How do we trust strangers? The neural correlates of decision making and outcome evaluation of generalized trust

Yiwen Wang,1,2,† Zhen Zhang,1,2 Yiming Jing,3,†,‡ Emilio A. Valadez,3 and Robert F. Simons3

1School of Humanities and Social Sciences, Institute of Psychological and Cognitive Sciences, Fuzhou University, China, 2Academy of Psychology and Behavior, Tianjin Normal University, China, and 3Psychological and Brain Sciences, University of Delaware, Newark, DE, USA

†These authors contributed equally to this work as co-first authors.
‡Present address: Institute for US-China Issues, University of Oklahoma. Norman, OK, 73019.
Correspondence should be addressed to Robert F. Simons, Department of Psychological and Brain Sciences, University of Delaware, 108 Wolf Hall, Newark, DE 19716, USA. E-mail: rsimons@psych.udel.edu

Abstract

This study investigates the brain correlates of decision making and outcome evaluation of generalized trust (i.e. trust in unfamiliar social agents)—a core component of social capital which facilitates civic cooperation and economic exchange. We measured 18 (9 male) Chinese participants’ event-related potentials while they played the role of the trustor in a one-shot trust game with unspecified social agents (trustees) allegedly selected from a large representative sample. At the decision-making phase, greater N2 amplitudes were found for trustors’ distrusting decisions compared to trusting decisions, which may reflect greater cognitive control exerted to distrust. Source localization identified the precentral gyrus as one possible neuronal generator of this N2 component. At the outcome evaluation phase, principal components analysis revealed that the so called feedback-related negativity was in fact driven by a reward positivity, which was greater in response to gain feedback compared to loss feedback. This reduced reward positivity following loss feedback may indicate that the absence of reward for trusting decisions was unexpected by the trustor. In addition, we found preliminary evidence suggesting that the decision-making processes may differ between high trustors and low trustors.

Key words: trust; event-related potentials; N2; feedback-related negativity; reward positivity

Introduction

Previous research has revealed the critical role of trust in fostering cooperation and solidarity in human society (e.g. Fukuyama, 1995; Hardin, 2002; Putnam, 2000; Uslaner, 2004). In particular, it is generalized trust (i.e. trust in unfamiliar others) which breeds voluntary associations, civic engagement and large-scale economic exchange, and in turn promotes a modern society’s civility, democracy and economic growth (e.g. Fukuyama, 1995; Putnam, 2000; Zak and Knack, 2001; Uslaner, 2004). To date, considerable research from a broad range of disciplines has been conducted to understand generalized trust, especially its social, economic and psychological foundations (e.g. Fukuyama, 1995; Zak and Knack, 2001; Yamagishi, 2011; Jing and Bond, 2015), which bears important practical implications for improving human society’s functioning and the individual’s well-being.

A newly developed trend in trust research, social neuroscience provides us with a unique tool to discover the cognitive and biological mechanisms that facilitate human trust. However, much of the extant brain research merely investigated the trust building processes in which the trustor can learn the trustee’s
trustworthiness through repeated socio-economic transactions (history-based or knowledge-based trust; e.g. McCabe et al., 2001; King-Casas et al., 2005; Tomlin et al., 2006; Krueger et al., 2007; Fouragnan et al., 2013; Fett et al., 2014). On the other hand, although a few brain studies have looked into a single round of trust transaction between two strangers (e.g. Chang et al., 2011; van den Bos et al., 2009), these studies were largely focused on how the trustee reciprocates the trustor’s trust rather than how the trustor decides to trust unknown social agents. Furthermore, there is a paucity of brain studies exploring the temporal process of trust decision making. The present study is an attempt to address these limitations using event-related potentials (ERPs) to examine the neural correlates of trustors’ decision making and outcome evaluation in a one-shot trust game (a trust game in which the trustor’s decisions are in response to a different trustee in each round; Berg et al., 1995). The results of this study will provide new information regarding individuals’ decision-making processes that facilitate civic cooperation and economic exchange.

Why do people trust strangers?

Trusting others can be a costly endeavor because it renders the trustor vulnerable due to his or her dependence on the trustee’s actions (Mayer et al., 1995). In particular, trust may indicate the trustor’s positive expectation on the trustee’s goodwill; however, such an expectation may not be confirmed, and the trustor may suffer unpleasant consequences accordingly (e.g. being cheated). This risk inherent to trusting is especially salient when the trustor is unfamiliar with the trustee’s morality and reputation (Yamagishi, 2011). As a result, from an economically rational point of view, people should not trust strangers who may behave opportunistically to exploit the trust (Berg et al., 1995; Dunning et al., 2014).

Nonetheless, empirical studies have revealed a great tendency of people to trust each other, even strangers (for reviews, see Johnson and Mislin, 2011; Balliet and Van Lange, 2013). Researchers have hypothesized a number of possible mechanisms that may account for the existence of generalized trust in human society. First, evolutionary biologists and psychologists believe that cooperation and trust reflect adaptive instincts evolved for human social life (Richerson and Boyd, 2005, ch. 6). Specifically, even if social encounters will not repeat between two strangers, a good reputation of being trusting and trustworthy will enhance the individual’s chance of being selected and rewarded by other partners (indirect reciprocity; Nowak, 2006). In contrast, the violation of cooperation and trust is universally punished in human communities which, in turn, greatly reduces the fitness of violators (Fehr and Fischbacher, 2003).

Second, modern society has developed social and economic institutions which further promote cooperation between unrelated people (e.g. Fukuyama, 1995; Putnam, 2000; Yamagishi, 2011; Camera et al., 2013). For instance, modern regulatory institutions (e.g. the universalistic legal system) establish and enforce the norm of civic cooperation, which boosts citizens’ confidence in trusting people with different characteristics such as demographic background and group memberships (Knack and Keefer, 1997; Jing and Bond, 2015). Importantly, there has been a growing body of research suggesting that trusting strangers is a moral standard in modern society (e.g. Dunning et al., 2014). In particular, people trust strangers not only because such cooperative relationships may bring them better economic payoffs but also because they desire to experience happiness and personal satisfaction from being cooperative and trusting (Yamagishi et al., 2015).

Third, people can also develop generalized trust based on their early experiences (Rotter, 1980). Being raised in a stable and secure environment, for instance, a child will first learn to trust people around him or her (e.g. the caregiver) and will later generalize such trust to social agents in other environments (Uslaner, 2002; Stolle and Hooghe, 2004). From this perspective, there may exist certain stable individual differences in generalized trust which can be attributed to the socialization environment. Some research has suggested that childhood socialization context fostering self-determination and civic concern, rather than fostering obedience and practicality, may facilitate a person’s development of generalized trust (Jing and Bond, 2015).

The neural correlates of trust-game responses

If cooperation and trust are adaptive tendencies in modern society, what are the cognitive and biological mechanisms that facilitate decisions to trust unfamiliar social agents? In recent years, a number of studies have been conducted to understand the genetic, endocrinological and brain mechanisms of trust (Riedl and Javor, 2012) that may provide some insight into this question. Most of these studies utilized Berg et al.’s (1995) trust game as the behavioral task (for a review, see Tzieropoulos, 2013). The trust game operationalizes the behavioral definition of trust, assessing trustors’ willingness to engage in economic exchange in a socially interdependent situation, and assume its attendant risks. Past research has provided evidence supporting this game’s external validity by linking trusting behaviors within and outside the experiment (e.g. Glaser et al., 2000). In this section, we will review findings from functional magnetic resonance imaging (fMRI) and ERP studies to highlight the brain correlates of trust-game responses.

fMRI studies

There exist a few fMRI studies which have looked into one-shot trust transactions between unfamiliar participants (e.g. van den Bos et al., 2009; Chang et al., 2011). However, these studies were largely focused on the brain correlates of trustees’ reciprocating decisions, and revealed little about trustors’ decision making of generalized trust.

On the other hand, more fMRI studies have examined trustors’ brain activation during the repeated trust game, an adaptation of Berg et al.’s (1995) original (one-shot) trust game (e.g. Delgado et al., 2005; King-Casas et al., 2005; Krueger et al., 2007; Fett et al., 2014). In the adaptation, each trustor played with the same trustee over multiple trials, as opposed to with a different trustee in each trial. As a result, what these studies investigated was the trust building processes in which trustors can track trustees’ reputations and make trusting/distrusting decisions (i.e. history-based trust).

Specifically, it was shown that mentalizing (theory of mind [ToM]), reward and affective processing, and cognitive control were important for the trustor to understand the trustee’s intentions and to evaluate his or her reputation over trials, which in turn facilitated the trust vs. distrust decision making. Accordingly, brain regions associated with mentalizing (e.g. the tempoparietal junction, anterior paracingulate cortex and posterior cingulate/prefrontal cortex; see van Overwalle and Baetens, 2009, for a review), reward learning (e.g. the head of the caudate nucleus and orbitofrontal cortex) and affective processing (e.g.
insula, amygdala), and conflict monitoring and cognitive control [e.g. anterior cingulate cortex (ACC)] were activated during the repeated trust game (e.g. Delgado et al., 2005; King-Casas et al., 2005; Krueger et al., 2007; Kosick and Tranel, 2011; Fett et al., 2014; for a review, see Tzineropoulos, 2013). It is possible that some of these brain mechanisms are also involved in generalized trust. For instance, intentionality inference and outcome evaluation seem to be necessary for the trustor to engage in the one-shot trust transaction with a stranger.

**ERP studies**

In contrast to fMRI research, few studies have utilized ERP methodology and attempted to understand trust-related brain activity. Furthermore, the extant ERP studies only focused on trustors’ brain activity at the outcome evaluation stage (i.e. when the trustor was shown whether his or her trust was reciprocated or exploited), but ignored the decision-making stage (i.e. when the trustor was making the decision to trust/distrust).

With respect to outcome evaluation, ERP researchers examined trustors’ reward expectation (Chen et al., 2012; Long et al., 2012), and focused on the feedback-related negativity (FRN) that peaks at frontocentral recording sites between 200 and 350 ms after feedback onset (Holroyd and Coles, 2002; Miltner et al., 1997). The FRN was selected based on ample evidence suggesting that the FRN is sensitive to reward prediction errors and can distinguish between positive and negative outcomes (for a review, see Sambrook and Goslin, 2015). Specifically, the FRN is thought to be relatively more negative if the outcome was worse than expected. Previous research has source-localized FRN to brain regions involved in reward processing, such as striatum and dorsal ACC (e.g. Carlson et al., 2011; Hauser et al., 2014). As reviewed earlier, the activation of the brain’s reward network has been frequently reported in fMRI studies examining trust.

In Chen et al.’s (2012) experiment using the trust game, they concluded that trustors experienced a greater violation of reward expectation with betrayals of the attractive (trustworthy) trustees than with betrayals of the unattractive (untrustworthy) trustees, as indicated by a more negative-going FRN in response to loss feedback present in the former condition than in the latter condition.1 Additionally, Long et al. (2012) found that trustors’ FRNs were more negative going for no-gain outcomes than for gain outcomes, by which they suggested that reward expectation drove trusting decision making.2 However, these studies have some experimental design limitations, which make it unclear to what extent their behavioral findings are tied to generalized trust or even trust at all (Footnotes 1 and 2). In addition, an unresolved ERP issue in these studies lies with how the FRN should be interpreted in relation to the valence of feedback. Conventionally, the FRN is interpreted as a negativity driven by losses, but some recent work utilizing principal components analysis (PCA) has challenged this view, suggesting that the FRN may actually be an artifact of a positivity enhanced by rewards (e.g. Foti et al., 2011; for a review, see Proudfit, 2015). Thus, a more negative-going FRN following loss feedback may in fact reflect a reduced positivity in the absence of expected reward (e.g. when trust is not reciprocated).

**The current study**

The present study has two major purposes: first, we explore trustors’ ERP components at the decision-making stage (i.e. decide to trust or distrust a stranger) of his or her one-shot trust game play. Given a dearth of existing studies, no hypothesis is made regarding this process. Second, to address the ambiguity in whether the outcome evaluation ERP, viz., FRN, in the trust game is best characterized as positivity to reward or as negativity to non-reward, we subject trustors’ ERP responses following gain and loss feedback to PCA, which can reveal the underlying ERP components in the time window of FRN.

**Materials and methods**

**Participants**

Participants were 18 Chinese undergraduate and graduate students (nine male and nine female) from a university in northern China. Their ages ranged from 18 to 26 years (M = 21.20, s.d. = 2.60). All participants had normal or corrected-to-normal vision, and none had a history of any neurological or psychiatric disorders. Each participant gave written consent prior to participation.

**Behavioral task**

The behavioral task is a modified version of Berg et al.’s (1995) trust game, which is aimed at assessing the participant’s generalized trust. The participant played the role of the trustee with the alleged trustee. At the beginning of each round, both the trustee and the trustee are given 10 game points as an initial endowment. The trustee needs to decide whether to send all his or her 10 points to the trustee or to keep this endowment. If the trustee chooses to keep it, this round ends and both players will receive 10 points. If the trustee chooses to send the initial endowment, the points sent will be tripled to 30 points, and the trustee will then decide how to allocate these tripled points plus his or her own initial endowment (40 points in total). The trustee also has two options: to divide the 40 points equally and send back 20 points to the trustee, or to keep the 40 points and send nothing back to the trustee. Given the possibility of being exploited by the trustee, the trustee’s decision to send money reflects his or her willingness to be vulnerable to the trustee’s allocation decision, which is the behavioral operationalization of trust.

The current trust game is a one-shot version wherein the trustor plays with different trustees with complete anonymity. At the beginning of the task, participants were informed that in each round the trustee was a different adult randomly selected from a large and representative subject pool (N = 400), and that the experimenter sampled and interviewed these adults before this experiment and asked them to imagine participating in a single round of this trust game. Participants were told that these adults indicated their choices between sending 20 points back and keeping all 40 points, if being entrusted 30 points by a stranger. Participants were also told that the experimenter recorded all these adults’ choices and the computer would randomly select one to respond to the participant’s investment in each round. In reality, however, all trustees’ responses were set up by a pre-programmed procedure (same across all participants), such that the decision to reciprocate was made randomly across trials, and the overall reinforcement rates for the trustor (i.e. the rates of receiving 20 points if the trustor makes the trusting choice) were approximately 50%.

**Stimuli and procedure**

Participants completed 150 rounds of the trust game while their brain potentials were recorded using electroencephalography (EEG). In each round, as illustrated in Figure 1, the participant...
first sees a picture of a simplified trust game decision tree showing possible outcomes for his or her single decision for 1500 ms. After a variable 800–1000 ms fixation cross, a picture indicating decision options is displayed in the center of the screen for 2000 ms. During this time, the participant chooses either to keep (cued by the number ‘10’) or to send (cued by the number ‘30’) his or her initial endowment by using his or her index finger to press either the ‘1’ or ‘3’ key on the keyboard, respectively. The position of decision options is counterbalanced between subjects. If the participant fails to respond within 2000 ms, a warning message is displayed and indicates to the participant that he or she responded too slowly. Following a variable 800–1200 ms interval with a black screen, the outcome of the participant’s current trial as well as his or her current total score are displayed for 1200 and 2000 ms, respectively.

Upon a participant’s arrival, the experimenter described the rules of the trust game in detail. In particular, to make participants treat game points seriously, they were told that their participation compensation would be tied to the total number of game points earned in the game. However, regardless of their choices, participants earned similar number of points over trials, and we actually paid them a flat rate for their participation (about 8 dollars each person). Then, participants were fitted with an electrode cap and seated comfortably 1 m from a computer screen in an electromagnetically shielded room. A practice block of 10 trials was administered before the formal task.

**EEG recording and analysis**

We measured brain electric activity from 64 channels with the averaged bilateral mastoid reference and a forehead-ground using a modified 10–20 system electrode cap (Neuroscan Inc.). Vertical electrooculography (EOG) activity was recorded with electrodes placed above and below the left eye, and horizontal EOG activity was recorded with electrodes placed on the outer-boards of both eyes. All electrode sites were cleaned with alcohol and inter-electrode impedance was maintained below 5 kΩ. Bi-signal were extracted with a 0.05–100 Hz band-pass filter and continuously sampled at 1000 Hz/channel for off-line analysis. Eye blink artifacts were removed automatically using Scan software (Neuroscan Inc.). All trials in which EEG voltages exceeded a threshold of ±75 μV during recording were excluded from further analysis. After artifact rejection, the average numbers of trials being analyzed at the decision-making phase were 96 for the participant’s trusting decisions, and 50 for distrusting decisions. Because the outcome for distrusting decisions was always the same (i.e. 10 points), only trials with trusting decisions were analyzed at the outcome evaluation phase. The average numbers of trials being analyzed at the outcome evaluation phase were 46 for loss feedback, and 49 for gain feedback.

ERPs were analyzed separately for the decision-making and outcome evaluation phases. For the decision-making phase (i.e. when the participant makes a choice between keeping and sending the initial endowment), epochs were extracted from pre-stimulus to 600 ms post-stimulus. ERPs were then constructed by separately averaging the trials for the two decision conditions (i.e. trust vs distrust). For each ERP, activity in the –200 ms to 0 ms time window prior to the decision stimulus served as baseline. Based on the results of previous studies, we used 15 electrodes (F3/Fz/F4, FC3/FCz/FC4, C3/Cz/C4, CP3/CPz/CP4, P3/Pz/P4) to yield two factors: anteriority (frontal, frontocentral, central, centroparietal and parietal) and laterality (left, midline and right).

For the outcome evaluation phase (i.e. when the participant sees the outcome feedback), epochs were extracted from 200 ms before to 1000 ms after each feedback presentation and adjusted for baseline activity. The FRN was quantified using temporospatial PCA following the two-step procedure recommended by Dien (2010a). Using the ERP PCA Toolkit (version 2.47; Dien, 2010b), PCA was applied to participants’ trial-averaged, post-stimulus ERPs (the pre-stimulus periods were not included in order to eliminate a source of feedback-irrelevant variance) following gain and loss feedback. In the first step, a temporal PCA with promax rotation resulted in 12 factors based on the resulting Scree plot (Cattell, 1966). In the second step, each of the 12 temporal factors was analyzed using a spatial PCA with infomax rotation, resulting in two spatial factors (based on the averaged Scree plot for all 12 temporal factors), for a total of 24 unique factors.
temporal/spatial factor combinations. Factor combinations were considered if they shared the time course of the FRN (i.e. peaks at approximately 200–300 ms), shared the topography of the FRN (i.e. peaks at frontocentral electrode sites), and accounted for more than 1% of the variance in the averaged ERP. Three factor combinations met all these criteria, and were compared between the two feedback conditions (gain vs loss).

Results

Behavioral results

Cross-sectional analysis. The overall percentages of trials in which participants chose to send (trust) or to keep (distrust) were 66.1% and 33.9% (s.d. = 14.8%), respectively. Past fMRI studies using the repeated trust game have found that the trust rates range from about 50% to 84% (M = 62.20%) when interacting with neutral counterparts (Delgado et al., 2005; Krueger et al., 2007; Phan et al., 2010; Farioli et al., 2012; Fouragnan et al., 2013; Wardle et al., 2013). The trust rates in our experiment fall within this range. Importantly, a one-sample t-test indicated that the average trust rates are significantly higher than the manipulated 50% reinforcement rates, t(17) = 4.60, P < 0.001. This is consistent with previous research (e.g. Dunning et al., 2014), suggesting an excessive amount of generalized trust (compared to the actual amount of reward) that people exhibited in the trust game.

Table 1 presents each individual’s trusting tendency in the game. A paired-samples t-test indicated that across all trials participants tended to trust the stranger more than to distrust, t(17) = 4.60, P < 0.001. The mean response times for choosing to trust or distrust were 478.89 ms (s.d. = 150.39 ms) and 472.49 ms (s.d. = 123.56 ms), respectively. There was no significant difference in response time between trusting and distrusting choices across trials, t(17) = 0.45, P > 0.05.

Longitudinal and individual differences analysis. Research based on the repeated trust game could reveal participants’ learning curves over time, such that trustors’ levels of trust may increase or decrease as a result of tracking the same trustee’s reputation (e.g. Fett et al., 2014). To further validate our manipulation on generalized trust, we did longitudinal analysis to see whether such time-related or trial-related effects existed in our experiment. Specifically, we conducted a two-level hierarchical linear modeling (Raudenbush and Bryk, 2002) analysis in which trial-by-trial decision data were nested within each participant. This analysis was performed using HLM7 software (Raudenbush et al., 2011). Although the sample size (N = 18) for our level-two analysis (i.e. analysis of individual differences) was small, a simulation study (Maas and Hox, 2005) has suggested that HLM’s estimation of regression coefficients may be robust to small sample sizes.

A hierarchical generalized linear modeling (HGLM) analysis of binary outcome did not reveal any cubic (γ = −0.00, t = −0.22, P > 0.05, odds ratio = 1.00), quadratic (γ = 0.00, t = 0.07, P > 0.05, odds ratio = 1.00) or linear (γ = 0.00, t = 0.01, P > 0.05, odds ratio = 1.00) growth curve for trusting/distrusting choices over time. This indicates that, unlike in the repeated trust game (e.g. Fett et al., 2014), no time-related learning curves were observed in our one-shot trust game; as expected, participants seemed to consider trustees as different persons.

On the other hand, the HGLM analysis of trial-by-trial data found a significant cross-level interaction suggesting that participants’ trust rates in the game moderate the linkage between prior outcome feedback and later trusting choices, γ = 0.24, t = 2.34, P < 0.05, odds ratio = 1.27. As illustrated in Figure 2, the simple regression line (Cohen et al., 2003, pp. 268–270) indicated that for high trustors (1 s.d. above the mean trust rates, i.e. 81%) outcome feedback received in the prior trial did not influence the decisions made in the next trial, γ = 0.02, t = 0.96, P > 0.05, odds ratio = 1.02. In contrast, for low trustors (1 s.d. below the mean trust rates, i.e. 51%), trusting choices were more likely to be made following loss feedback (expected probability = 0.63) whereas distrusting choices were more likely to be made following gain feedback (expected probability = 0.61), γ = −0.05, t = −2.69, P < 0.05, odds ratio = 0.95.

An HLM analysis of continuous outcome also found a significant cross-level interaction indicating that participants’ trust rates in the game moderated the linkage between trial-by-trial trusting/distrusting choices and reaction time, γ = 242.79, t = 3.09, P < 0.01. Further plotting of the simple regression line (Figure 3) indicated that for high trustors (81% trust) it took more time to make trusting choices than to make distrusting choices across trials, γ = 38.87 (i.e. mean time difference = 38.87 ms), t = 2.27, P < 0.05. In contrast, for low trustors (51% trust) it took more time to make distrusting choices than to make trusting choices across trials, γ = −33.95 (i.e. mean time difference = 33.95 ms), t = −2.23, P < 0.05.

ERP results of decision-making phase

By visually inspecting the pattern of difference waves between trusting and distrusting decisions (Figure 4A), we observed a negative-going wave between 200 and 400 ms, which may be identified as an N2 component typically peaking between 200 and 350 ms after stimulus onset. In this time window, no other components can be distinguished from this N2 component in difference waves. As a result, we focused our analysis on the N2 component at the decision-making phase.

A 2 (decision choice: trust vs distrust) × 5 (anterior: frontal, frontocentral, central, centro-parietal, parietal) × 3 (laterality: left, midline, right) repeated-measures ANOVA on N2 amplitudes revealed a significant main effect of decision choice, F(1, 17) = 5.01, P < 0.05, partial η² = 0.23, such that a more
negative N2 was observed for the distrusting decision (2.45 ± 0.99 μV) than for the trusting decision (3.75 ± 0.69 μV) (Figure 4A). The main effect of anteriority was also significant, F(4, 68) = 8.56, P < 0.01, partial $\eta^2 = 0.34$. Post hoc analyses revealed a linear trend with more negative amplitudes at more anterior sites. The main effect of laterality was significant, F(2, 34) = 6.82, P < 0.01, partial $\eta^2 = 0.29$. Post hoc analyses reveal a more positive N2 at right sites (3.99 ± 0.78 μV) than midline (2.66 ± 0.91 μV) and left sites (2.66 ± 0.91 μV), P’s < 0.05. The P-values of main and interaction effects were corrected using the Greenhouse–Geisser method for repeated-measures effects. The Bonferroni correction was used for multiple comparisons.

Importantly, an additional analysis of covariance (ANCOVA) controlling for trust rates suggested that these main effects were not qualified by participants’ exhibited trusting levels in the game; the significant main effects of trust/distrust choices, F(1, 16) = 4.86, P < 0.05, partial $\eta^2 = 0.23$, of anteriority, F(4, 64) = 8.38, P < 0.01, partial $\eta^2 = 0.34$, and of laterality, F(2, 32) = 7.14, P < 0.01, partial $\eta^2 = 0.31$, were also observed in the ANCOVA.

In order to explore the possible neuronal generators of this N2 component during 275–325 ms, source localization was performed using standardized low resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui et al., 1994). Our results (Figure 4B and C) suggested that (i) the highest level of activation of the N2 component was localized to the right superior parietal lobule for the trusting choice (Brodmann area 7, Montreal Neurological Institute [MNI] coordinates = [25, −55, 70]) and for the distrusting choice (Brodmann area 7, MNI coordinates = [25, −60, 65]), and (ii) significantly more positive current difference was detected between the distrusting choice and the trusting choice in three voxels (peak t-score = 1.57, P < 0.01), clearly clustered in left precentral gyrus regions (two voxels within the left BA 6, MNI coordinates = [−60, 0, 20], [−60, 0, 15]; one voxel within the left BA 4, MNI coordinates = [−60, −5, 20]).

**ERP/PCA results of outcome evaluation phase**

The grand average ERPs (i.e. before PCA) from the outcome evaluation phase are presented in Figure 5A. For the three temporal/spatial PCA factor combinations that fit the signature of the FRN, peak voltages were compared across feedback conditions (gain and loss) statistically using a robust ANOVA conducted with the ERP PCA Toolkit. This ANOVA is more robust against violations of statistical assumptions than the traditional ANOVA and provides improved Type I error control, due in part to bootstrapping procedures (Keselman et al., 2003). Only one factor combination significantly differentiated gain from loss feedback based on the results of the robust ANOVA, F(1, 17) = 48.28, P < 0.001, partial $\eta^2 = 0.74$, compared to F(1, 17) = 2.42, P = 0.14, partial $\eta^2 = 0.13$ and F(1, 17) = 0.93, P = 0.34, partial $\eta^2 = 0.05$, such that peak scores on the significant factor were greater for gain (m = 24.31) than for loss (m = 18.09) feedback. An additional ANCOVA controlling for trust rates also confirmed that this factor’s main effect was not qualified by participants’ trusting levels, F(1, 16) = 46.03, P < 0.001, partial $\eta^2 = 0.74$. As illustrated in Figure 6A, this factor peaked at approximately 280 ms and is positive in both gain and loss conditions. Consequently, we hereafter refer to this factor as the ‘reward positivity’ (Foti et al., 2011; Proudfoot, 2015).

**Correlations between trust rates, N2 and reward positivity peak scores**

To explore the potential individual differences in N2 amplitudes and reward positivity peak scores, we correlated each participant’s trust rates in the game with his or her N2 and reward positivity peak scores. As illustrated in Table 2, there existed moderate, negative correlations between trustors’ trust rates and their reward positivity peak voltages. Such associations were slightly stronger for the response to losses (r = −0.47, P = 0.05) than those to gains (r = −0.41, P = 0.09). These results suggest that more trusting people may exhibit a reduced reward positivity in response to both unexpected loss and expected gain. On the other hand, the correlations between trust rates and N2 amplitudes were weak or trivial (P’s > 0.10), which in
Fig. 4. Event-related potential at the decision-making phase. (A) Grand average waveforms. (B) Scalp topography for the N2 in the 250–350 ms post-stimulus time window. (C) sLORETA images of the standardized current density maximum for Trust/Distrust and source localization (Distrust minus Trust during 275–325 ms). The results at the peak latency of the N2 (325 ms) are illustrated. A, anterior; P, posterior; R, right; L, left; LH, left hemisphere; RH, right hemisphere; LV, left view; RV, right view; TV, top view; BV, bottom view.

Fig. 5. Event-related potential at the outcome evaluation phase. (A) Grand average waveforms. (B) Scalp topography for the FRN in the 200–300 ms post-stimulus time window.
turn provides insufficient evidence about N2’s differences between high vs low trustors.

Discussion

Large-scale cooperation and generalized trust are adaptive to people’s social life in modern society. But what are the cognitive and biological mechanisms that facilitate citizens’ decisions to trust strangers? To answer this question, we assessed ERP correlates of trustors’ decision making and outcome evaluation of generalized trust. At the decision-making phase, our results revealed differential waves of the N2 component (peaking at 250–350 ms after the onset) accompanying the decision to trust or to distrust a stranger, such that a more negative-going N2 was observed for choosing to distrust than for choosing to trust. At the outcome evaluation phase, PCA revealed a reward positivity (see also Foti et al., 2011) that was reduced in response to loss feedback compared to in response to gain feedback.

As one of a few brain studies examining generalized trust, we provide some novel and intriguing findings. In particular, we revealed the ERP correlates of trustors’ decision making when interacting with a large number of unfamiliar trustees. More specifically, a frontocentrally distributed N2 component, which peaked at 250–350 ms after the stimulus onset, differentiated trustors’ trusting choices from their distrusting choices. Given that a greater N2 is often associated with increased attention or cognitive control in cognitively demanding tasks (e.g. the go/no go task and the flanker task; for a review, see Folstein and Van Petten, 2008), a more negative-going N2 for a distrusting choice than a trusting choice in our study may suggest greater cognitive effort to inhibit trusting or resolve conflict elicited by making distrusting decisions. From this perspective, trusting and cooperating with strangers may be a default strategy for the majority of participants in our experiment. Importantly, the average trusting rates in our trust game were significantly greater than the manipulated reward rates. The trial-by-trial analysis also revealed that high trustors’ choices to trust in the game seemed not to be affected by the outcome feedback (gain vs loss) that they received in the prior trial. These results are consistent with some recent research findings suggesting that generalized trust is not merely a strategy for maximizing one’s self-interest, but may also be motivated by people’s pro-social desire for being cooperative and trusting (Dunning et al., 2014; Yamagishi et al., 2015). Especially for people who possess high levels of generalized trust, their internalized moral codes for pro-sociality may override instrumental concerns in making decisions to trust.

In addition, a more positive current difference between the distrusting choice and the trusting choice was localized to pre-central gyrus (BA6). This is consistent with the extant fMRI research, revealing that (i) precentral gyrus is a neuro-generator of N2 (e.g. Kiehl et al., 2001) and (ii) precentral gyrus is involved in cognitive control (e.g. Booth et al., 2003; Li et al., 2006. For a review, see Riddervold et al. 2004). On the other hand, source localization identified that both trusting and distrusting choices the highest activation of the N2 was generated from the superior parietal lobule where the precuneus (part of mentalizing system) is located (van Overwalle and Baetens, 2009). Previous ERP studies have revealed a similar ERP component, viz., N270, which is enhanced during ToM reasoning (e.g. Sabbagh et al., 2004). The source localization result is also consistent with previous fMRI findings, suggesting that the ToM system, including precuneus (BA7), is activated in trust decision making (e.g. Delgado et al., 2005; Fett et al., 2014). It appears that both trusting and distrusting decisions rely on the trustor’s

![Fig. 6. Reward positivity factor from principal components analysis at the outcome evaluation phase. (A) Grand average waveforms. (B) Scalp topography at peak latency (280 ms post-stimulus).](image)

| Table 2. Correlations between trust rates in the trust game and ERP data |
|---------------------------------|----------------|-----------------|-----------------|-----------------|
|                                  | N2 (distrust) | N2 (trust)      | Reward positivity (loss) | Reward positivity (gain) |
| Trust rates                      | –0.08         | –0.25           | –0.47*           | –0.41†          |
| N2 (distrust)                    | 0.84***       | 0.09            | 0.22             | 0.33            |
| N2 (trust)                       |               |                 |                 |                 |
| Reward positivity (loss)         |               |                 | 0.90***          |                 |

Notes: Trust rates, the percentage of trials in which the participant chose to trust in the game; N2 (distrust), N2 amplitudes associated with distrusting choices; N2 (trust), N2 amplitudes associated with trusting choices; Reward positivity (loss), PCA peak voltage on reward positivity underlying FRN following loss feedback; Reward positivity (gain), PCA peak voltage on reward positivity underlying FRN following gain feedback.

†P < 0.10, *P < 0.05, **P < 0.001.
inferences about the trustee’s intentions. But the decisions to distrust as opposed to trust strangers may especially require the trustor’s cognitive control suppressing the injunctive norm for being cooperative (Dunning et al., 2014).

With regard to outcome evaluation of trust game decisions, the results of our PCA are inconsistent with the conventional view that the FRN is a negativity driven by unexpected losses. Instead, in line with some recent work (e.g. Foti et al., 2011), we found that, in the context of this experiment, the feedback ERP was in fact driven by a reward positivity which was reduced in the absence of expected reward (i.e. unexpected loss). As a result, we observed a reduced reward positivity response when trustors’ reward expectation was violated. In addition, our correlational analysis revealed two trend-level effects suggesting that the reward positivity following gain and loss feedback may have been reduced for participants who exhibited greater trust in the game compared to those who were less trusting. Following gain feedback, the reward positivity was larger among low trustors (i.e. unexpected reward) compared to high trustors (i.e. expected reward). Following loss feedback, the reward positivity was reduced among high trustors (i.e. unexpected loss) compared to low trustors (i.e. expected loss). This is consistent with a large body of evidence suggesting that the FRN is sensitive to reward prediction errors and can distinguish between positive and negative outcomes (for a review, see Sambrook and Goslin, 2015). High trusting people likely have greater expectations for trust reciprocation and are more surprised by other social agents’ violation of pro-social norms than low trusting people and, conversely, are less surprised by other social agents’ upholding of pro-social norms than low trusting people.

Our trial-by-trial analysis revealed some intriguing behavioral results, also suggesting potential individual differences in decision making of generalized trust. First, high trustors exhibited a relatively stable trusting pattern over trials, regardless of whether they were reciprocated in the prior trial. This not only supports the argument that the reasons to trust may go beyond economic rationality (Dunning et al., 2014; Yamagishi et al., 2015) but also corroborates previous fMRI research revealing that, even in the repeated trust game, people’s prior belief in the partner’s benevolence can reduce their reliance on feedback mechanisms to make trusting decisions (Delgado et al., 2005). On the other hand, our analysis indicated that low trustors adopted a more varying game strategy, such that they were more inclined to make investment (to trust) following prior losses but more inclined to keep endowment (to distrust) following prior gains. Unlike high trustors, low trustors’ sensitivity to previous payoffs may suggest that their trust game decisions were based more on economic calculations than social concerns. Further investigation is required to determine why low trustors become more risk seeking after losses but more risk averse after gains.

Second, we found that high trustors tended to take more time making trusting choices, probably due to our actual, relatively low reinforcement rates which are contradictory to high trustors’ expectation of benign humanity (e.g. most people are trustworthy). On the contrary, low trustors tended to take more time making distrust choices, probably due to some conflicts between reward seeking (trust) and risk aversion (distrust). These findings may also provide some new insights into how people’s variations on social preference, such as pro-social preference (e.g. Yamagishi et al., 2015) and betrayal aversion (e.g. Bohnet and Zeckhauser, 2004), could moderate their decision-making processes related to generalized trust. However, given our limited sample size, these behavioral findings are not conclusive, and replications are needed in larger samples.

Additionally, due to the limited number of valid trials for trusting vs distrustng choices, we were not able to conduct trial-by-trial analysis to explore the individual differences in ERP correlates. Future research should utilize a more sensitive experimental design allowing single-trial EEG analyses of trust.

**Limitations**

In addition to small sample size and lack of single-trial EEG analysis, our study has other limitations that need to be acknowledged. First, our results are based on a Chinese college student sample, thus their generalizability to other populations is in question. In particular, according to the psychological and behavioral literature, there exist some age differences and cultural differences (e.g. West vs East Asia) in people’s levels of generalized trust (e.g. Delhey et al., 2011; Li and Fung, 2012). The type of trustees (e.g. strangers vs known people) may also modulate trust-related responses. Future studies should examine how these potential, contextual factors influence ERP correlates of trust, and may qualify our current findings. Second, it is an empirical question to what extent different pay-off structures (e.g. invest tripled money vs invest doubled money) and responding scales (e.g. binary choices of investment vs continuous amount of investment) may affect participants’ generalized trust decisions and ERP responses. Future studies should look into this issue. Third, our source localization analysis (sLORETA) is a reverse inference technique. The specific role of precentral gyrus and superior parietal lobule in generalized trust decision making demands further investigations based on fMRI research. Finally, our study did not assess trustors’ emotional reactions at the outcome evaluation phase. Given that emotional processing is an integral part of trust decision making (Chang et al., 2011; Koscit and Tanel, 2011; Aimone et al., 2014), especially for distrust (Dimoka, 2010), future studies should consider trustors’ emotions as covariates in the ERP analysis of generalized trust.

**Funding**

This work was supported by the National Natural Science Foundation of China [31371045]; the Program for New Century Excellent Talents in Universities [NCET-11-1065], and the MOE Project of Key Research Institute of Humanities and Social Sciences at Universities [12JJD190004] to Y.W.

**Conflict of interest.** None declared.

**References**

Aimone, J.A., Houser, D., Weber, B. (2014). Neural signatures of betrayal aversion: an fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20132127.

Balliet, D., Van Lange, P.A. (2013). Trust, conflict, and cooperation: a meta-analysis. *Psychological Bulletin*, 139, 1090–112.

Berg, J., Dickhaut, J., McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122–42.

Bohnet, I., Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55, 467–84.

Booth, J.R., Burman, D.D., Meyer, J.R., et al. (2003). Neural development of selective attention and response inhibition. *Neuroimage*, 20, 737–51.

Boudreau, C., McCubbins, M.D., Coulson, S. (2009). Knowing when to trust others: an ERP study of decision making after receiving information from unknown people. *Social Cognitive and Affective Neuroscience*, 4, 23–34. doi: 10.1093/scan/nss034.

Dimoka, J.A., Houser, D., Weber, B. (2014). Neural signatures of betrayal aversion: an fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20132127.

Balliet, D., Van Lange, P.A. (2013). Trust, conflict, and cooperation: a meta-analysis. *Psychological Bulletin*, 139, 1090–112.

Berg, J., Dickhaut, J., McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122–42.

Bohnet, I., Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55, 467–84.

Booth, J.R., Burman, D.D., Meyer, J.R., et al. (2003). Neural development of selective attention and response inhibition. *Neuroimage*, 20, 737–51.

Boudreau, C., McCubbins, M.D., Coulson, S. (2009). Knowing when to trust others: an ERP study of decision making after receiving information from unknown people. *Social Cognitive and Affective Neuroscience*, 4, 23–34. doi: 10.1093/scan/nss034.
Dunning, D., Anderson, J.E., Schlosser, T., Ehlebracht, D., Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *Neuroimage*, 57, 1608–16. doi:10.1016/j.neuroimage.2011.05.037.

Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245–76.

Chang, L.J., Smith, A., Dufwenberg, M., Sanfey, A.G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, 70, 560–72.

Chen, J., Zhong, J., Zhang, Y., et al. (2012). Electrophysiological correlates of processing facial attractiveness and its influence on cooperative behavior. *Neuroscience Letters*, 517, 65–70.

Cohen, J., Cohen, P., West, S.G., Aiken, L.S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Delgado, M.R., Frank, R.H., Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, 8, 1611–8.

Delhey, J., Newton, K., Welzel, C. (2011). How general is trust in “most people”? Solving the radius of trust problem. *American Sociological Review*, 76, 786–807.

Dien, J. (2010a). Evaluating two-step PCA of ERP data with geomin, oblimin, promax, and varimax rotations. *Psychophysiology*, 47, 170–83.

Dien, J. (2010b). The ERP PCA Toolkit: an open source program for advanced statistical analysis of event-related potential data. *Clinical Neurophysiology*, 121, 138–45.

DiMoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *Mis Quarterly* 34, 373–96.

Dunning, D., Anderson, J.E., Schlosser, T., Ehlebracht, D., Fetchenhauer, D. (2014). Trust at zero acquaintance: more a matter of respect than expectation of reward. *Psychological Review*, 121, 1–17.

Fehr, E., Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425, 785–91.

Fett, A.K.J., Gromoff, P.M., Giampietro, V., Shergill, S.S., Krabbendam, L. (2014). Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Social Cognitive and Affective Neuroscience*, 9, 395–402.

Folstein, J.R., Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, 45, 152–70.

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. *Human Brain Mapping*, 32, 2207–16.

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., Coricelli, G. (2013). Reputational priors magnify striatal responses to violations of trust. *The Journal of Neuroscience*, 33, 3602–11.

Fukuyama, F. (1995). Trust: The Social Virtues and the Creation of Prosperity. New York: Free Press.

Gaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115, 811–46.

Hardin, R. (2002). Trust and Trustworthiness. New York: Russell Sage Foundation.

Hauser, T.U., Iannaccone, R., Stampfli, P., et al. (2014). The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *Neuroimage*, 84, 159–68.

Holroyd, C.B., Coles, M.G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–708.

Jing, Y., Bond, M.H. (2015). Sources for trusting most people: How national goals for socializing children promote the contributions made by trust of the in-group and the out-group to non-specific trust. *Journal of Cross-Cultural Psychology*, 46, 191–210.

Johnson, N.D., Mislin, A.A. (2011). Trust games: a meta-analysis. *Journal of Economic Psychology*, 32, 865–89.

Keselman, H.J., Wilcox, R.R., Lix, L.M. (2003). A generally robust approach to hypothesis testing in independent and correlated groups designs. *Psychophysiology*, 40, 586–96.

Kiehl, K.A., Laurens, K.R., Duty, T.L., Forster, B.B., Liddle, P.F. (2001). Neural sources involved in auditory target detection and novelty processing: an event-related fMRI study. *Psychophysiology*, 38, 133–42.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S.R., Montague, P.R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, 308, 78–83.

Knack, S., Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, 112, 1251–88.

Kosick, T.R., Tranel, D. (2011). The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia*, 49, 602–11.

Krueger, F., McCabe, K., Moll, J., et al. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences*, 104, 20084–9.

Li, C.S.R., Huang, C., Constable, R.T., Sinha, R. (2006). Imaging response inhibition in a stop-signal task: neural correlates independent of signal monitoring and post-response processing. *The Journal of Neuroscience*, 26, 186–92.

Li, T., Fung, H.H. (2013). Age differences in trust: An investigation across 38 countries. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 68, 347–55.

Long, Y., Jiang, X., Zhou, X. (2012). To believe or not to believe: trust choice modulates brain responses in outcome evaluation. *Neuroscience*, 200, 50–8.

Maas, C.J., Hox, J.J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1, 86–92.

Mayer, R.C., Davis, J.H., Schoorman, F.D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, 20, 709–34.

McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, 98, 11832–5.

Miltner, W.H., Braun, C.H., Coles, M.G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, 9, 788–98.

Nowak, M.A. (2006). Five rules for the evolution of cooperation. *Science*, 314, 1560–3.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18, 49–65.
Phan, K.L., Sripada, C.S., Angstadt, M., McCabe, K. (2010). Reputation for reciprocity engages the brain reward center. *Proceedings of the National Academy of Sciences, 107*, 13099–104.

Proudft, G.H. (2015). The reward positivity: from basic research on reward to a biomarker for depression. *Psychophysiology, 52*, 449–59.

Putnam, R.D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Raudenbush, S.W., Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd edn. Thousand Oaks, CA: Sage.

Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., Congdon, R.T., du Toit, M. (2011). *HLM 7*. Lincolnwood, IL: Scientific Software International Inc.

Richerson, P.J., Boyd, R. (2005). *Not by Genes Alone*. Chicago, IL: The University of Chicago Press.

Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*, 443–7.

Riedl, R., Javor, A. (2012). The biology of trust. *Journal of Neuroscience, Psychology, and Economics, 5*, 63–91.

Rotter, J.B. (1980). Interpersonal trust, trustworthiness, and gullibility. *American Psychologist, 35*, 1–7.

Sabbagh, M.A., Moulson, M.C., Harkness, K.L. (2004). Neural correlates of mental state decoding in human adults: an event-related potential study. *Journal of Cognitive Neuroscience, 16*, 415–26.

Sambrook, T.D., Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin, 141*, 213–35.

Stolle, D., Hooghe, M. (2004). The roots of social capital: attitudinal and network mechanisms in the relation between youth and adult indicators of social capital. *Acta Politica, 9*, 422–41.

Tomlin, D., Kayali, M.A., King-Casas, B., et al. (2006). Agent-specific responses in the cingulate cortex during economic exchanges. *Science, 312*, 1047–50.

Tzirpeopoulos, H. (2013). The Trust Game in neuroscience: a short review. *Social Neuroscience, 8*, 407–16.

Uslaner, E.M. (2002). *The Moral Foundations of Trust*. New York: Cambridge University Press.

Uslaner, E.M. (2004). Trust and corruption. In: Lambsdorf, J. G., Taube, M. and Schramm, M., editors. *Corruption and the New Institutional Economics*, (pp. 76–92). London: Routledge.

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S.A.R.B., Crone, E.A. (2009). What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Social Cognitive and Affective Neuroscience, 4*, 294–304.

van Overwalle, F., Baetens, K. (2009). Understanding others’ actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage, 48*, 564–84.

Wardle, M.C., Fitzgerald, D.A., Angstadt, M., Sripada, C.S., McCabe, K., Phan, K.L. (2013). The caudate signals bad reputation during trust decisions. *PLoS One, 8*, e68884.

Yamagishi, T. (2011). *Trust: The Evolutionary Game of Mind and Society*. New York: Springer.

Yamagishi, T., Akutsu, S., Cho, K., Inoue, Y., Li, Y., Matsumoto, Y. (2015). Two-component model of general trust: predicting behavioral trust from attitudinal trust. *Social Cognition, 33*, 436–58.

Zak, P.J., Knack, S. (2001). Trust and growth. *Economic Journal, 111*, 295–321.