

Neural correlates of trust

Frank Krueger*, Kevin McCabe†, Jorge Moll‡, Nikolaus Kriegeskorte§, Roland Zahn¶, Maren Strenziok*, Armin Heinecke||, and Jordan Grafman*.*.*

*Cognitive Neuroscience Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892-1440; †The Center for the Study of Neuroeconomics, George Mason University, Fairfax, VA 22030; ‡Cognitive and Behavioral Neuroscience Unit, LABS-D'Or Hospital Network, 2228-080, Rio de Janeiro, Brazil; §Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892-1148; ¶Neuroscience and Aphasia Research Unit, School of Psychological Sciences, University of Manchester, Manchester M13 9PL, United Kingdom; and ||Brain Innovation B.V., Universiteitssingel 40, 6201 BC Maastricht, The Netherlands

Communicated by Vernon L. Smith, George Mason University, Fairfax, VA, October 25, 2007 (received for review July 12, 2007)

Trust is a critical social process that helps us to cooperate with others and is present to some degree in all human interaction. However, the underlying brain mechanisms of conditional and unconditional trust in social reciprocal exchange are still obscure. Here, we used hyperfunctional magnetic resonance imaging, in which two strangers interacted online with one another in a sequential reciprocal trust game while their brains were simultaneously scanned. By designing a nonanonymous, alternating multi-round game, trust became bidirectional, and we were able to quantify partnership building and maintenance. Using within- and between-brain analyses, an examination of functional brain activity supports the hypothesis that the preferential activation of different neuronal systems implements these two trust strategies. We show that the paracingulate cortex is critically involved in building a trust relationship by inferring another person's intentions to predict subsequent behavior. This more recently evolved brain region can be differently engaged to interact with more primitive neural systems in maintaining conditional and unconditional trust in a partnership. Conditional trust selectively activated the ventral tegmental area, a region linked to the evaluation of expected and realized reward, whereas unconditional trust selectively activated the septal area, a region linked to social attachment behavior. The interplay of these neural systems supports reciprocal exchange that operates beyond the immediate spheres of kinship, one of the distinguishing features of the human species.

attachment | neuroeconomics | reward | social | functional MRI

Unlike other species, humans are trustful and cooperate with genetically unrelated strangers, with individuals they will never meet again, or even when reputation and gains are absent (1, 2). Recent studies in experimental economics and social neuroscience have started to explore the neurobiology of trust (2–6) and cooperation (7–10) in reciprocal exchange. Reciprocal behavior allows the formation of partnerships that can produce mutual advantages for cooperators and thus can be selected for maximizing evolutionary fitness (11). Reciprocity generally involves a first mover who must trust another person to give the other person an opportunity to reciprocate (12). Typically in a partnership, the person who moves first will vary frequently. In laboratory experiments, trusting behavior can be reliably reproduced (13, 14) although with significant individual variation with respect to both experience (3, 15) and context (5, 16).

In this paper we look at first movers' decisions to trust. Trusting is always risky given the unpredictability of the intentions of the partner in a social exchange (17). A trust relationship is built on each partner's decisions to trust and reciprocate. To build a trust relationship, partners must learn that they can depend on each other. One model of this process is the goodwill accounting model (18), which is based on the empirical practice of taking into account the value of ongoing partnerships. Partners accumulate goodwill toward each other and evaluate it against the constantly changing risk of defection. Without balanced goodwill, partners cannot synchronize their mutual cooperation. In this regard, individuals can use one of two strat-

egies that imply different benefits and costs (19), conditional trust or unconditional trust.

Conditional trust assumes that one's partner is self-interested and estimates the expected value of one's strategy with respect to the benefits of cooperating, the risk of defection, and the future value of past decisions; it causes less balanced goodwill and results in greater variance in cooperative decisions and, therefore, is cognitively more costly to maintain. In contrast, unconditional trust assumes that one's partner is trustworthy and updates the value of one's partner with respect to their characteristics and past performance; balanced goodwill occurs more quickly, allowing the partners to attain high levels of synchronicity in their decisions and, therefore, is cognitively less costly to maintain. In this work, an examination of functional brain activity supports the hypothesis that the preferential activation of different neuronal systems implements these two trust strategies.

We used event-related hyperfunctional magnetic resonance imaging (20) (hyperfMRI), in which two strangers of the same gender (44 participants, 11 female and 11 male pairs), each in a separate MRI scanner, interacted with one another in a sequential reciprocal trust game while their brains were simultaneously scanned [see *Materials and Methods* and [supporting information \(SI\) Fig. 5](#)]. Participants were asked to make sequential decisions for monetary payoffs (low, medium, or high in cents) presented in a binary game tree (Fig. 1a). The first mover can either quit the game by not trusting the second mover, resulting in a small equal payoff for both; or the first mover can continue the game by trusting the second mover, hoping to receive a better payoff. The second mover can reciprocate the first mover's trust, giving them both a higher payoff, or defect on the first mover's trust, resulting in an even larger payoff for the second mover and a payoff of zero for the first mover. Partners played 36 voluntary trust games and 16 control games (SI Fig. 6). In the control games, partners followed the same time line as in trust games, but they did not have to interact with one another and merely had to choose between lower and higher monetary rewards (SI Fig. 7).

Previous studies used anonymous single or multi-round interactions, in which individuals maintained their roles as first and second mover throughout reciprocal exchange (3, 4, 8, 21). In their natural environment, however, partners are not anonymous and often alternate in their roles while interacting over long time periods. To improve the ecological validity of the task, we let

Author contributions: F.K., K.M., J.M., and J.G. designed research; F.K., J.M., R.Z., and M.S. performed research; N.K. and A.H. contributed analytic tools; F.K., N.K., and A.H. analyzed data; and F.K., K.M., and J.G. wrote the paper.

The authors declare no conflict of interest.

**To whom correspondence should be addressed at: Cognitive Neuroscience Section, National Institute of Neurological Disorders and Stroke, Building 10, Room 7D43, MSC 1440, National Institutes of Health, Bethesda, MD 20892-1440. E-mail: Grafman@ninds.nih.gov.

This article contains supporting information online at www.pnas.org/cgi/content/full/0710103104/DC1.

© 2007 by The National Academy of Sciences of the USA

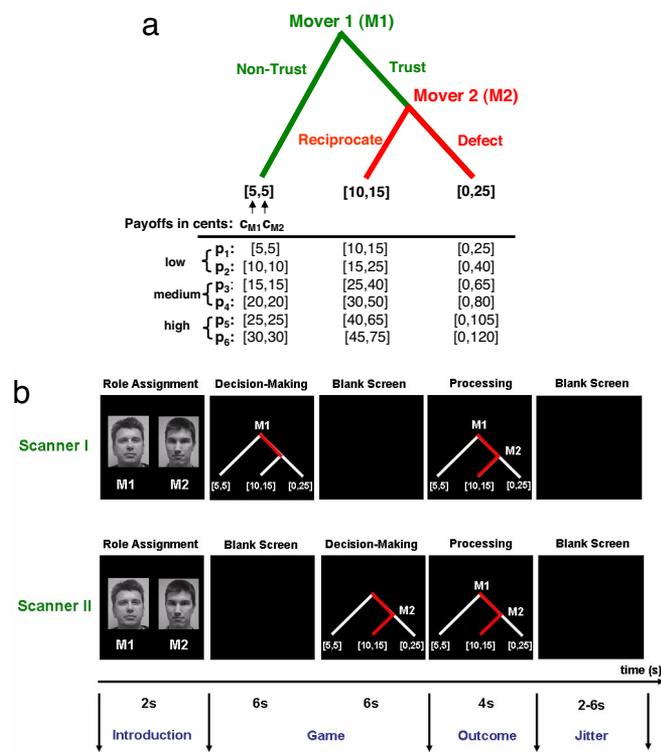


Fig. 1. Experimental design. (a) Voluntary trust game. Partners made sequential decisions as first mover (M1) and second mover (M2) for payoffs in cents [c_{M1}, c_{M2}] presented in a binary decision tree. M1 can choose left (nontrust) and quit the game with a small payoff for M1 and M2 (e.g., [5, 5]) or can choose right (trust) to continue the game. M2 can then choose left (reciprocate), giving them both a higher payoff (e.g., [10, 15]) or choose right (defect), resulting in an even larger payoff to M2 and a payoff of zero to M1 (e.g., [0, 25]). Payoffs (p_1 – p_6) were split into three types: low (p_1 – p_2), medium (p_3 – p_4), and high (p_5 – p_6). (b) Time line for a single trust game. Partners were introduced by seeing each other by webcam, and digital photographs were taken to be used for game trials. A 2-s introductory screen informed partners of the role that they were playing (M1 or M2). M1 saw the game tree, had to make a decision (nontrust or trust) within 6 s, and waited 6 s for M2's decision while seeing a blank screen. M2 saw a blank screen for 6 s, saw the game tree with M1's decision, and had to make a decision (reciprocate or defect) within 6 s. If M1 had chosen not to trust M2, the game was over, and M2 saw M1's decision for 6 s. Partners saw the outcome of the game for 4 s followed by a blank screen with a jittered interstimulus interval of 2–6 s.

pairs of strangers play multirounds of nonanonymous voluntary trust games while alternating their roles as first and second mover (13) (Fig. 1b and see *SI Procedures*). Therefore, trust becomes bidirectional for both partners, allowing us to explore partnership building and maintenance while partners develop mental models of one another (6, 14). Previous research has shown that the striatum (caudate head) of second movers in a social reciprocal exchange encodes a signal coding for the expectation of trusting behavior by their partners (3). However, in this work we are interested in first movers' decisions to trust and the underlying mechanisms of conditional and unconditional trust in developing a trust partnership. The design of our experiment allowed us to address two questions: (i) which brain regions modulate decisions to trust in a partnership and (ii) which brain regions modulate different trust strategies over time in a partnership.

Results

Multisubject Level Analyses. Data on decisions in voluntary trust games showed that first movers decided to trust more often than not to trust (84% vs. 16%) ($t_{21} = 7.16$, $P < 0.001$), and second

movers reciprocated more often than they defected (77% vs. 7%) ($t_{21} = 9.55$, $P < 0.001$) (SI Fig. 8a). For the control games, participants chose more often the higher monetary reward (93%) ($t_{43} = 18.67$, $P < 0.001$). Decision times (mean \pm SEM) between trust (2,349 \pm 73 ms) and control (2,435 \pm 70 ms) games did not differ ($t_{43} = -1.61$, $P = 0.114$). Before and after scanning, partners were asked to rate their closeness and partnership to one another on 11-point Likert scales. Participants felt closer to each other ($t_{43} = -3.86$, $P < 0.001$) and ranked themselves more as a partner to the other person ($t_{43} = -2.21$, $P < 0.033$) after the experiment (SI Fig. 8b). Using a general linear model (GLM) analysis, we first sought brain regions whose blood oxygenation level-dependent (BOLD) responses were recruited for decisions to trust. Decisions to trust contrasted with the control condition activated the paracingulate cortex (PcC) [peak voxel (x, y, z): 5, 39, 22] and the septal area (SA) (together with the adjoining hypothalamus) (peak voxel: -4, 4, -3) (Fig. 2a and b and SI Table 1).

Group-Level Analyses. Next we explored the dynamic role of the PcC and SA in supporting conditional and unconditional trust strategies. Therefore, we arbitrarily divided the experiment into two stages under the assumption that ongoing participation in games during stage I represented partnership building and during stage II, partnership maintenance (SI Fig. 6). In addition, we split pairs of participants into two equal-sized groups based on their decision patterns in the experiment. For the nondefector group (11 pairs; 6 female pairs; mean age, 28.9 \pm 7.0 years, range 22–47; mean education level, 17.4 \pm 2.1 years, range 16–23) neither player ever defected their partner's decision to trust, whereas for the defector group (11 pairs; 5 female pairs; mean age, 27.7 \pm 7.3 years, range 21–51; mean education level, 17.3 \pm 2.3 years, range 12–23) partners experienced some defections during the experiment. The nondefector and defector groups did not differ significantly in age ($t_{22} = 0.55$; $P = 0.588$) or education ($t_{22} = 0.79$; $P = 0.787$).

Trust in the nondefector group was higher than in the defector group ($F_{1,42} = 26.62$, $P < 0.001$) and increased across stages ($F_{1,21} = 5.86$, $P < 0.025$) (SI Fig. 8c). In contrast, trust in the defector group decreased across stages ($F_{1,21} = 4.37$, $P < 0.048$) and depended on the payoff type ($F_{2,42} = 9.57$, $P < 0.001$). In the maintenance stage, trust in the defector group occurred more often in the low-payoff games compared with the medium- and high-payoff games ($F_{1,21} = 23.25$, $P < 0.001$) and in the medium-compared with the high-payoff games ($F_{1,21} = 4.91$, $P < 0.038$) (Fig. 3a). Decision times for trust games became faster for the nondefector group across stages ($F_{1,21} = 5.86$, $P < 0.025$), and decision times accelerated by 20% for first movers ($t_{21} = 5.15$, $P < 0.001$) and by 10% for second movers ($t_{21} = 2.71$, $P < 0.013$) (Fig. 3b). After the experiment, partners in the nondefector group felt closer to each other ($t_{21} = -3.24$, $P < 0.004$) and ranked themselves as more of a partner to the other person ($t_{21} = -2.99$, $P < 0.007$) (SI Fig. 8d). Finally, the defector group earned less money than the nondefector group ($F_{1,42} = 9.08$, $P < 0.005$), and earnings decreased for the defector group but increased for the nondefector group across stages ($F_{1,42} = 5.79$, $P < 0.021$) (SI Fig. 8e).

In a planned follow-up region of interest (ROI) analysis, we derived the parameter estimates from the beforehand identified PcC to investigate how first movers in the nondefector and defector groups engaged a PcC-supported mentalizing system (22) to build different trust strategies across stages. First movers in the nondefector and defector groups made different use of the mentalizing system across stages ($F_{1,42} = 9.14$, $P < 0.004$). The nondefector group showed a higher activation in the PcC compared with the defector group in the building stage ($t_{42} = 2.72$, $P < 0.010$). The nondefector group showed a decrease in activation ($t_{21} = 2.10$, $P <$

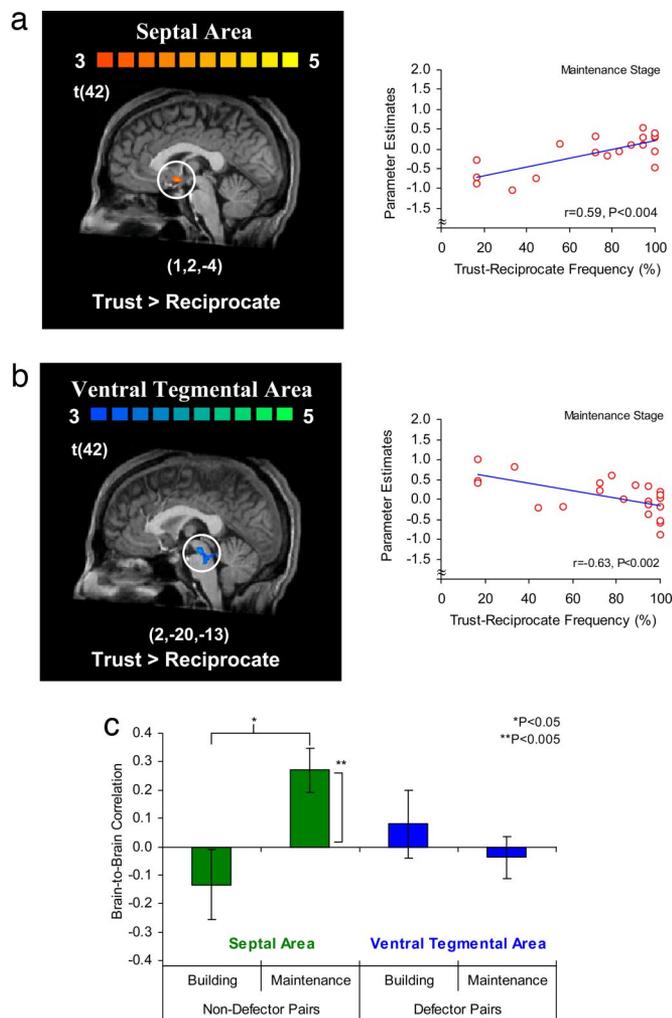


Fig. 4. Brain responses for trust maintenance. (a) Unconditional trust. In the nondefector group, decisions to trust contrasted with decisions to reciprocate revealed a higher activation in the SA compared with the defector group. Pairs who showed the highest trust-reciprocate history (frequency) in their decisions also showed the highest activation (parameter estimates) in the SA. (b) Conditional trust. In the defector group, decisions to trust contrasted with decisions to reciprocate revealed a higher activation in the VTA compared with the nondefector group. Pairs who showed the lowest trust-reciprocate history (frequency) in their decisions also showed the highest activation (parameter estimates) in the VTA. (c) Brain-to-brain correlation (\pm SEM). In the nondefector group, brain-to-brain correlations increased in the SA across stages. In the maintenance stage, partners in the nondefector group became synchronized in their SA BOLD amplitudes as first movers in adjacent trials of trust games.

in the building and maintenance stages were obtained (SI Fig. 10). After confirming that the obtained population distributions (D1–D4) were normally distributed (one-sample Kolmogorov–Smirnov test), mean brain-to-brain correlations for the nondefector and defector groups were computed for the building and maintenance stage of the experiment and then compared with their population distribution means (one-sample t test). If the mean brain-to-brain correlations for the nondefector and defector groups differed significantly from its population distribution means, then we assumed that partners became “synchronized” in their decision patterns. Brain-to-brain correlations only increased in the SA region for the nondefector group across stages ($t_{10} = -2.40, P < 0.038$), and only partners in the nondefector group became synchronized in their SA BOLD

amplitudes as first movers in adjacent trials of trust games ($r = 0.27, P < 0.005$) (Fig. 4c).

Discussion

We used event-related hyperfMRI to investigate the neural correlates of conditional and unconditional trust in two-person reciprocal exchange. First, we identified two distinct regions that underlie decisions to trust in a partnership. Decisions to trust contrasted with the control condition activated the PcC. Previous research has shown that the PcC not only represents our own thoughts, feelings, and beliefs, but also represents the mental states of other people (6, 21, 23–25). Mentalizing (22) is a unique human characteristic and can be observed only in a most rudimentary form in great apes (26) and has never been observed in monkeys (27). In building mutual goodwill, partners must infer each other’s intentions to determine whether to trust their partners and whether their partners will reciprocate their trust in the future.

Decisions to trust contrasted with the control condition also activated the SA (together with the adjoining hypothalamus), a limbic region that has been demonstrated to modulate various aspects of social behavior including social memory and learning (28). In addition, the SA plays a putative role in controlling anterior hypothalamic functions and the release of the neuropeptides vasopressin and oxytocin and itself contains receptors for those neuropeptides (29–31). Besides the well known physiological functions of oxytocin in milk letdown and during labor, oxytocin is a key mediator in facilitating various complex social behaviors, including maternal care (31), pair bonding (31), social recognition (32), and the ability to form social attachment (33–35). There is evidence that greater first mover trust can be induced in strangers by the nasal administration of synthetic oxytocin (36). Because synthetic oxytocin increases trust, we surmised that partners recruited the SA to encode goodwill to maintain their trust partnership. Results from pre- and postquestionnaire ratings support our view demonstrating that partners felt significantly closer to each other and ranked themselves as being more of a partner to the other person after the experiment.

After identifying two distinct regions that underlie decisions to trust in a partnership, we next explored the dynamic role of these regions in supporting conditional and unconditional trust strategies. We arbitrarily divided the experiment into two stages: partnership-building stage and partnership stage. In addition, we identified two equal-sized groups based on their decision patterns throughout the experiment: a nondefector group in which neither player ever defected on their partners’ decision to trust, and a defector group in which partners experienced some defections during the experiment. We hypothesized that the nondefector and defector groups would adapt different trust strategies across stages of the experiment. Results revealed that first movers in the nondefector and defector groups made different use of the mentalizing system, resulting in two different neural systems for maintaining unconditional and conditional trust.

Unconditional trust assumes that one’s partner is trustworthy. During the building stage, first movers in the nondefector group showed higher activation in the PcC compared with first movers in the defector group. Through mentalizing, partners of this group verified their prior trustworthy assumption, updated the value of one’s partner’s strategy with respect to their past performance, and maintained a balanced goodwill toward each other, allowing them to avoid defections. By developing “better” mental models in this early stage, partners in the nondefector group accumulated sufficient mutual goodwill to become socially attached to each other and adopted an unconditional trust strategy.

During the maintenance stage, the nondefector group showed a higher activation in the SA compared with the defector group.

Across groups, pairs who showed the highest trust-reciprocate history in their decisions also showed the highest activation in this region. Furthermore, analyses of pre- and postscan behavioral ratings confirmed that only nondefector pairs felt significantly closer to each other and ranked themselves as being more of a partner to the other person after the experiment. Through early mentalizing, partners in the nondefector group must have balanced goodwill more quickly, allowing them to become synchronized in their decision patterns. Brain-to-brain correlations only increased in the SA region for the nondefector group across stages, and only partners in the nondefector group became synchronized in their SA BOLD amplitudes as first movers in adjacent trials of trust games. Synchronization in the SA led to social attachment associated with a significant decrease in activation in the PcC during the maintenance stage. By adopting this cognitively less costly strategy, decision times became significantly faster for the nondefector group across stages of the experiment.

Conditional trust assumes that one's partner is self-interested. During the building stage, first movers in the defector group showed less activation in the PcC compared with the nondefector group. Through less mentalizing in the building stage, partners in this group produced higher errors in the inferences of second movers' goodwill toward them, resulting in less balanced goodwill and, therefore, in less overall trust compared with the nondefector group. More importantly, they started to trust more in the low-payoff games and less in the high-payoff games. This decision pattern implies that defectors were adapting a conditional trust strategy by evaluating the expected value of one's strategy with respect to the risks and benefits of cooperation.

During the maintenance stage, the defector group showed higher activations in the VTA compared with the nondefector group, a region linked to the dopaminergic mesolimbic reward system providing a general reinforcement mechanism to encode expected and realized reward (37, 38). Across groups, pairs who shared the lowest trust-reciprocate history in their decisions also showed the highest activation in this region. By adopting a cognitively more costly strategy, partners in the defector group showed a significant increase in activation in the PcC over the experiment. Through more mentalizing in this late stage, first movers in the defector group tried to develop more accurate models about the likelihood of their partner's choices so that they could make a more advantageous decision about when to trust. The conditional trust strategy paid off less over time as total earnings decreased for the defector group (but increased for the nondefector group) across stages.

In conclusion, we applied event-related hyperfMRI to identify the neural correlates of conditional and unconditional trust when paired strangers interacted with one another in a sequential reciprocal trust game. By designing a nonanonymous, alternating multiround game, trust became bidirectional, and partnership building and maintenance were explored. Our findings extend previous knowledge of the neural basis of trust in two-person reciprocal exchange and broaden our understanding of how trust relationships are built and maintained over time. First, the PcC is critically involved in building a trust relationship by inferring another person's intentions to predict subsequent behavior. Second, this more recently evolved brain region can be differently engaged to recruit more primitive neural systems in maintaining conditional and unconditional trust in a partnership. Conditional trust selectively activated the VTA, a region linked to the evaluation of expected and realized reward, whereas unconditional trust selectively activated the SA, a region linked to social attachment behavior. The interplay of these neural systems supports reciprocal exchange that operates beyond the immediate spheres of kinship, one of the distinguishing features of the human species.

Materials and Methods

Subjects. Forty-four normal volunteers (22 women and 22 men matched by age (mean age 28.3 ± 7.1 years, range 21–51) and education (mean education level 17.3 ± 2.2 years, range 12–23) participated for financial compensation in the fMRI experiment. All participants were native English speakers and right-handed as determined by the Edinburgh Handedness Inventory (95.3 ± 8.7 , range 65–100) (39). All volunteers underwent a neurological examination and had normal or corrected-to-normal vision, no history of medical, psychiatric, or neurological diagnoses, and were not taking any medication. Informed consent was obtained according to procedures approved by the National Institute of Neurological Disorders and Stroke (NINDS) Institutional Review Board.

Data Acquisition and Analysis. Imaging was performed on two 3-Tesla MRI whole-body scanners (General Electric) equipped with standard circularly polarized head coils located in the NMR Research Center at the National Institutes of Health. Head motion was restricted by using foam pads placed around the participant's head. Anatomical images (T1-weighted three-dimensional MP-RAGE sequence: TR, 9.7 ms; TE, 4.0 ms; flip angle, 12° ; field of view, 240 mm; matrix size, 256×256 ; thickness, 1.2 mm; in-plane resolution, 0.8594×0.8594 mm²) and functional images (two-dimensional gradient EPI sequence, optimized for BOLD contrast: TR, 2s; TE, 30 ms; flip angle, 90° ; thickness, 6 mm; number of slices, 22; field of view, 240 mm; voxel dimensions, $3.75 \times 3.75 \times 6$ mm) were acquired. For each of the two functional runs, 291 volume images per run were taken parallel to AC–PC line. The first five volumes were discarded to allow for T1 equilibration effects.

Behavioral data analyses were carried out by using SPSS software (SPSS, Inc.). $P < 0.05$ was used for all behavioral analyses (two-tailed). Image analyses were performed by using BrainVoyager QX (Brain Innovation) and custom-written scripts in MATLAB (The MathWorks, Inc.). The following data-preprocessing steps were applied: slice scan time correction ("sinc" interpolation), linear trend removal, temporal high-pass filtering to remove low-frequency nonlinear drifts of three or fewer cycles per time course, spatial smoothing (8-mm FWHM), and three-dimensional motion correction to detect and correct for small head movements by spatial alignment of all participants to the first volume by rigid body transformation. Estimated translation and rotation parameters were inspected and never exceeded 2 mm or 2° . To transform the functional data into Talairach space (40), the functional time series data of each subject were first coregistered with the subject's three-dimensional anatomical data set and resampled to $3 \times 3 \times 3$ mm³ isotropic voxels, resulting in a normalized four-dimensional volume time course data.

A GLM corrected for first-order serial correlation was applied (41). Random-effect analyses were performed on the multisubject level (group data: $n = 44$) and group-level (subgroup data: nondefectors (ND), $n = 22$; defectors (D), $n = 22$) to explore brain regions that are associated with decisions to trust and trust development over time. For each participant, regressors were created based on individual decisions as M1 and M2 in voluntary trust games (VTGs) (M1, T or NT; M2, R or D) and control games (CGs) (M1, C1; M2, C2) over both functional runs (run1: building stage, SI and run2: maintenance stage, SII). The regression model consisted of a set of 26 regressors (8 VTG and 5 CG regressors per stage). Regressor time courses were adjusted for the hemodynamic response delay by convolution with a double-gamma hemodynamic response function (42). Multiple regression analyses were performed independently for the time course of each individual voxel. After computing the coefficients for all regressors, t tests were performed between coefficients of

different conditions. A statistical model on the multisubject level was fit for one linear contrast to explore brain regions that were associated with decisions to trust: trust > control ($T_{SI} + T_{SII} > C_{SI} + C_{SII}$). Furthermore, statistical models on the group level (ND, D) were fit for two linear contrasts to explore trust development over time (SI, SII): (1) Trust > Reciprocate for (i) partnership building [$(T_{SI} > R_{SI})_{ND} > (T_{SI} > R_{SI})_D$] and (ii) partnership maintenance [$(T_{SII} > R_{SII})_{ND} > (T_{SII} > R_{SII})_D$].

Functional analyses were performed at the whole-brain level and within predefined *a priori* regions of interest: PcC, SA, VTA, and striatum. The Talairach coordinates (x, y, z) for the centers of those regions (left and right side) were based on the most significant voxel of the clusters activated in previous fMRI studies. The coordinates for the PcC (1, 46, 20) were based on a metaanalysis that mapped medial frontal cortex activations during mentalizing (23). The coordinates for the SA (6, 11, 4) and the striatum (17, 4, 11) were based on our previous study on decision-making about charitable donations (33). The coordinates for the VTA (0, -10, -10) were based on a study that reviewed fMRI studies that investigated reward such as cocaine injection and receipt of money (35). *A priori* regions hypothesized to be active were tested for activity by using a small-volume correction of a sphere of 10 mm and a false-discovery rate (FDR) (43) with a threshold of $q(\text{FDR}) < 0.05$ (small volume-corrected) and a cluster size threshold of 100 mm³. Because of a lack of *a priori* hypotheses for other brain areas, non-*a priori* regions that were activated were reported in a whole-brain analysis by using a FDR with a threshold of $q(\text{FDR}) < 0.005$ (corrected) and a cluster size threshold of 100 mm³. Statistical images were superimposed on a template structural brain in Talairach space and thresholded at $P < 0.005$, uncorrected, with extent threshold of 100 mm³ ($t = 3.00$, random effects). Parameter estimates (mean β weights) from ROIs were derived from the peak voxel activation and surrounding voxels encompassing 54 mm³.

Finally, brain-to-brain correlation between partners' BOLD amplitude responses in ROIs were computed to measure partners' intrapair synchronization when they were first movers in

adjacent trials of trust games (SI Fig. 9). For each ROI, time series of BOLD responses were derived from subjects' normalized four-dimensional brain datasets after identifying the peak of activation and surrounding voxels encompassing 54 mm³. After the functional time series were averaged and Z-transformed, BOLD values selected from the time of peak responses and the peak's two flanking points were averaged for partners' decisions as first movers. Mean BOLD amplitudes were arranged pairwise, and brain-to-brain correlations were computed for each pair of the nondefector and defector group separately for the building and maintenance stage of the experiment.

Furthermore, to rule out the possibility that brain-to-brain correlations for nondefector and defector pairs were introduced by the design of the game, scanner noise, or the order of preprocessing procedures, brain-to-brain correlations were also computed for randomly reassigned pairs ($k = 2$) of subjects ($n = 44$). Permutation steps were repeated until all $[(n)_k = n! / (k!(n-k)!)] = 946$ combinations had been examined and the population distributions for both brain regions in the building and maintenance stages were obtained (44). Afterward, mean brain-to-brain correlations for the nondefector and defector groups were computed for the building and maintenance stage of the experiment and compared with their population distribution means. If the mean brain-to-brain correlations of the nondefector and defector groups differed significantly from its population distribution means, then we assume that partners became synchronized in their BOLD amplitudes decision patterns as first movers in adjacent trials of trust games.

We thank E. Wassermann for performing the neurological exams, N. Armstrong and M. Ulrich for help in various stages of this project, and A. Butz for help in programming the fMRI experiment. We thank J. Bodurka and S. Marrett for technical assistance for the hyperfMRI setup and E. Condon for scanning assistance. The work was supported in part by a National Institute of Neurological Disorders and Stroke/National Institutes of Health (NINDS/NIH) postdoctoral competitive fellowship award (to F.K.) and the Intramural Research Program of the Cognitive Neuroscience Section, NINDS/NIH.

- Smith V (1998) *South Econ J* 65:1–19.
- de Quervain DJ, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, Buck A, Fehr E (2004) *Science* 305:1254–1258.
- King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR (2005) *Science* 308:78–83.
- Tomlin D, Kayali MA, King-Casas B, Anen C, Camerer CF, Quartz SR, Montague PR (2006) *Science* 312:1047–1050.
- Delgado MR, Frank RH, Phelps EA (2005) *Nat Neurosci* 8:1611–1618.
- McCabe K, Houser D, Ryan L, Smith V, Trouard T (2001) *Proc Natl Acad Sci USA* 98:11832–11835.
- Decety J, Jackson PL, Sommerville JA, Chaminade T, Meltzoff AN (2004) *Neuroimage* 23:744–751.
- Rilling J, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C (2002) *Neuron* 35:395–405.
- Singer T, Kiebel SJ, Winston JS, Dolan RJ, Frith CD (2004) *Neuron* 41:653–662.
- Singer T, Seymour B, O'Doherty JP, Stephan KE, Dolan RJ, Frith CD (2006) *Nature* 439:466–469.
- Axelrod R, Hamilton WD (1981) *Science* 211:1390–1396.
- Klin A, Volkmar FR, Sparrow SS, Cicchetti DV, Rourke BP (1995) *J Child Psychol Psychiatry* 36:1127–1140.
- Marchionni E, Galimberti CA, Soragna D, Ferrandi D, Maurelli M, Ratti MT, Bo P, Montalbetti L, Albergati A, Savoldi F (1995) *Neurology* 45:33–37.
- Camerer CF (2003) *Behavioral Game Theory: Experiments in Strategic Interactions* (Princeton Univ Press, Princeton).
- McCabe K, Smith V (2000) *Proc Natl Acad Sci USA* 97:3777–3781.
- Burnham T, McCabe K, Smith V (2000) *J Econ Behav Organ* 43:57–73.
- Fehr E, Fischbacher U (2003) *Nature* 425:785–791.
- McCabe K, Smith V (2001) in *Bounded Rationality: The Adaptive Toolbox*, eds Gigerenzer G, Selten R (MIT Press, Cambridge), pp 319–340.
- Williamson OE (1993) *J Law Econ* 36:453–486.
- Battig WF, Montague WE (1969) *J Exp Psychol Monogr* 80:1–46.
- Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD (2004) *Neuroimage* 22:1694–1703.
- Jahanshahi M, Dirnberger G, Fuller R, Frith CD (2000) *Neuroimage* 12:713–725.
- Amodio DM, Frith CD (2006) *Nat Rev Neurosci* 7:268–277.
- Gallagher HL, Jack AI, Roepstorff A, Frith CD (2002) *Neuroimage* 16:814–821.
- Goel V (1995) *Sketches of Thought* (MIT Press, Cambridge).
- Povinelli DJ, Preuss TM (1995) *Trends Neurosci* 18:418–424.
- Cheney DL, Seyfarth RM (1990) *How Monkeys See the World: Inside the Mind of Another Species* (University of Chicago Press, Chicago).
- Numan R (2000) *The Behavioral Neuroscience of the Septal Region* (Springer, New York).
- Loup F, Tribollet E, Dubois-Dauphin M, Dreifuss JJ (1991) *Brain Res* 555:220–232.
- Powell EW, Rorie DK (1967) *Am J Anat* 120:605–610.
- Rilling JK, Insel TR (1999) *J Hum Evol* 37:191–223.
- Choleris E, Gustafsson JA, Korach KS, Muglia LJ, Pfaff DW, Ogawa S (2003) *Proc Natl Acad Sci USA* 100:6192–6197.
- Moll J, Krueger F, Zahn R, Pardini M, de Oliveira-Souza R, Grafman J (2006) *Proc Natl Acad Sci USA* 103:15623–15628.
- Bartels A, Zeki S (2004) *Neuroimage* 21:1155–1166.
- Aron A, Fisher H, Mashek DJ, Strong G, Li H, Brown LL (2005) *J Neurophysiol* 94:327–337.
- Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E (2005) *Nature* 435:673–676.
- Andreasen NC, Nopoulos P, Schultz S, Miller D, Gupta S, Swayze V, Flaum M (1994) *Acta Psychiatr Scand Suppl* 384:51–59.
- Fiorillo CD, Tobler PN, Schultz W (2003) *Science* 299:1898–1902.
- Oldfield RC (1971) *Neuropsychologia* 9:97–113.
- Talairach J, Tournoux P (1988) *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York).
- Friston KJ, Holmes AP, Worsley KJ (1999) *NeuroImage* 10:1–5.
- Buchel C, Holmes AP, Rees G, Friston KJ (1998) *Neuroimage* 8:140–148.
- Genovese CR, Lazar NA, Nichols T (2002) *Neuroimage* 15:870–878.
- Good P (1995) *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses* (Springer, New York).