Effect of relationship experience on trust recovery following a breach

Oliver Schilke[a,1], Martin Reimann[b], and Karen S. Cook[c,1]

[a]Department of Sociology, University of California at Los Angeles, Los Angeles, CA 90095; [b]Department of Marketing, Eller College of Management, University of Arizona, Tucson, AZ 85721-0108; and [c]Department of Sociology, Stanford University, Stanford, CA 94305

Contributed by Karen S. Cook, August 6, 2013 (sent for review April 24, 2013)

A violation of trust can have quite different consequences, depending on the nature of the relationship in which the trust breach occurs. In this article, we identify a key relationship characteristic that affects trust recovery: the extent of relationship experience before the trust breach. Across two experiments, this investigation establishes the behavioral effect that greater relationship experience before a trust breach fosters trust recovery. A neuroimaging experiment provides initial evidence that this behavioral effect is possible because of differential activation of two brain systems: while decision making after early trust breaches engages structures of a controlled social cognition system (C-system), specifically the anterior cingulate cortex and lateral frontal cortex, decision making after later trust breaches engages structures of an automatic social cognition system (X-system), specifically the lateral temporal cortex. The present findings make contributions to both social psychological theory and the neurophysiology of trust.

Trust is known to facilitate collective undertakings across a variety of contexts (1–3). Unfortunately, few social relationships endure without a violation of trust (i.e., trust breach), and once broken, trust is notoriously hard to reestablish (4–6). This article aims to improve our understanding of the process of trust recovery after a trust breach (7, 8). Specifically, we address the question of why certain kinds of relationships recover better from a trust breach than others and focus on the role of prior relationship experience, one of the most basic and fundamental characteristics of social relationships (9). We propose that greater relationship experience before a trust breach facilitates trust recovery. In other words, the longer the relationship history before a trust breach, the more likely is recovery from such a breach. However, if the trust breach occurs in an earlier stage when trust is still partial, tentative, and fragile, we expect trust to be particularly susceptible to enduring damage by a trust breach, suggesting a weaker recovery of trust.

Beyond identifying a direct relationship between prebreach experience and postbreach trust levels, we analyze key mechanisms underlying this relationship. We propose that if little relationship experience exists and a trust breach occurs, an individual engages in more conscious learning, complex planning, and increased problem solving with respect to the social relationship. Prior research implies that such cognitive processes may be key to trust recovery after early trust breaches, but this research has not yet provided actual empirical support (7). Furthermore, as relationships mature, they become increasingly habitualized and "taken for granted," fostering reconciliation after a trust breach. This notion is supported by prior literature, which implies that over time trustors tend to develop mental models of their counterpart that provide a basis for habitualized decision making (10) and make a negative deviation (such as a trust breach) more likely to be seen as the exception rather than the rule (7, 11).

We initially tested the hypothesis that prior relationship experience increases the amount of trusting behavior after a trust breach occurs. We tested this hypothesis in two behavioral studies, one of which was conducted by means of an online experiment among adult participants from the general population, and another, which was conducted under laboratory conditions while adult student participants were undergoing functional MRI (fMRI). Previous investigators have used both online experiments conducted over the Internet (12, 13) and neuroimaging experiments (10, 14–16) to study contemporary issues of social exchange and trust, and we followed their methodological choices. Particularly, the latter method (fMRI) has allowed previous research efforts to tap into the neurophysiological correlates of trust, and thus to gain some insights into the inner functional processes that precede trusting behavior (17).

Participants in our study engaged in an established repeated-measures trust–honor game (18). In this game, participants could either keep \$8 on a given trial or transfer it to a partner, in which case the money would be tripled and the partner would decide whether to reciprocate and equally share the \$24 or to defect and keep all of the money. Participants were told they would be playing with other study participants. In reality, for the purposes of experimental control, participants were actually playing against a computer with a preprogrammed set of choices that were identical across conditions, with the exception of the manipulation. In one condition, the computer defected early, whereas in the other condition the computer violated the participant's trust only later in the experiment. This experimental procedure allowed us to analyze to what extent participants would recover from their partners' trust breach and transfer money again to their partner, indicating renewed trust. We consider the deception involved in our procedures methodologically necessary, as it allowed us to cleanly implement our relationship experience manipulation while also avoiding excessive waste of data collection resources (19). The behavioral results obtained provide consistent empirical support for the hypothesized positive main effect of relationship experience on trust recovery.

Significance

Will people be more likely to forgive a breach of trust in an earlier or later stage of an interpersonal relationship? The present article reports behavioral and neurophysiological experiments that speak to this important question. Results show that trust recovery is facilitated with increasing relationship experience. Differential activation in the controlled social cognition system (C-system) and the automatic social cognition system (X-system) indicate that decision making is less controlled and more automatic following a later as opposed to an earlier trust breach. These findings have important implications for the study of trust recovery after a breach, as well as the neuroscience of trust.

SOCIAL SCIENCES

Author contributions: O.S., M.R., and K.S.C. designed research; O.S. and M.R. performed research; O.S. and M.R. analyzed data; and O.S., M.R., and K.S.C. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: schilke@ucla.edu or kcook@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1314857110/-/DCSupplemental.

In addition to the behavioral effect, we tested the hypothesis that blood-oxygen-level-dependent (BOLD) responses in two specific brain systems—the "C-system" and the "X-system" (20)—are differentially engaged depending on the relationship experience before the trust breach. Building on social cognitive neuroscience theory, we first hypothesized that trust recovery after early trust breaches is associated with increased activation of structures of the C-system, which previous research describes as a controlled social cognition system (20, 21). Because we argue that processing a trust breach when little relationship experience exists requires more conscious learning, more complex planning, and increased problem solving, we expect that increased activation of structures of the C-system will be associated with less prior relationship experience.

We further expect that a later trust breach, occurring after more relationship experience exists, is associated with activation of structures of the X-system, which prior research has described as an automatic social cognition system (20). In particular, the X-system has been argued to automatically generate the habits and impulses that guide people's daily activities (22, 23). Because we argue that relationships become habitualized as they mature and that such habitualization will then guide decision making after trust breaches, we expect that increased activation of structures of the X-system will be associated with more prior relationship experience.

Study 1: Prior Experience and Recovery from Trust Breach

Data showed that, overall, participants decided to trust more often than not to trust (66% vs. 34%; $t_{99} = 11.17$, $P < 0.001$). In line with previous research (e.g., ref. 24), a survey measure of general trust was significantly correlated with trust decisions during the experiment ($r = 0.22$, $P < 0.05$). We considered sex differences in trust recovery and found no significant effect ($t_{99} = 0.09$, $P > 0.1$).

As expected, we found a main effect of relationship experience: the proportion of postbreach money transfers was significantly greater in the high relationship-experience condition ($M_{\text{trust proportion}} = 0.79$, $SE_{\text{trust proportion}} = 0.04$) compared with the low relationship-experience condition ($M_{\text{trust proportion}} = 0.63$, $SE_{\text{trust proportion}} = 0.04$, $t_{99} = 2.66$, $P < 0.01$). We also conducted a more fine-grained analysis of trust rates in the individual posttrust breach trials to explore the trust recovery dynamics in greater detail. Fig. S1 shows that trust rates first drop more sharply but then recover to higher levels in the high (versus low) relationship-experience condition.

To establish the robustness of the effect of the manipulation while accounting for the level of prebreach trust, we conducted a regression with "trust recovery" as the dependent variable and "relationship experience" as the independent variable, controlling for the proportion of money transfers in the four trials before the trust breach. While the effect of prebreach money transfers was significant ($b = 0.28$, $t = 2.37$, $P < 0.05$), the effect of relationship experience remained significant as well ($b = 0.12$, $t = 1.99$, $P < 0.05$), suggesting that its influence is robust to preexisting trust levels.

Study 2: Neurophysiological Process

Behavioral Results of Study 2. We replicated our behavioral finding in a dataset obtained in a controlled laboratory setting: results showed significant differences in trust recovery between the early and late trust breach conditions ($t_{20} = 2.88$, $P < 0.01$); participants' posttrust breach behavior was significantly more trusting if the breach occurred late ($M_{\text{trust proportion}} = 0.77$, $SE_{\text{trust proportion}} = 0.05$) rather than early ($M_{\text{trust proportion}} = 0.58$, $SE_{\text{trust proportion}} = 0.07$) in a series of trust decisions. In addition, reaction times differed significantly between conditions ($t_{20} = 3.56$, $P < 0.01$), with reaction times being shorter in the late ($M_{\text{reaction time}} = 670.92$ ms, $SE_{\text{reaction time}} = 47.40$ ms) than in the early trust breach condition ($M_{\text{reaction time}} = 931.05$ ms, $SE_{\text{reaction time}} = 88.95$ ms), providing initial support for our proposition that trust recovery is more habitualized when the trust breach occurs later rather than earlier.

Neuroimaging Results of Study 2. We contrasted two specific trial phases of the behavioral task. In line with previous trust research (10, 14) and investigations in decision neuroscience (25–27), we focused on the four trials right after the trust breach occurred, specifically analyzing the phase during which the specific partner was announced and choices were entered (i.e., the "partner announcement" phase shown in Fig. 1*B*). In particular, we contrasted brain activation for the four relevant partner announcement

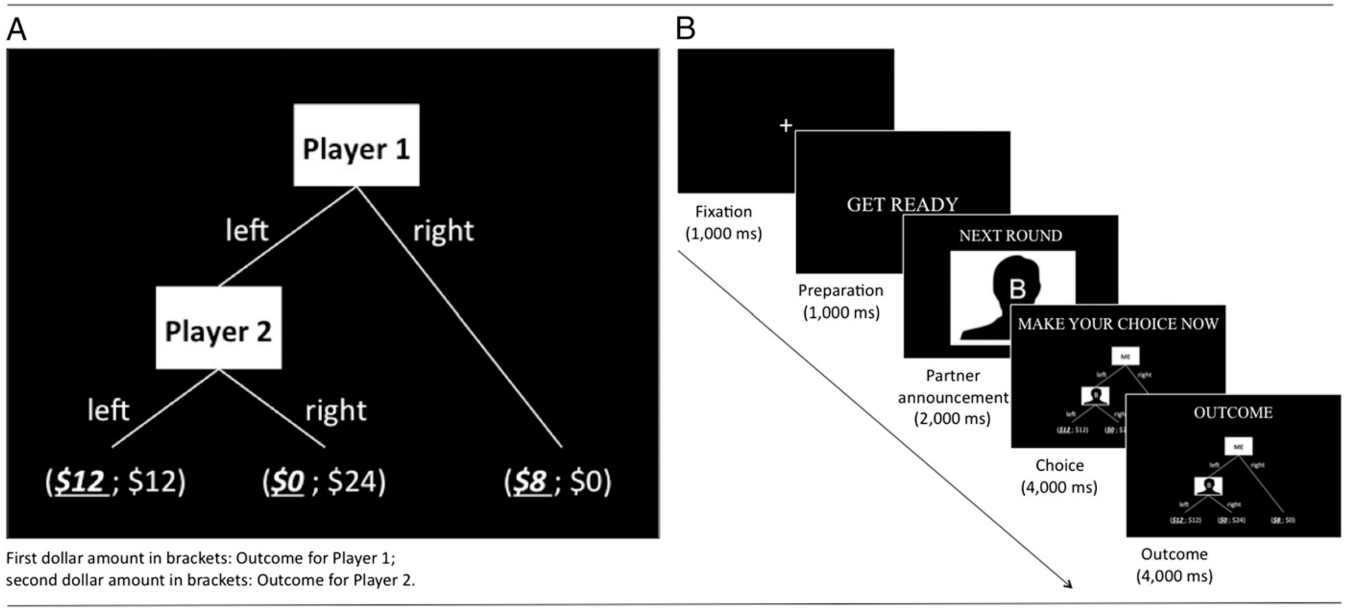

Fig. 1. Studies 1 and 2: Outcome structure (*A*) and trial phase of the behavioral task (*B*). The payment structure and trial structure was adapted from extant research (10).

phases after the early trust breach occurred to brain activation for the four relevant partner announcement phases after the late trust breach occurred.

Differences in brain activation following early compared with late trust breaches. Results of the whole-brain analysis revealed significant differences in the BOLD responses within 18 peak activation voxels and their surrounding voxel clusters for announcing partner B (the early trust-breaching partner) compared with partner C (the late trust-breaching partner) (see Tables S1 and S2 for more details). Of particular interest is the fact that our results revealed increased activation within subareas of the C-system, specifically the anterior cingulate cortex and lateral frontal cortex (Fig. 2 *A* and *B*). Moreover, decreased activation was found in a subarea of the X-system, specifically the lateral temporal cortex (Fig. 2*C*).

Region of interest analyses. Furthermore, a more stringent region of interest (ROI) analysis, which only considers the BOLD response data within a predefined voxel cluster and not all voxels of the

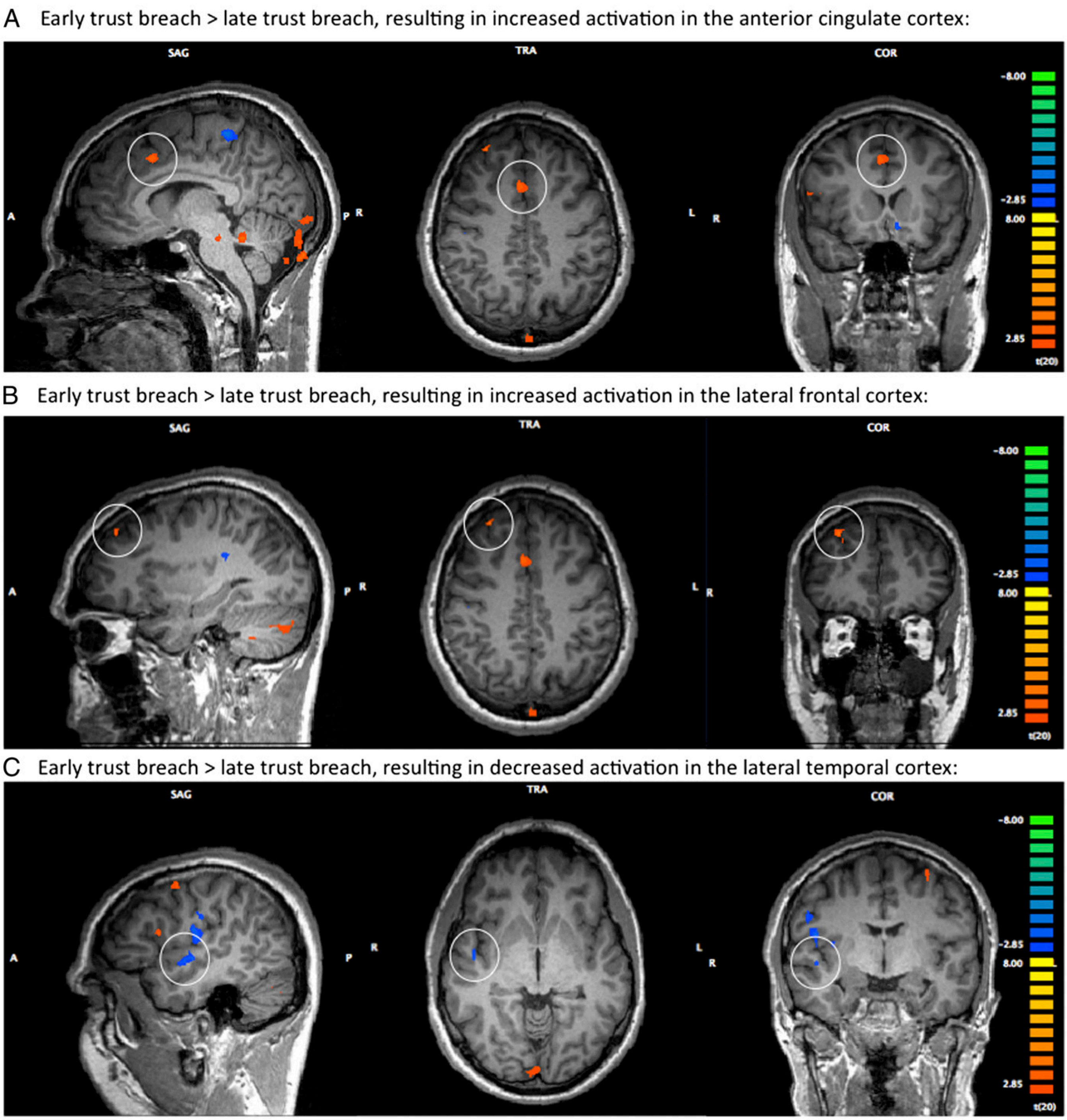

Fig. 2. (*A*–*C*) Study 2: BOLD responses for early trust breaches compared with late trust breaches ($P < 0.01$).

whole brain, confirmed our whole-brain analysis results (Table S1). In particular, results of nine separate ROI analyses again revealed significant differences in BOLD responses when comparing early to late trust breaches at significance levels of $P < 0.01$.

Discussion

This investigation provides convergent behavioral evidence that greater relationship experience before a trust breach supports trust recovery. Two experiments were conducted and reported here, which offer support for our account. It was shown that this behavioral effect is not only observable in data obtained from a controlled laboratory setting with a student sample (study 2), but also in data collected online via the Internet (uncontrolled) from a general population sample (study 1).

The results of our neuroimaging experiment (study 2) support our proposition that brain activation in a controlled social cognition system (C-system)—specifically, the anterior cingulate cortex and lateral frontal cortex—are associated with decision making after early trust breaches, whereas BOLD responses in an automatic social cognition system (or X-system)—specifically, the lateral temporal cortex—are associated with decision making after later trust breaches. In particular, the increased activations in the anterior cingulate cortex and lateral frontal cortex provide support for the notion that participants whose trust was breached early engaged in early learning, more complex planning, and more problem solving. Prior research has accumulated proof that the anterior cingulate cortex plays a vital role in these psychological functions (cf. ref. 28). Furthermore, previous research has shown that the lateral frontal cortex, specifically Brodmann area 8, is engaged when participants experience uncertainty (29), suggesting that participants in the low experience condition were arguably more uncertain about whether they should trust their partner.

Like most cognitive neuroscience research, the present fMRI study relies on reverse inference in that activation of a particular brain area (e.g., prefrontal cortex or the anterior cingulate cortex) is interpreted as support for engagement of a particular psychological process (e.g., increased planning or problem solving). In dealing with this controversial issue, we closely followed pertinent recommendations (30). Rather than pursuing an exploratory approach that would rely on post hoc explanation of a particular result, our fMRI study was set up to test more specific predictions pertaining to activation in two specific brain systems: that is, the C-system and the X-system. While these two systems incorporate several different structures, in the present research, we were able to provide neuroimaging evidence for how these systems are differentially processing controlled social cognition (i.e., less relationship experience and less trust recovery) versus automatic social cognition (i.e., more relationship experience and more trust recovery). Of course, we do not claim that recovery from a trust breach engages all structures that prior research has implicated in controlled versus automatic social cognition (cf. ref. 20 for a list of all brain areas implied). One reason why we did not identify all of these brain areas may be because of the nature of the behavioral task used in this research. Another reason could be that trust recovery employs this unique network of brain areas, which differs from other social cognition phenomena. Only future research will ultimately answer these questions. For example, functional connectivity studies could provide additional evidence concerning the functional relationship underlying social interactions, for which investigations have been started only recently (31).

The present research makes contributions on several fronts. First, this research contributes to social psychological theory on trust. Although trust is acknowledged to be critical for many social relations, scholars have only recently begun to focus on the recovery of trust following a breach. Our research adds to this emerging stream of work by focusing specifically on a victim's willingness to reconcile following a trust violation in situations varying in relationship experience before the violation. Our account emphasizes that trust decisions—even after trust breaches—become habitualized, more automatic, and less reflexive over time. Overall, this research contributes to understanding how a key relationship characteristic affects the recovery of trust, and the processes through which that occurs.

Second, we make a contribution to the neurophysiology and neurobiology of trust, which has been extensively studied over the past 10 y (6, 10, 14, 15, 17, 32). We add to this line of research by highlighting the fact that breaching or violating trust is associated with activation of brain areas that prior research has implicated in other aspects of trust. For example, McCabe et al. have associated regions of the frontal cortex (more medial and prefrontal than found in our research) in a cooperative game (14). Furthermore, Krueger et al. have identified the paracingulate cortex (an area adjacent to the anterior cingulate cortex, which we identified here) as well as areas of the brainstem in trusting relationships (10). Specifically, these investigators have linked their brainstem finding—particularly, the ventral tegmental area—to the release of dopamine, which in turn generates a "rewarding" experience through its projections throughout the mesolimbic dopamine system. In our research, although not directly providing evidence for an implication of the ventral tegmental area (which is an extremely small area), we provide similar evidence by showing engagement of the midbrain.

Finally, this research adds to the emerging field of organizational neuroscience, which has only recently started to contrast a more controlled, planning-based decision system with a more automatic habit-based decision system at the brain level, and speculates that the habit system arguably biases decision makers to persist engaging in past behaviors (33). The present research contributes to this newly forming domain of inquiry by providing initial empirical evidence for differentiation of these two systems and its consequences for trust-related behaviors.

Materials and Methods

Study 1. One-hundred adults (53% female; $M_{age} = 45.52$, $SE_{age} = 1.33$, ranging from 19 to 74 y) were recruited from a general-population participant panel of a commercial Web survey provider, completed an online consent form approved by Universität Magdeburg, answered a survey, engaged in a behavioral task shown on their computer screens, and received reimbursement through the Web survey provider.

Before engaging in the behavioral task, participants responded to a survey, which asked them to report various demographics and respond to an eight-item survey measure for general trust (34), that is, trust in others in general. Subsequently, the experimental task was presented in Java on each participant's computer screen. On several different sequentially presented slides, participants first received the following instructions:

> WELCOME! Thank you very much for volunteering to participate in the DECISION-MAKING GAME! In this game, you will play several rounds with a partner, who we will call "B." To avoid introducing potential biases, you will neither get to know this partner personally nor see pictures of him/her but only interact with the partner over the Internet. In each round, you receive \$8 and decide whether to invest this money with your partner or whether to keep it. If you invest, the money is tripled. However, the partner decides whether to share the \$24 with you (so that each of you receives \$12) or whether to keep the entire amount (so that you end up with nothing). If you invest, your return is thus determined by your partner. An overview of the game structure and pay-offs is provided on the next slide.

Participants then saw the outcome structure in a picture, which is depicted in Fig. 1*A*.

> Prior research has demonstrated that it makes a lot of sense for the players to send money to their partners in order to maximize their pay-offs over all rounds. Each of the decision rounds is structured as follows: (*i*) Both you and your partner simultaneously make a choice ("send" or "keep" money) within a time window of no more than 4 seconds. Please make sure to enter your choice within this time window! If you don't, you will receive \$0 for the round independent of

your partner's choice. (*ii*) You will see both your's and your partner's choices and the resulting outcome. Sometimes, you will not play with your partner but instead you will have to chance to play with a slot machine. You will be asked whether or not you would like to put $8 in a slot machine. Subsequently, you will see the outcome of the lottery. At the end of the session today, the computer will randomly select one of the rounds (either from the decision-making game or the lottery) and you will be paid according to the outcome of that round. As such, your choices during this task have real economic consequences for you. You can choose between the two available alternatives by either pressing "1" on your keyboard for the left alternative or pressing "2" on your keyboard for the right alternative. Do you understand these instructions and are you ready to start? (Once you have responded, you cannot pause the decision-making task until it is completed.) YES (1) or NO (2). PLEASE WAIT A FEW MOMENTS FOR US TO SYNCHRONIZE YOU WITH YOUR PARTNER!

After the presentation of the instructions, the behavioral task started automatically. Participants first saw a white cross to center their attention for 1 s (intertrial phase), were then prompted to "Get ready" for 1 s (preparation phase), then told whom they would play with on this trial (i.e., either player B or a slot machine) for 2 s (partner announcement phase), then shown the decision tree and asked to "Make your choice now" for 4 s (choice phase), and then presented with the outcome for 4 s (outcome phase). Note that in the variant of the trust game used here, participants learned about their partner's decision regardless of their own choice (8, 35). That is, even when participants chose not to transfer their $8 endowment to their partner, they would still learn whether their partner would have returned $12. Thus, the information about partner trustworthiness was independent of participants' own choices. After completing the experimental task, participants were directed to a posttask survey and were finally debriefed. Given that their partner was not real (which was revealed only after finishing the task), they were paid the highest possible outcome of $12.

In this between-subjects experimental design, participants engaged in a total of 24 trials either with an early- or a late-defecting partner. A trust breach was operationalized in terms of the partner not sharing in two consecutive trials (36, 37). Relationship experience was manipulated by varying the number of trials before that breach. While the trust breach occurred during trials 5 and 6 in the low relationship-experience condition, the trust breach only occurred during trials 15 and 16 in the high relationship-experience condition (38). The experimental software randomly assigned participants to one of the two conditions: 45 participants to the low relationship condition and 55 participants to the high relationship condition.

The dependent variable of trust recovery was operationally defined as the proportion of money-transfer decisions in the four trust-game trials immediately following the trust breach (36).

The 24 trials of the trust game were intermixed with an additional six trials of slot-machine gambles (10, 39, 40). When gambling on the slot machine, participants were offered two choices: they could either gamble with a pool of $8 or they could keep $8. Participants were informed they would engage in several trials but, to avoid end-game effects, they were unaware of the exact number of trials. Trials were pseudorandomized and participants responded on their computer keyboards. Fig. 1*B* illustrates a trial sequence of the behavioral task.

Study 2. Twenty-two adults (71% female; $M_{age} = 19.71$, $SE_{age} = 0.39$, ranging from 18 to 26 y) gave written informed consent to participate in this fMRI study, were checked for medical eligibility, engaged in the behavioral task while undergoing fMRI, and received a $20 show-up fee. The data file of one participant was corrupt and was not usable in the subsequent analyses, resulting in a final sample size of $n = 21$.

Before the scan session, participants performed a short training task to alleviate task-related confusion during scanning. Subsequently, participants were placed supine and task stimuli were presented in E-Prime 2.0 (Psychology Software Tools) and projected into the scanner, and participants could see the stimuli in a mirror located directly in front of their eyes.

Instructions and stimuli were analogous to those reported in study 1, with the exception that study 2 used a within-subject design. For this purpose, participants were playing with two different partners that were intermixed throughout subsequent trials: player B (low relationship-experience condition, trust breach in fifth and sixth partner-specific trial) and player C (high relationship-experience condition, trust breach in 15th and 16th partner-specific trial). Twenty-four trials $\times$ 2 conditions equaled 48 trials. During the other 12 trials, participants again gambled on a slot machine. Trials were pseudorandomized and participants responded using a response box they held in both hands. The experiment was incentive compatible by having participants expect actual payments for one randomly selected trial (in reality, each participant eventually received the maximum outcome of $12 in addition to the $20 show-up fee).

ACKNOWLEDGMENTS. We thank Bernd Weber for constructive feedback on an earlier version of the manuscript. We also thank the reviewers for their comments, which helped us improve the manuscript. This research was performed while M.R. was affiliated with Universität Magdeburg.

1. Cook KS, Hardin R, Levi M (2005) *Cooperation Without Trust?* (Russell Sage Foundation, New York, NY).
2. Cook KS, Schilke O (2010) The role of public, relational and organizational trust in economic affairs. *Corp Reputation Rev* 13(2):98–109.
3. Hardin R (2002) *Trust and Trustworthiness* (Russell Sage Foundation, New York, NY), pp xxi, 234 pp.
4. Rempel JK, Holmes JG, Zanna MP (1985) Trust in close relationships. *J Pers Soc Psychol* 49(1):95–112.
5. Slovic P (1993) Perceived risk, trust, and democracy. *Risk Anal* 13(6):675–682.
6. King-Casas B, et al. (2008) The rupture and repair of cooperation in borderline personality disorder. *Science* 321(5890):806–810.
7. Tomlinson EC, Dineen BR, Lewicki RJ (2004) The road to reconciliation: Antecedents of victim willingness to reconcile following a broken promise. *J Manage* 30(2):165–187.
8. Haselhuhn MP, Schweitzer ME, Wood AM (2010) How implicit beliefs influence trust recovery. *Psychol Sci* 21(5):645–648.
9. Krackhardt D (1992) The strength of strong ties: the importance of philos in organizations. *Networks and Organizations: Structure, Form, and Action*, eds Nohria N, Eccles RG (Harvard Business School, Boston, MA), pp 216–239.
10. Krueger F, et al. (2007) Neural correlates of trust. *Proc Natl Acad Sci USA* 104(50): 20084–20089.
11. Kramer RM (1996) Divergent realities and convergent disappointments in the hierarchic relation: trust and the intuitive auditor at work. *Trust in Organizations: Frontiers of Theory and Research*, eds Kramer RM, Tyler TR (Sage, Thousand Oaks, CA), pp 216–245.
12. Liu JH, et al. (2011) Unbalanced triangle in the social dilemma of trust: Internet studies of real-time, real money social exchange between China, Japan, and Taiwan. *Asian J Soc Psychol* 14(4):246–257.
13. Riedl R, Hubert M, Kenning P (2010) Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *Manage Inf Syst Q* 34(2):397–428.
14. McCabe K, Houser D, Ryan L, Smith V, Trouard T (2001) A functional imaging study of cooperation in two-person reciprocal exchange. *Proc Natl Acad Sci USA* 98(20): 11832–11835.
15. King-Casas B, et al. (2005) Getting to know you: Reputation and trust in a two-person economic exchange. *Science* 308(5718):78–83.
16. Takagishi H, et al. (2009) Neural correlates of the rejection of unfair offers in the impunity game. *Neuroendocrinol Lett* 30(4):496–500.
17. Adolphs R (2002) Trust in the brain. *Nat Neurosci* 5(3):192–193.
18. Cook KS, Cooper RM (2003) Experimental studies of cooperation, trust and social exchange. *Trust, Reciprocity and Gains from Association: Interdisciplinary Lessons from Experimental Research*, eds Ostrom E, Walker J (Russell Sage Foundation, New York, NY), pp 277–333.
19. Cook KS, Yamagishi T (2008) A defense of deception on scientific grounds. *Soc Psychol Q* 71(3):215–221.
20. Lieberman MD (2007) Social cognitive neuroscience: A review of core processes. *Annu Rev Psychol* 58:259–289.
21. Satpute AB, Lieberman MD (2006) Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Res* 1079(1):86–97.
22. Lieberman MD, Jarcho JM, Satpute AB (2004) Evidence-based and intuition-based self-knowledge: An fMRI study. *J Pers Soc Psychol* 87(4):421–435.
23. Lieberman MD, Eisenberger NI (2004) Conflict and habit: A social cognitive neuroscience approach to the self. *On Building, Defending and Regulating the Self: A Psychological Perspective*, eds Tesser A, Wood JV, Stapel DA (Psychology Press, New York, NY), pp 77–102.
24. Yamagishi T, Cook KS, Watabe M (1998) Uncertainty, trust, and commitment formation in the United States and Japan. *Am J Sociol* 104(1):165–194.
25. Bechara A, Damasio H, Tranel D, Damasio AR (1997) Deciding advantageously before knowing the advantageous strategy. *Science* 275(5304):1293–1295.
26. Ernst M, Paulus MP (2005) Neurobiology of decision making: A selective review from a neurocognitive and clinical perspective. *Biol Psychiatry* 58(8):597–604.
27. Knutson B, Greer SM (2008) Anticipatory affect: Neural correlates and consequences for choice. *Philos Trans R Soc Lond B Biol Sci* 363(1511):3771–3786.
28. Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P (2001) The anterior cingulate cortex. The evolution of an interface between emotion and cognition. *Ann N Y Acad Sci* 935(1):107–117.
29. Volz KG, Schubotz RI, von Cramon DY (2005) Variants of uncertainty in decision-making and their neural correlates. *Brain Res Bull* 67(5):403–412.

30. Poldrack RA (2006) Can cognitive processes be inferred from neuroimaging data? *Trends Cogn Sci* 10(2):59–63.
31. Rilling JK, et al. (2008) The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* 46(5):1256–1266.
32. Fouragnan E, et al. (2013) Reputational priors magnify striatal responses to violations of trust. *J Neurosci* 33(8):3602–3611.
33. Becker WJ, Cropanzano R (2010) Organizational neuroscience: The promise and prospects of an emerging discipline. *J Organ Behav* 31(7):1055–1059.
34. Mayer RC, Davis JH (1999) The effect of the performance appraisal system on trust for management: A field quasi-experiment. *J Appl Psychol* 84(1):123–136.
35. van den Bos W, van Dijk E, Crone EA (2012) Learning whom to trust in repeated social interactions: A developmental perspective. *Group Process Intergroup Relat* 15(2):243–256.
36. Schweitzer ME, Hershey JC, Bradlow ET (2006) Promises and lies: Restoring violated trust. *Organ Behav Hum Decis Process* 101(1):1–19.
37. Dirks KT, Kim PH, Ferrin DL, Cooper CD (2011) Understanding the effects of substantive responses on trust following a transgression. *Organ Behav Hum Decis Process* 114(2):87–103.
38. Bottom WP, Gibson K, Daniels SE, Murnighan JK (2002) When talk is not cheap: Substantive penance and expressions of intent in rebuilding cooperation. *Organ Sci* 13(5):497–513.
39. Delgado MR, Frank RH, Phelps EA (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat Neurosci* 8(11):1611–1618.
40. Chang LJ, Doll BB, van 't Wout M, Frank MJ, Sanfey AG (2010) Seeing is believing: Trustworthiness as a dynamic belief. *Cognit Psychol* 61(2):87–105.

Supporting Information

Schilke et al. 10.1073/pnas.1314857110

SI Materials and Methods

Study 1. To explore the dynamics of trust recovery in greater detail, we analyzed trust rates in each of the four trials following the trust breach individually. As shown in Fig. S1, trust rates are at low levels immediately following the trust breach but recover thereafter and to a greater extent in the high than in the low relationship-experience condition.

Study 2. ***Functional MRI data acquisition.*** Functional MRI (fMRI) was conducted using a full-body 3.0 Tesla Siemens scanner, which is fitted with a 12-channel matrix head coil. For structural imaging, a high-resolution image of the brain was acquired with a 3D, T1-weighted sequence [echo time (TE)/repetition time (TR)/inversion time = 3.1/2,530/800 ms, flip angle = 10°, matrix = 256 × 256, field of view (FOV) = 256 mm, slice thickness = 1 mm without gap]. For functional imaging, a time series of 464 volumes, with 41 slices in the transverse plane, was obtained using single shot gradient-echo planar imaging (TR = 2,000 ms, TE = 25 ms, flip angle = 90°, resolution = 3.0 mm × 3.0 mm × 2.5 mm, and FOV = 192 mm).

fMRI data analyses. Imaging data were preprocessed and analyzed using the BrainVoyager QX 2.20 software (1). Changes in the blood oxygen level-dependent (BOLD) responses were assessed for each voxel using the volume map (i.e., the map of brain function over the course of the experiment), which involved a multistep analyses procedure: First, we created both an anatomical data file (i.e., volumetric magnetic resonance, or VMR, file) and a functional data file (i.e., functional magnetic resonance, or FMR, file) for each participant. As per recommendation by the BrainVoyager software, we skipped the first two volumes of the FMR file because the first volume scans in any functional run may contain data with very high intensity values as a result of T1 saturation. This analysis step resulted in a time series of 462 volumes, contained in each participant's FMR file, which we used for the subsequent analyses.

Second, for each participant the FMR file was preprocessed, applying standard slice-scan time correction, standard 3D motion correction, standard spatial smoothing with a 3D Gaussian kernel, 4-mm full-width at half-maximum, and standard temporal high-pass filtering. Specifically, the 3D motion correction was performed by spatial transformation of all volumes of a participant to the (first) volume by rigid body transformations. Estimated translation and rotation parameters never exceeded 3 mm or 2°. In addition to linear trend removal, low-frequency nonlinear trends of three or fewer cycles per time course were removed by temporal high-pass filtering. These preprocessing steps are standard procedures in the BrainVoyager software (1), except spatial smoothing, which is not a default preprocessing step.

Third, the VMR and FMR files were coregistered, followed by both an anterior commissure–posterior commissure (AC–PC) transformation and an AC–PC to Talairach (2) transformation. These transformation steps again represent standard transformation procedures in the BrainVoyager software (1).

Fourth, on the basis of the previously generated files, volume time course (VTC) files were created for each participant.

Fifth, a time-course protocol was created, defining each volume of the task. Thirteen unique predictors were created, representing the different trial phases of the behavioral task: (*i*) the four relevant partner announcement phases right before playing with partner B after the trust breach ("Present_PartnerB_PostTb_4rele"); (*ii*) the four relevant presentation phases right before playing with partner C after the trust breach ("Present_PartnerC_PostTb_4rele"); (*iii*) all outcome phases right after playing with partner B during the trust breach ("Outcome_PartnerB_Tb"); (*iv*) all outcome phases right after playing with partner C during the trust breach ("Outcome_PartnerC_Tb"); (*v*), the three relevant presentation phases before playing on the slot machine ("Present_SlotM_74-99-124"); (*vi*) the three relevant presentation phases before playing on the slot machine ("Present_SlotM_199-224-249"); (*vii*) the two relevant losing outcome phases after playing on the slot machine ("Outcome_SlotM_Loss_51-101"); (*viii*) the two relevant losing outcome phases after playing on the slot machine ("Outcome_SlotM_Loss_176-201"); (*ix*) all present, choice, and outcome phases before the trust breach ("PreTb"); (*x*) all other present and choice phases during the trust breach ("Others_Tb"); (*xi*) all other presentation and choices phases after the trust breach ("Others_PostTb"); (*xii*) all other presentation, choice, and outcome phases for playing on the slot machine ("Others_SlotM"); and (*xiii*) all remaining task phases, including introduction, goodbye sequence, fixation phases, and get-ready phases ("Others_IntroGoodbyeFixationGetReady").

Sixth, a single-design matrix was created by convolving the onset of each predictor with a hemodynamic response function. This analysis step aimed at identifying voxels with blood flow that correlated with the predictors.

Seventh, a multisubject-design matrix was created, which included all 22 VTC files and the single-design matrix.

Eighth, data were submitted to a random-effects general linear model with percent-wise time course transformation applied to each run.

Ninth, in line with previous decision neuroscience research (3), we focused our analyses on the comparison of the four 2-s phases leading up to the different choices after the trust breach occurred; that is, the so called "partner announcement" phases right before participants played with either partner B or C by making a choice on their button box (i.e., predictors 1 and 2). We contrasted these two predictors—predictor 1 and predictor 2—against each other by subtracting the BOLD responses during the 2-s partner announcement phase for player C from those of player B. The goal was to identify differences in BOLD responses for partner B, which breached participants trust early during the behavioral task, compared with partner C, which breached participants trust late during the behavioral task. The statistical threshold was $P < 0.01$ and the default cluster threshold option of four continuous voxels was applied. All clusters of brain activation shown on the brain map (Fig. 1) were transformed into volumes-of-interest, and then Talairach coordinates for each peak activation voxel were read-out. These coordinates, in turn, served as an input for the automated Talairach client in which each single coordinate was searched (4). The outcome of this final-analysis step are summarized in Table S1, which lists the Talairach coordinates, information on the identified lobe, brain area, Brodmann area (5), t-value, and significance level.

Tenth, regions of interest (ROIs) were defined both on the basis of the Talairach client automatically (4) and a standard neuroanatomical atlas in which we identified each brain structure manually (6). Data were submitted to a ROI general linear model with separate subject predictors and a percent-wise time course transformation was applied to each run. The same multidesign matrix file was used as in the whole-brain analysis. Again, predictor 1 was contrasted against predictor 2 (i.e., same contrasts as used in the previous analysis: predictor 1 > predictor 2). ROI results confirmed the earlier whole-brain analysis results for all nine voxel clusters reported in Table S1 at $P < 0.01$.

1. Goebel R, Esposito F, Formisano E (2006) Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp* 27(5):392–401.
2. Talairach J, Tournoux P (1988) *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York, NY).
3. Ernst M, Paulus MP (2005) Neurobiology of decision making: A selective review from a neurocognitive and clinical perspective. *Biol Psychiatry* 58(8):597–604.
4. Lancaster JL, et al. (2000) Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp* 10(3):120–131.
5. Brodmann K (1909) *Vergleichende Lokalisationslehre der Grosshirnrinde: in ihren Principien dargestellt auf Grund des Zellenbaues [Comparative Study of the Localization Within the Cortex: Illustrated in its Principles on the Basis of Cell Structure]* (Johann Ambrosius Barth, Leipzig, Germany). German.
6. Haines DE (2008) *Neuroanatomy: An Atlas of Atructures, Sections, and Systems* (Lippincott Williams & Wilkins, Philadelphia, PA).

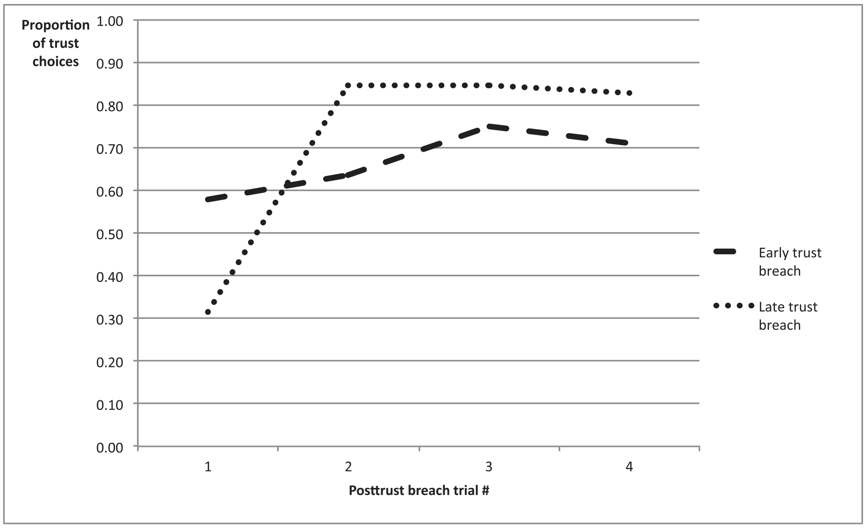

Fig. S1. Study 1: Mean proportion of trust choices in the four rounds following the trust breach.

Table S1. Study 2: Cerebral regions showing changes in BOLD responses during 2-s partner announcement phases after the trust breach occurred by partner B (early trust-breaching partner) compared with partner C (late trust-breaching partner)

Talairach coordinates								
x	*y*	*z*	Hemisphere	High-level brain region	Specific brain region	Brodmann area	*t*(20)	*P*
47	−8	0	Right	Temporal lobe	Superior temporal gyrus	22	−4.064273	0.000605
29	46	42	Right	Frontal lobe	Superior frontal gyrus	8	4.252369	0.000390
11	−68	18	Right	Occipital lobe	Precuneus	31	−3.627354	0.001678
11	10	66	Right	Frontal lobe	Superior frontal gyrus	6	3.897135	0.000895
5	19	42	Right	Limbic lobe	Cingulate gyrus	32	3.931459	0.000826
−1	−38	−33	Left	Brainstem	Pons	N/A	4.716577	0.000132
−10	−29	−18	Left	Brainstem	Midbrain	N/A	3.770300	0.001203
−28	−5	66	Left	Frontal lobe	Superior frontal gyrus	6	4.413431	0.000268
−40	−32	15	Left	Temporal lobe	Superior temporal gyrus	41	−3.683345	0.001473

Talairach coordinates and corresponding information for the peak activation voxels (searched by single point) are shown. The output is based on the automated Talairach client (4). Increased brain activation following early compared with late trust breaches: Results revealed increased activation within the cingulate gyrus at and around 5/19/42, $t(20) = 3.93$, $P < 0.01$, corresponding to the anterior cingulate cortex and Brodmann area 32. Furthermore, participants responded with increased brain activation within the superior frontal gyrus [at and around Talairach coordinates of the peak activation voxel of 29/46/42, $t(20) = 4.25$, $P < 0.001$, at and around −28/−5/66, $t(20) = 4.41$, $P < 0.001$, and at and around 11/10/66, $t(20) = 3.90$, $P < 0.01$]. These activation clusters correspond to the lateral frontal cortex, specifically Brodmann areas 8 and 6. Moreover, two regions of the brainstem revealed significant increases in brain activation for partner B compared with partner C, particularly within the midbrain [at and around −10/−29/−18, $t(20) = 3.77$, $P < 0.01$] and the pons [at and around −1/−38/−33, $t(20) = 4.72$, $P < 0.001$]. Decreased brain activation following early compared with late trust breaches: Results of the whole-brain analysis also showed decreased brain activation for announcing partner B compared with partner C with the superior temporal gyrus in both the right hemisphere [at and around 47/−8/0, $t(20) = -4.06$, $P < 0.01$] and the left hemisphere [at and around −40/−32/15, $t(20) = -3.68$, $P < 0.01$]. The precuneus [at and around 11/−68/18, $t(20) = -3.63$, $P < 0.01$] further revealed significantly decreased activation for announcing partner B compared with partner C after they breached participants trust.

Table S2. Study 2: Cerebellar regions showing changes in BOLD responses during 2-s partner announcement phases after the trust breach occurred by partner B (early trust-breaching partner) compared with partner C (late trust-breaching partner)

Talairach coordinates						
x	*y*	*z*	Hemisphere	Specific brain region	*t*(20)	*P*
50	−68	−24	Right	Tuber	3.643402	0.001617
38	−77	−27	Right	Tuber	4.206315	0.000434
35	−56	−33	Right	Cerebellar tonsil	4.186854	0.000454
26	−50	−21	Right	Culmen	4.748166	0.000123
11	−74	−30	Right	Pyramis	4.334925	0.000322
5	−44	−15	Right	Cerebellar lingual	4.228548	0.000412
−28	−68	−33	Left	Pyramis	4.259907	0.000383
−28	−80	−24	Left	Uvula	4.006740	0.000693
−43	−50	−36	Left	Cerebellar tonsil	4.397260	0.000278

Talairach coordinates and corresponding information for the peak activation voxels (searched by single point) are shown. The output is based on the automated Talairach client (4). These cerebellar results were omitted from further analyses because they were not considered ROI because the focus of our analysis was on the C-system and the X-system.