International Journal of Psychophysiology 96 (2015) 183–190

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

You did not mean it: Perceived good intentions alleviate sense of unfairness

Qingguo Ma [a,b,1], Liang Meng [a,b,1], Zhexiao Zhang [c], Qing Xu [a,b], Yue Wang [d], Qiang Shen [a,b,e,*]

[a] School of Management, Zhejiang University, Hangzhou, China
[b] Neuromanagement Lab, Zhejiang University, Hangzhou, China
[c] College of Economics, Zhejiang University, Hangzhou, China
[d] Neuroengineering Lab, Mechanical Engineering Department, National University of Singapore, Singapore
[e] National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

A R T I C L E   I N F O

Article history:
Received 9 November 2014
Received in revised form 16 March 2015
Accepted 17 March 2015
Available online 8 April 2015

Keywords:
Fairness
Intention
Ultimatum game
Social norm
Event-related potentials
FRN
P300

A B S T R A C T

Previous research has reported that feedback-related negativity (FRN) may represent the degree of perceived unfairness in the ultimatum game (UG). However, few studies have incorporated intention-related consideration in examining the neural correlates of fairness perception. To address this issue, the present study introduced an intentional UG paradigm to disentangle the effect of perceived intention from fairness concerns, using an event-related potential (ERP) analysis. Consistent with the hypothesis, the behavioral results indicated that good intention could markedly reduce rejection rates, and this intention effect was modulated by the degree of fairness, which was more prominent under unfair scenarios. Further electrophysiological results showed that, for the unfair division schemes, FRN and P300 amplitudes were significantly different between offers proposed with good intention and those with bad intention, while such discrepancies were not observed for the fair condition. In summary, converging results demonstrated that perceived intention can modulate the effect of fairness in social decision-making.

1. Introduction

The capability to evaluate and respond to social behaviors in a quick and efficient manner is deeply rooted in human nature. Nevertheless, during dynamic social interactions, individuals are not merely concerned with objective outcomes. Instead, they may go beyond that concern and evaluate people's actions leading to specific social outcomes, as well as the perceived intentions behind such actions. For example, in a recent study (Ames and Fiske, 2013), participants were asked to read a vignette that describes a small company in which employees' salaries partially depended on its profits. Importantly, the CEO made a poor investment at a cost to the employees' compensation. When participants were told that the investment was considered to have great prospect, and the CEO did not expect that the investment would fail, they thought that the employees were less harmed. By contrast, when participants learned that the CEO intentionally made that poor investment to urge the employees to work even harder to increase future profits, they deemed that the employees were harmed to a greater extent. Therefore, this study showed that interest-neutral third parties consider intended social harms to be much worse than unintended ones.

Besides serving as third parties, we are quite often directly involved in similar scenarios in social contexts. Social pain can be induced internally when we are either economically or psychologically harmed by others, and neuroscientific studies have suggested that social pain shares common neural mechanisms with primary physical pain (Eisenberger, 2012). By contrast, when we feel loved by and socially connected to others, social warmth is experienced. In a similar manner, the underlying neural representation of social warmth is found to be similar with that of physical warmth (Inagaki and Eisenberger, 2013). It is worth noting that in daily life, there are occasions when others' actions are not explicit, and we can only infer their intentions from observed outcomes. Sometimes, we mistakenly speculate others' intentions and blame others unjustly, which may even cost us a precious friendship. However, when we later identify the underlying truth, we may feel deeply guilty and regretful. This further suggests that perceived intentions play an important role in the evaluation of outcomes as well as the judgment of action agents.

In the present study, we modified the Ultimatum game (UG) (Güth et al., 1982) and engaged participants in social interactions in which monetary payoffs are distributed between two players. In the standard UG task, the proposer has the privilege of dividing the stake between himself/herself and the paired anonymous responder. Once the offer is

* Corresponding author at: School of Management, Zhejiang University, Hangzhou, China. Tel.: +86 571 88206833; fax: +86 571 87995372.
E-mail address: johnsonzhj@gmail.com (Q. Shen).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ijpsycho.2015.03.011

proposed, the responder can either accept it or reject it. While acceptance would lead to the division of the stake as proposed, rejection of the offer would leave both the proposer and the responder empty-handed. In this modification, both the outcome and the other player's action leading to the outcome are revealed, which may provide a valuable cue for the participants to infer the intention of others. We orthogonally manipulated outcomes and the perceived intentions of others and examined participants' behavioral and neural responses toward varied outcomes, especially when the outcome itself and the perceived intention were in conflict (either bad outcome as a result of good intention, or good outcome resulting from bad intention). Specifically, we intended to investigate whether perceived good intentions would alleviate the social pain brought by unsatisfactory outcomes.

In regard to asset allocation, both lab-based experiments and field studies have shown that individuals do not merely care about their own payoffs. In contrast, they are also sensitive to self-others' comparisons, and fairness plays a key role in such considerations. As a long-established social norm, a sense of fairness is prevalent in various circumstances, from sports competition to promotions at work, and it is generally universal across diverse cultures (Henrich et al., 2006). Defined as the phenomenon of inequity aversion, violation of the social norm of fairness can elicit negative emotions and give rise to subsequent strong reactions, including punishment or revenge toward unequal distributions of resources (Fehr and Schmidt, 1999). UG is widely used in the fields of economics and psychology to examine people's responses to unfairness. In recent years, in the emerging interdisciplinary field of neuroeconomics, several studies have explored the neural underpinnings of fairness perception and evaluation. For example, using fMRI, Sanfey et al. (2003) examined the responders' reactions toward unfair offers and found that the evaluation of unfair offers resulted in increased brain activations in the bilateral insula, dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC).

Beyond the degrees of fairness of actual outcomes, previous behavioral findings also reported that people's reactions were highly dependent on perceived intentions of others (Falk et al., 2008; Nelson, 2002; Radke et al., 2012; Sutter, 2007). For example, Blount (1995) discovered that people tend to accept an unfair offer proposed by a computer rather than a human agent, as they perceived the unfair offer with the same payoff to be unintentional under the first scenario. However, until now, little attention has been attached to the role of perceived intentions in fairness perception, with the aim of parsing its behavioral and neural underpinnings.

To date, two fMRI studies have examined the neural correlates associated with behavioral responses to unfair offers within a modified UG paradigm (mini-UG) that manipulates the perceived intention of offers (Güroglu et al., 2010; Güroglu et al., 2011), focusing on the responder's decision-making process in such contexts. In the first study, the proposer was always faced with two alternatives, with one fixed alternative being unfair. It was discovered that brain responses underlying behavioral reactions to unfair offers depended heavily on perceived intentions. Specifically, the insula and the dorsal medial prefrontal cortex were activated both when an unfair offer was rejected in the non-alternative condition and when the same offer was accepted in the fair- or hyperfair-alternative condition. However, it is worth mentioning that because subjects were presented with the offer and the alternatives at the same time, perception of the proposer's intention and evaluation of an offer's fairness were intertwined in these studies.

To tease apart the cognitive processes of intention and fairness perception, we developed a revised version of the mini-UG with orthogonally manipulated intention and fairness (coined as intentional UG). Subjects were asked to evaluate two offers with varying fairness (fair: 6, 4; unfair: 8, 2; throughout this paper, the former number in a division set refers to the proposer's outcome, while the latter one refers to that of the responder) in two context conditions (fairer-alternative, more unfair-alternative). In addition, to avoid potential confounds between offer evaluation and the choice action, subjects were only allowed to respond after the hints of "accept" and "reject" appeared on the screen. Based on results from prior behavioral studies (Radke et al., 2012; Sutter, 2007), we predicted that the perceived good intention of the proposer (choosing the more favorable option for the responder from the two alternatives) would effectively reduce the rejection rate at the behavioral level, especially when the offer was objectively unfair.

To our knowledge, no previous research has directly applied electrophysiological methods to examining intention-related fairness perception. In the current study, ERPs were used to obtain a better idea of temporal dynamics of brain activities as a result of the evaluation of fairness and intention. In recent years, there have been many ERP studies examining brain responses to various offers using the UG task, and the FRN has been implicated in fairness evaluation (Boksem and De Cremer, 2010; Ma et al., 2015a; Mussel et al., 2014; Polezzi et al., 2008). The FRN is a negative deflection that generally peaks between 250 ms and 350 ms at frontocentral electrodes. Compared with positive feedback, the FRN has been shown to be more pronounced for negative feedback that indicates unfavorable outcomes (San Martin, 2012).

According to reinforcement learning theory, in the non-social domain, human brains have developed specific mechanisms to detect reward prediction errors (Harris and Fiske, 2010). Similar mechanisms may have been developed to detect deviations from expected social norms in the social domain (Montague and Lohrenz, 2007; Wang et al., 2013; Wu et al., 2011). Specifically, because equal division of assets is commonly accepted as a social norm (Deutsch, 1975), unfair offers that deviate from the fairness social norm would generally induce more negative-going FRN than fair offers (Boksem and De Cremer, 2010; Ma et al., 2015a; Mussel et al., 2014; Polezzi et al., 2008). Similarly, being kind to other people and treating them well might be another widely accepted social norm, the violation of which might lead to a larger prediction error as reflected in enhanced FRN signals.

Another mainstream theory of the FRN considers that motivational significance could explain the FRN discrepancy toward losses and gains (d-FRN), suggesting that increased d-FRN will be elicited when outcomes in a certain condition bear more motivational significance to participants (Ma et al., 2014; Ma et al., 2015b; Zhou et al., 2010). In the current study, we assume that when an offer is unfair, whether it was proposed out of good intention or bad intention will be more motivationally significant. Correspondingly, in line with the motivational significance account of FRN, we predict that the good/bad intention comparison in the unfair condition might induce a larger FRN discrepancy than that in the fair condition.

2. Methods

2.1. Participants

Twenty healthy, right-handed subjects aged 19–26 years ($M = 22.17$ years, $SD = 1.92$ years), 12 of whom were male, participated in this study. All subjects were registered students of Zhejiang University and had normal or corrected-to-normal vision. They reported no history of neurological disorders or mental diseases. This study was approved by the Institutional Review Board of the Zhejiang University Neuromanagement Lab. Informed consent was obtained from all subjects before the experiment. Data from one subject were discarded because of excessive recording artifacts. In the post-experiment debriefing, another subject cast doubt on the existence of the prior behavioral study. Thus, data from 18 valid subjects were used for the final analysis.

2.2. Experimental procedure

The subjects were comfortably seated in a dimly lit, sound-attenuated and electrically shielded room. The stimuli were presented at the center of a computer screen at a distance of 100 cm, with a visual angle of $8.69° \times 6.52°$ (15.2 cm × 11.4 cm, width × height). Subjects were instructed to use the keypad to make their choices. The experiment consisted of 4 blocks, each containing 60 trials. Before the experiment, participants were informed of the rules of the intentional UG. They were told that offers from anonymous proposers were collected in a prior behavioral study, and they would play off-line with them. The computer randomly displayed all the division schemes we had designed, to guarantee that there would be equal numbers of trials in all conditions to be analyzed.

The present study modified the mini-UG paradigm (Falk et al., 2003). Similar to the manipulation by Falk and his colleagues, the proposer always has two alternatives to choose from. However, as shown in Fig. 1, there are two target allocation schemes in the present task rather than only the unfair division of the stake (8, 2) in the original mini-UG design. The other target allocation scheme is a comparatively fair division (6, 4), according to which the proposer would get ¥6, while the responder would get the remaining ¥4. Beyond these, there are three other schemes, which are (5, 5), (7, 3) and (9, 1), respectively. The intention manipulation is associated with the alternative scheme along with the two target schemes. When the target scheme is paired with an advantageous alternative for the proposer, the act of choosing the target scheme would be deemed as being of good intention, and vice versa.

The four different pairs of proposals, (5, 5; 6, 4), (6, 4; 7, 3), (7, 3; 8, 2) and (8, 2; 9, 1), each appeared in 60 rounds during the experiment. We also manipulated the likelihood that target allocation schemes were chosen. For each pair of options, the target scheme would be chosen in 36 rounds, whereas the other scheme would be chosen in 24 rounds. In the previous fMRI studies, target schemes were twice as likely to be chosen compared with non-target schemes (Güroglu et al., 2010; Güroglu et al., 2011). Only trials in which target schemes were chosen went into the behavioral and ERP analyses (see Table 1).

As illustrated in Fig. 2, a fixation appeared as a cue for 800 ms on the blank screen, indicating the beginning of each trial. The given name of the proposer was presented afterwards. After the division scheme set was shown for 1500 ms, the proposer's choice was highlighted. Subjects could respond after the hints of "accept" and "reject" appeared on the screen. If they decided to accept the division scheme, then ¥10 would be divided as proposed. Otherwise, both the proposer and the responder would receive no compensation in that round.

Table 1
Number of trials and offers made in the experiment.

Allocation set	Condition/offer	Number of trials
(5, 5; 6, 4)	5, 5	24
	Fair–bad intention (6, 4)[a]	36
(6, 4; 7, 3)	Fair–good intention (6, 4)[a]	36
	7, 3	24
(7, 3; 8, 2)	7, 3	24
	Unfair–bad intention (8, 2)[a]	36
(8, 2; 9, 1)	Unfair–Good intention (8, 2)[a]	36
	9, 1	24

[a] Target trials that are analyzed in the behavioral and ERP data.

Blank screens separating sequential stimuli within a trial lasted for 600–800 ms, and the inter-trial interval lasted for 1000–1500 ms. To familiarize subjects with the experimental procedure, a four-trial pilot experiment was administered before the experiment formally started. Subjects received ¥30 as compensation for attendance, and two trials were randomly chosen to determine their decision-based payment. At the end of the experiment, subjects were debriefed and paid accordingly. Stimuli, recording triggers, and responses were controlled by E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, USA).

2.3. EEG recordings

EEGs were recorded (band-pass 0.05 Hz to 70 Hz, sampling rate 500 Hz) from 64 scalp sites with Neuroscan Synamp2 Amplifier. The left mastoid served as an on-line reference. EEGs were off-line re-referenced to the average of the left and the right mastoid. The electrode

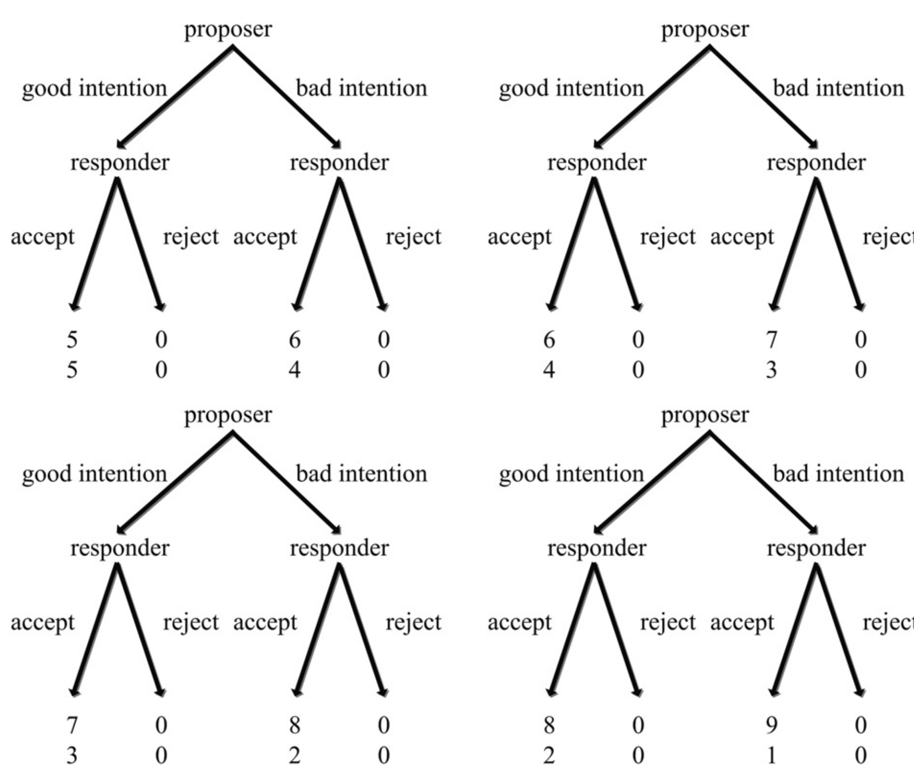

Fig. 1. The four original division sets are (5, 5; 6, 4), (6, 4; 7, 3), (7, 3; 8, 2) and (8, 2; 9, 1). Each proposer was faced with one division set and could choose from the two alternatives. (8, 2) was considered an unfair offer, while (6, 4) was deemed a fair one. If the proposer chose the advantageous offer for the responder, it was defined as the good intention condition, and vice versa. The responder can either accept or reject the offer. If an offer is rejected, the proposer and the responder each receive nothing in that round. In each condition, the proposer receives the outcome presented on the first row, while the responder receives the outcome displayed on the second row.

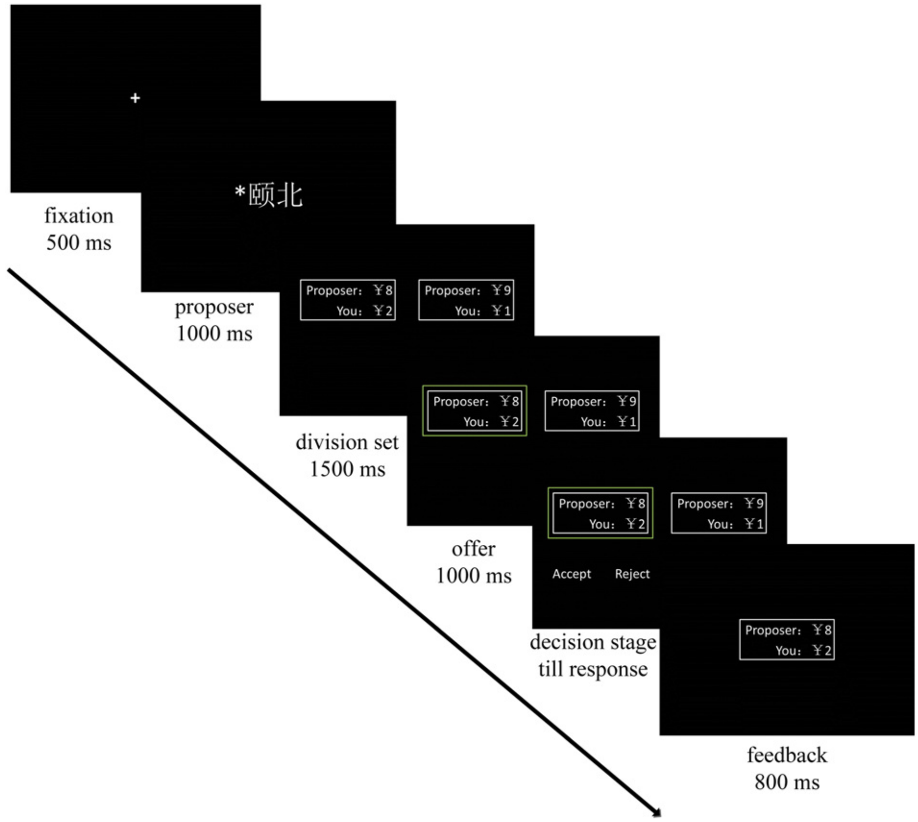

Fig. 2. Experimental task. The proposer's offer was revealed after the original division set was presented. Subjects could either accept or reject the offer. The payoffs for both the proposer and the responder in that round were presented at the feedback stage.

on the cephalic region was applied as a ground. The vertical electrooculogram (EOG) was recorded supra- and infra-orbitally at the left eye, and the horizontal EOG was recorded at the left versus right orbital rim. Electrode impedance was maintained below 5 kΩ during the experiment.

2.4. Data analysis

During the offline EEG analysis, ocular artifacts were removed, followed by digital filtering through a zero phase shift (low pass at 30 Hz, 24 dB/octave). Time windows of 200 ms before and 800 ms after onset of the offer were segmented, and the whole epoch was baseline-corrected using the 200 ms interval prior to offer onset. Trials containing amplifier clipping, bursts of electromyography activity, or peak-to-peak deflection exceeding ±80 μV were excluded. For each subject, recorded EEGs were separately averaged over each recording site under each condition. Specifically, EEG epochs were separately averaged for fairness (fair/unfair) × intention (good intention/bad intention) conditions. All of these conditions included a minimum of 32 valid trials. For the analysis of behavioral data, repeated measures ANOVA was adopted to compare rejection rates across four conditions.

Considering that the maximum FRN amplitudes appeared at frontal sites, data from the electrode sites F1, Fz, F2, FC1, FCz and FC2 were analyzed. Because the most negative peak of FRN appeared approximately 330 ms after onset of the offer, mean amplitudes in the 280–380 ms time window post-onset of feedback went into a 2 (fairness) × 2 (intention) × 6 (electrode) repeated-measures ANOVA. Given that the maximum P300 amplitudes were observed at parietal sites, data from the electrode sites C1, Cz, C2, CP1, CPz, CP2, P1, Pz, and P2 were analyzed. Because the most positive peak appeared approximately 480 ms post-onset of feedback, mean amplitudes in the time window of 380–580 ms were calculated, and ANOVA was conducted for P300 with three within-subject factors: fairness, intention and electrodes. Simple effect analysis was conducted when the interaction effect achieved significance. The Greenhouse–Geisser correction was applied in all statistical analyses when necessary.

3. Results

3.1. Behavioral performance

The rejection rates for different division schemes, presented in Fig. 3, are 5.40% (fair–bad intention), 1.08% (fair–good intention), 71.91% (unfair–bad intention) and 26.70% (unfair–good intention), respectively. A 2 (fairness) × 2 (intention) repeated-measures ANOVA revealed significant main effects of both fairness ($F_{1,\ 1} = 38.424$; $p < 0.001$; $\eta^2 = 0.693$) and intention ($F_{1,\ 1} = 24.746$; $p < 0.001$; $\eta^2 = 0.593$). The main effect of offer type suggested higher rejection rates for unfair offers than for fair ones, which interacted with intention ($F_{1,\ 1} = 13.861$; $p = 0.002$; $\eta^2 = 0.449$). Subjects reacted differently to fair and unfair offers, with significant fairness effects for both the bad intention condition ($t = 44.543$; $p < 0.001$) and the good intention condition ($t = 12.207$; $p = 0.003$). We further examined the effect of intention in fair and zunfair conditions, respectively. The intention effect is significant in the unfair condition ($t = 20.190$; $p < 0.001$) but not in the fair condition ($t = 3.278$; $p = 0.088$).

3.2. ERPs

The evaluation of division schemes is mainly reflected in the FRN and the P300. As presented in Fig. 4, the mean FRN amplitudes in the 2 (fairness) × 2 (intention) conditions are 0.602 μV (fair–good intention), 0.480 μV (fair–bad intention), 2.888 μV (unfair–good intention) and 0.345 μV (unfair–bad intention), respectively. ANOVA analysis for the FRN revealed main effects of fairness ($F_{1,\ 17} = 5.395$; $p = 0.033$; $\eta^2 = 0.241$), intention ($F_{1,\ 17} = 9.490$; $p = 0.007$; $\eta^2 = 0.358$) and

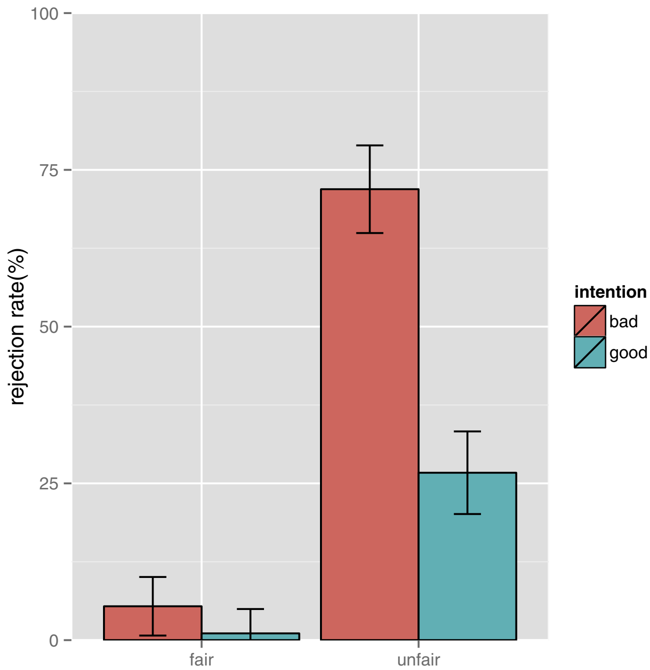

Fig. 3. Average rejection rates across the 2 (fair vs. unfair) by 2 (good intention vs. bad intention) conditions in the intentional UG task.

electrode ($F_{5,\,80} = 2.942$; $p = 0.017$; $\eta^2 = 0.658$). The interaction effect of fairness and intention is significant ($F_{1,\,17} = 6.207$; $p = 0.025$; $\eta^2 = 0.262$), with enhanced FRN amplitude toward bad intention compared with good intention when the offer is unfair ($F_{1,\,17} = 11.239$; $p = 0.004$; $\eta^2 = 0.398$) but not when the offer is fair ($F_{1,\,17} = 0.052$; $p = 0.823$; $\eta^2 = 0.003$). We further examined the fairness effect in the good and bad intention conditions, respectively; it is significant in the good intention condition ($F_{1,\,17} = 6.932$; $p = 0.017$; $\eta^2 = 0.290$) but not in the bad intention condition ($F_{1,\,17} = 0.113$; $p = 0.741$; $\eta^2 = 0.007$).

As shown in Fig. 5, the mean P300 amplitudes in the 2 (fairness) × 2 (intention) conditions are 4.311 μV (fair–good intention), 4.457 μV (fair–bad intention), 7.145 μV (unfair–good intention) and 4.960 μV (unfair–bad intention), respectively. The ANOVA for the P300 revealed main effects of fairness ($F_{1,\,17} = 10.071$; $p = 0.006$; $\eta^2 = 0.372$) and electrode ($F_{8,\,128} = 17.758$; $p < 0.001$; $\eta^2 = 0.945$). The main effect of intention approaches marginal significance ($F_{1,\,17} = 4.393$; $p = 0.051$; $\eta^2 = 0.205$). The interaction effect of fairness and intention is significant ($F_{1,\,17} = 5.003$; $p = 0.039$; $\eta^2 = 0.227$), with enhanced P300 magnitude toward good intention than bad intention when the offer is unfair ($F_{1,\,17} = 6.156$; $p = 0.024$; $\eta^2 = 0.266$) but not when the offer is fair ($F_{1,\,17} = 0.083$; $p = 0.777$; $\eta^2 = 0.005$). In addition, the main effect of fairness is not significant in the bad intention condition ($F_{1,\,17} = 0.168$; $p = 0.687$; $\eta^2 = 0.010$) but is significant in the good intention condition ($F_{1,\,17} = 12.717$; $p = 0.002$; $\eta^2 = 0.428$).

4. Discussion

As one of the most fundamental facets of social preferences, fairness is profoundly affected by the perceived intentions of others. In the present study, using an intentional UG paradigm, we investigated how perceived intentions of others influenced responders' behaviors as well as their brain responses to various offers in social interactions. The results demonstrate that good intentions produced more positive reactions at both the behavioral and neural levels. We found that participants were more likely to reject schemes when the proposer could have chosen a fairer alternative, especially when the schemes themselves were disappointing. Electrophysiologically, offers proposed out of good intentions elicited less negative-going FRN responses than the same offers proposed with bad intentions, and this effect was modulated by the degree of fairness of the proposed offer. These findings reveal the temporal dynamics of neural activity in considerations of fairness and intention, nicely complementing previous behavioral and neuroimaging studies (Güroglu et al., 2010; Sanfey et al., 2003; Tabibnia et al., 2008).

In the past, scholars have commonly adopted the mini-UG task to test whether intention matters beyond the fairness of offers. Previous researchers usually manipulated intentional consideration by applying one fixed unfair option (8, 2) along with varied alternatives (equal,

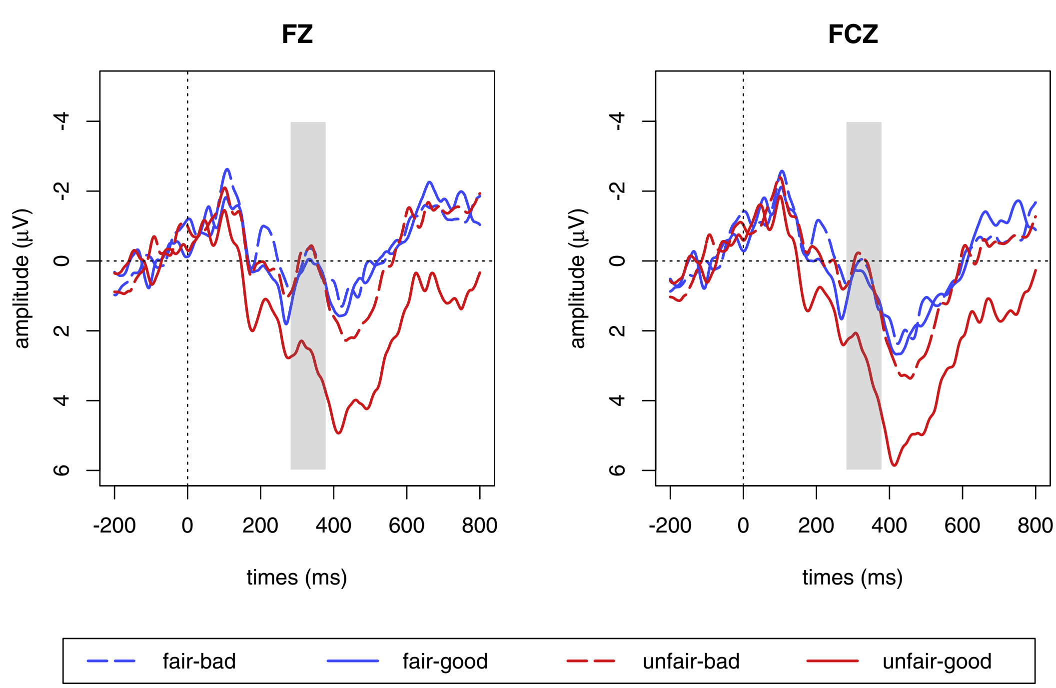

Fig. 4. FRN results during offer evaluation. For illustration, grand-averaged ERP waveforms of FRN from 2 midline frontal electrodes (Fz, FCz) are shown in relation to fairness (fair vs. unfair) and intention (good intention vs. bad intention).

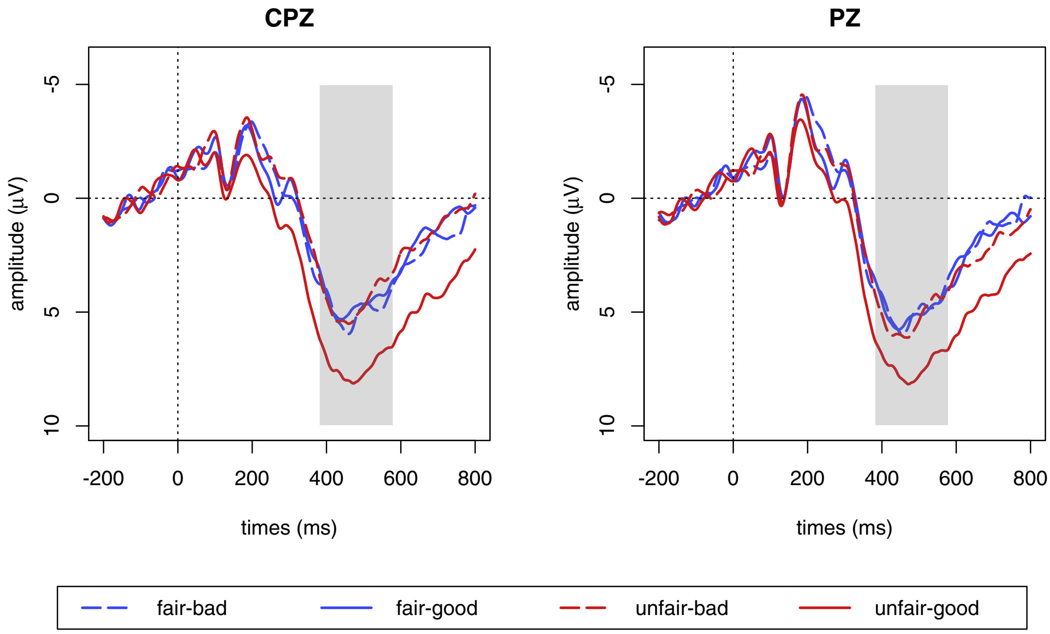

Fig. 5. P300 results during offer evaluation. For illustration, grand-averaged ERP waveforms of P300 from 2 midline frontal electrodes (CPz, Pz) are shown as a function of fairness (fair vs. unfair) and intention (good intention vs. bad intention).

fairer or more unfair options). To the best of our knowledge, no behavioral or neuroscience studies have ever tested how the perceived intention interacted with the degree of fairness. To track how responders perceive and respond to the intention of offers with varied degrees of fairness, we revised the traditional mini-UG and adopted a 2 (good/bad intention) by 2 (fair/unfair offer) experimental design to place subjects into such tradeoffs. In this design, there are two target offers with varying degrees of fairness, which are (6, 4) and (8, 2), respectively. While (8, 2) is an inherently unfair offer, (6, 4) is a comparatively fair one in the current setting. The four division sets of (5, 5; 6, 4), (6, 4; 7, 3), (7, 3; 8, 2) and (8, 2; 9, 1) make it possible for us to compare responders' reactions to fair (6, 4) and unfair (8, 2) offers with discrepant perceived intentions, which allows for the comparison of inequity aversion and intention consideration.

Behavioral results from the current study showed that subjects tended to reject unfair offers although this act would reduce their own monetary benefit. This was consistent with previous findings and further confirmed that fairness is an essential social norm in economic decision-making (Falk et al., 2003; Fehr and Schmidt, 1999). Compared with the situation in which the alternative offer is comparatively fair, rejection rates were dramatically reduced when proposers chose the responder-advantageous option from two objectively unfair division schemes. The drastic decline of rejection rates of unfair offers, from 71.91% to 26.70%, when the offer was perceived to be proposed out of good intention, demonstrated that intentional consideration indeed modulate rejection (Falk et al., 2003; Güroglu et al., 2010; Sutter, 2007). Nevertheless, the rejection rate in the unfair scenarios remained higher than that observed in the fair scenarios regardless of the perceived intention, which can be attributed to the endogenous inequity aversion of human beings (Fehr and Schmidt, 1999).

In terms of the electrophysiological data, we found a general FRN effect for fairness in the present study. At first glance, the more negative-going FRN responses to fair offers compared with unfair ones appeared to conflict with previous findings (Boksem and De Cremer, 2010; Polezzi et al., 2008). Based on visual inspection, we speculated that the reversed FRN pattern for fairness of the offer was mainly due to the drastically reduced FRN amplitude in the unfair–good intention condition, and this conjecture was verified in the subsequent statistical analysis.

Statistical results for the FRN showed that d-FRN in the unfair condition (FRN amplitude in the unfair–bad intention condition minus that in the unfair–good intention condition) is significantly larger than that in the fair condition, as reflected in the interaction effect analysis. According to the motivational significance model of the FRN, this reflects the more important role of intentionality consideration when an offer is unsatisfactory. From the observed FRN pattern, we speculate that, during social interactions, fairness might be the dominant factor to be considered in outcome evaluation. When the proposed offer is fair, people do not care much about the perceived intentions, as reflected in similar FRN amplitudes in the fair–good intention and the fair–bad intention condition (results of the simple effect analysis in the fair condition). However, when the offer is unfair, people may pay special attention to the perceived intentions of the offer, as reflected in the significantly different FRN amplitudes in the unfair condition when the perceived intention is good or bad.

A recent study demonstrated that facial expression could influence the evaluation of an offer as reflected in the FRN pattern, as a smile would alleviate people's perceived unfairness at the neural level and lead to lower rejection rates behaviorally (Mussel et al., 2014). In a similar vein, in the present study, it seems that perceived good intentions could alleviate people's sense of unfairness as well. During outcome presentation, the magnitudes of elicited FRNs might reflect subjective evaluation of outcome valence (San Martin, 2012). The finding of a less pronounced FRN when unfair offers, as opposed to fair offers, were proposed with good intentions demonstrated that outcomes in the unfair–good intention condition are subjectively evaluated as even more positive than those in the fair–good intention condition. The seemingly counter-intuitive finding led us to conclude that perceived good intentions would alleviate people's sense of unfairness (social pain) to a large extent and that this might outweigh the influence of the objective fairness of an offer.

Why was the subjective evaluation of offers in the unfair–good intention condition so positive? Previous studies have shown that close relationships with one's family, friends and colleagues, as well as a warm smile from a stranger, give rise to the feeling of social warmth (Bargh and Shalev, 2012; Inagaki and Eisenberger, 2013). In the unfair–good intention condition of the present study, proposers were faced with two disadvantageous division schemes for responders, and they could have chosen a more extreme option. However, they treated the responders kindly by selecting the comparatively fairer option. Treating others kindly is significant in social interactions. From this action, good intentions were perceived, possibly leading to the experience of social warmth, as reflected in the significantly reduced FRN amplitude in the unfair–good intention condition.

Beyond illuminating people's social interactions, the current findings also have important implications for understanding the modulation of FRN patterns. Although it is widely accepted that the FRN shows larger deflections for negative feedback than for positive feedback (San Martin, 2012), most previous studies directly compared the differentiated FRN (d-FRN) in different manipulations, which equates to the FRN evoked by positive outcomes subtracted from that induced by negative ones. Such a manipulation makes it impossible to determine whether the FRN pattern is mainly the result of the negative deflections toward negative outcomes or positive deflections toward positive outcomes, or whether both deflections jointly contribute to the observed FRN pattern. To overcome such a disadvantage, recent studies have calculated the FRN separately for positive and negative outcomes, and many of them have found greater deflection of the FRN toward positive outcomes than to negative ones (Foti et al., 2011; Kreussel et al., 2012; Qu et al., 2013; San Martin et al., 2010; Weinberg et al., 2014). In the present study, an outcome proposed out of good intentions would be deemed as more favorable than the same outcome proposed out of bad intentions. Consistent with the existing literature, it was found that the FRN mainly reflects positive deflections toward favorable outcomes rather than the other way round, as reflected in the drastically reduced FRN in the unfair–good intention condition, which further demonstrated the modulatory role of perceived intentions on considerations of fairness.

Consistent with findings of FRN, in a later time window, we also found a prominent positive deflection of P300 toward unfair offers proposed with good intentions. In terms of the P300, previous studies mainly reported that its magnitude is closely linked to the motivational significance of the stimuli (Ma et al., 2014; Zhou et al., 2010) or attentional resources allocated to the outcome (Duncan-Johnson and Donchin, 1977; Johnson and Donchin, 1980). However, recent studies, especially those in risky decision-making, began to report that P300 could also encode the valence of the outcome (Gu et al., 2011; Polezzi et al., 2008; Shen et al., 2013; Wu and Zhou, 2009), which is similar with the valence effect that is well known for the FRN. Therefore, the current finding was consistent with the proposition that P300 may also encode the favorable vs. non-favorable discrepancy in social decision-making (San Martin, 2012). Moreover, compared with the other conditions, the prominently increased P300 in the unfair–good intention condition further indicated the modulatory role of perceived intention on fairness consideration, complementing the behavioral findings and the observed FRN patterns.

5. Conclusion

Providing subjects with fair or unfair division schemes as well as the original allocation set from which the schemes were proposed, we found that the rejection frequency of unfair offers was modulated by the subjects' perception of proposers' intention in the intentional UG. In general, subjects were more likely to be influenced by the perceived intention of offers in unfair scenarios. Electrophysiological data for the FRN and P300 well corroborated the modulatory role of intention in the perception of fairness. It is suggested that such an effect is mainly attributable to the alleviation of unfairness resulting from the perceived kind intention of proposers. In summary, these results highlight the significance of intention considerations in fairness-related social decision-making.

Acknowledgments

Liang Meng was funded by Philosophy and Social Science Planning Project of Zhejiang Province (No. 14NDJC011Z). Qingguo Ma was supported by grant No. 71371167 from the National Natural Science Foundation of China. Qiang Shen was funded by Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning of Beijing Normal University. Qing Xu received a grant from the National Natural Science Foundation of China (No. 71202157).

References

Ames, D.L., Fiske, S.T., 2013. Intentional harms are worse, even when they are not. Psychol. Sci. 24 (9), 1955–1962.
Bargh, J.A., Shalev, I., 2012. The substitutability of physical and social warmth in daily life. Emotion 12 (1), 154–162.
Blount, S., 1995. When social outcomes aren't fair: the effect of causal attributions on preferences. Organ. Behav. Hum. Decis. Process. 63 (2), 131–144.
Boksem, M.A., De Cremer, D., 2010. Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. Soc. Neurosci. 5 (1), 118–128.
Deutsch, M., 1975. Equity, quality and need: what determines which value will be used as the basis of distributive justice? J. Soc. Issues 31 (3), 137–149.
Duncan-Johnson, C.C., Donchin, E., 1977. On quantifying surprise: the variation of event-related potentials with subjective probability. Psychophysiology 14 (5), 456–467.
Eisenberger, N.I., 2012. The neural bases of social pain: evidence for shared representations with physical pain. Psychosom. Med. 74 (2), 126–135.
Falk, A., Fehr, E., Fischbacher, U., 2003. On the nature of fair behavior. Econ. Inq. 41 (1), 20–26.
Falk, A., Fehr, E., Fischbacher, U., 2008. Testing theories of fairness: intentions matter. Games Econ. Behav. 62, 287–303.
Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition and cooperation. Q. J. Econ. 114 (3), 817–868.
Foti, D., Weinberg, A., Dien, J., Hajcak, G., 2011. Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity. Hum. Brain Mapp. 32, 2207–2216.
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, 3891–3897.
Güroglu, B., van den Bos, W., Rombouts, S.A., Crone, E.A., 2010. Unfair? It depends: neural correlates of fairness in social context. Soc. Cogn. Affect. Neurosci. 5, 414–423.
Güroglu, B., van den Bos, W., van Dijk, E., Rombouts, S.A., Crone, E.A., 2011. Dissociable brain networks involved in development of fairness considerations: understanding intentionality behind unfairness. NeuroImage 57, 634–641.
Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. J. Econ. Behav. Organ. 3 (4), 367–388.
Harris, L.T., Fiske, S.T., 2010. Neural regions that underlie reinforcement learning are also active for social expectancy violations. Soc. Neurosci. 5, 76–91.
Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., 2006. Costly punishment across human societies. Science 312, 1767–1770.
Inagaki, T.K., Eisenberger, N.I., 2013. Shared neural mechanisms underlying social warmth and physical warmth. Psychol. Sci. 24 (11), 2272–2280.
Johnson, R., Donchin, E., 1980. P300 and stimulus categorization: two plus one is not so different from one plus one. Psychophysiology 17 (2), 167–178.
Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M.G., Miltner, W.H., 2012. The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. Psychophysiology 49, 207–219.
Ma, Q., Meng, L., Wang, L., Shen, Q., 2014. I endeavor to make it: effort increases valuation of subsequent monetary reward. Behav. Brain Res. 261, 1–7.
Ma, Q., Hu, Y., Jiang, S., Meng, L., 2015a. The undermining effect of facial attractiveness on brain responses to fairness in the Ultimatum Game: an ERP study. Front. Neurosci. 9, 77.
Ma, Q., Meng, L., Shen, Q., 2015b. You have my word: reciprocity expectation modulates feedback-related negativity in the trust game. PLoS ONE 10 (2), e0119129.
Montague, P.R., Lohrenz, T., 2007. To detect and correct: norm violations and their enforcement. Neuron 56, 14–18.
Mussel, P., Hewig, J., Allen, J.J., Coles, M.G., Miltner, W., 2014. Smiling faces, sometimes they don't tell the truth: facial expression in the ultimatum game impacts decision making and event-related potentials. Psychophysiology 51, 358–363.
Nelson Jr., W.R., 2002. Equity or intention: it is the thought that counts. J. Econ. Behav. Organ. 48 (4), 423–430.
Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., Rumiati, R., 2008. Mentalizing in economic decision-making. Behav. Brain Res. 190, 218–223.
Qu, C., Huang, Y., Wang, Y., Huang, Y.X., 2013. The delay effect on outcome evaluation: results from an event-related potential study. Front. Hum. Neurosci. 7, 748.
Radke, S., Güroglu, B., de Bruijn, E.R., 2012. There's something about a fair split: intentionality moderates context-based fairness considerations in social decision-making. PLoS ONE 7, e31491.
San Martin, R., 2012. Event-related potential studies of outcome processing and feedback-guided learning. Front. Hum. Neurosci. 6, 304.
San Martin, R., Manes, F., Hurtado, E., Isla, P., Ibanez, A., 2010. Size and probability of rewards modulate the feedback error-related negativity associated with wins

but not losses in a monetarily rewarded gambling task. NeuroImage 51, 1194–1204.
Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the Ultimatum Game. Science 300, 1755–1758.
Shen, Q., Jin, J., Ma, Q., 2013. The sweet side of inequality: how advantageous status modulates empathic response to others' gains and losses. Behav. Brain Res. 256, 609–617.
Sutter, M., 2007. Outcomes versus intentions: on the nature of fair behavior and its development with age. J. Econ. Psychol. 28, 69–78.
Tabibnia, G., Satpute, A.B., Lieberman, M.D., 2008. The sunny side of fairness: preference for fairness activates reward circuitry. Psychol. Sci. 19 (4), 339–347.
Wang, Y., Roberts, K., Yuan, B., Zhang, W., Shen, D., Simons, R., 2013. Psychophysiological correlates of interpersonal cooperation and aggression. Biol. Psychol. 93 (3), 386–391.
Weinberg, A., Riesel, A., Proudfit, G.H., 2014. Show me the Money: the impact of actual rewards and losses on the feedback negativity. Brain Cogn. 87, 134–139.
Wu, Y., Zhou, X., 2009. The P300 and reward valence, magnitude, and expectancy in outcome evaluation. Brain Res. 1286, 114–122.
Wu, Y., Zhou, Y., van Dijk, E., Leliveld, M.C., Zhou, X., 2011. Social comparison affects brain responses to fairness in asset division: an ERP study with the ultimatum game. Front. Hum. Neurosci. 5, 131.
Zhou, Z., Yu, R., Zhou, X., 2010. To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. Neuropsychologia 48, 3606–3613.