya, J., Klucharev, V. (2013). Electrophysiological precursors of social conformity. *Social Cognitive and Affective Neuroscience*, **8**(7), 756–63.

Suzuki, S., Jensen, E.L., Bossaerts, P., O'Doherty, J.P. (2016). Behavioral contagion during learning about another agent's risk-preferences acts on the neural representation of decision-risk. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(14), 3755–60.

Toelch, U., Dolan, R.J. (2015). Informational and normative influences in conformity from a neurocomputational perspective. *Trends in Cognitive Sciences*, **19**(10), 579–89.

Trautmann-Lengsfeld, S.A., Herrmann, C.S. (2013). EEG reveals an early influence of social conformity on visual processing in group pressure situations. *Social Neuroscience*, **8**(1), 75–89.

Tump, A.N., Pleskac, T.J., Kurvers, R.H. (2020). Wise or mad crowds? The cognitive mechanisms underlying information cascades. *Science Advances*, **6**(29), eabb0266.

Tversky, A., Fox, C.R. (1995). Weighing risk and uncertainty. *Psychological Review*, **102**(2), 269.

Tymula, A., Belmaker, L.A.R., Roy, A.K., *et al.* (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(42), 17135–40.

Vazire, S. (2016). Editorial. *Social Psychological and Personality Science*, **7**(1), 3–7.

Vogel, E.K., Luck, S.J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, **37**(2), 190–203.

Wang, G., Suemitsu, K. (2007). Object recognition learning differentiates the representations of objects at the ERP component N1. *Clinical Neurophysiology*, **118**(2), 372–80.

Wu, H., Luo, Y., Feng, C. (2016). Neural signatures of social conformity: a coordinate-based activation likelihood estimation meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, **71**, 101–11.

Wu, L., Gu, R., Cai, H., Zhang, J. (2016). Electrophysiological evidence for executive control and efficient categorization involved in implicit self-evaluation. *Social Neuroscience*, **11**(2), 153–63.

Wu, Y., Leliveld, M.C., Zhou, X. (2011). Social distance modulates recipient's fairness consideration in the dictator game: an ERP study. *Biological Psychology*, **88**(2–3), 253–62.

Wu, Y., Zhou, X.L. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, **1286**, 114–22.

Xie, Y., Chen, M., Lai, H., Zhang, W., Zhao, Z., Anwar, C.M. (2016). Neural basis of two kinds of social influence: obedience and conformity. *Frontiers in Human Neuroscience*, **10**, 51.

Xiong, W., Gao, X., He, Z., Yu, H., Liu, H., Zhou, X. (2020). Affective evaluation of others' altruistic decisions under risk and ambiguity. *Neuroimage*, **218**, 116996.

Yang, Z., Sedikides, C., Gu, R., *et al.* (2018). Communal narcissism: social decisions and neurophysiological reactions. *Journal of Research in Personality*, **76**, 64–73.

Yeung, N., Sanfey, A.G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, **24**(28), 6258–64.

Yu, H., Dan, M., Ma, Q., Jin, J. (2018). They all do it, will you? Event-related potential evidence of herding behavior in online peer-to-peer lending. *Neuroscience Letters*, **681**, 1–5.

Yu, R., Sun, S. (2013). To conform or not to conform: spontaneous conformity diminishes the sensitivity to monetary outcomes. *PLoS One*, **8**(5), e64530.

Zaki, J., Schirmer, J., Mitchell, J.P. (2011). Social influence modulates the neural computation of value [Research Support, U.S. Gov't, Non-P.H.S.]. *Psychological Science*, **22**(7), 894–900.

Zanesco, J., Tipura, E., Posada, A., Clement, F., Pegna, A.J. (2019). Seeing is believing: early perceptual brain processes are modified by social feedback. *Social Neuroscience*, **14**(5), 519–29.

Zhang, D., Gu, R., Wu, T., *et al.* (2013). An electrophysiological index of changes in risk decision-making strategies. *Neuropsychologia*, **51**(8), 1397–407.

Zhang, L., Gläscher, J. (2020). A brain network supporting social influences in human decision-making. *Science Advances*, **6**(34), eabb4159.

Zhu, R., Wu, H., Xu, Z., *et al.* (2019). Early distinction between shame and guilt processing in an interpersonal context. *Social Neuroscience*, **14**(1), 53–66.