

Agent-Specific Responses in the Cingulate Cortex During Economic Exchanges

Damon Tomlin,^{1*} M. Amin Kayali,^{1*} Brooks King-Casas,¹ Cedric Anen,³ Colin F. Camerer,³ Steven R. Quartz,³ P. Read Montague^{1,2}

Interactions with other responsive agents lie at the core of all social exchange. During a social exchange with a partner, one fundamental variable that must be computed correctly is who gets credit for a shared outcome; this assignment is crucial for deciding on an optimal level of cooperation that avoids simple exploitation. We carried out an iterated, two-person economic exchange and made simultaneous hemodynamic measurements from each player's brain. These joint measurements revealed agent-specific responses in the social domain ("me" and "not me") arranged in a systematic spatial pattern along the cingulate cortex. This systematic response pattern did not depend on metrical aspects of the exchange, and it disappeared completely in the absence of a responding partner.

Social exchange occurs in species ranging from insects to humans (1–3). In primates, reciprocal interactions with nonkin occur repeatedly, thus necessitating the capacity to assign social credit or blame for shared outcomes and to act appropriately according to these assignments (4–6). In humans, reciprocity is a central feature of the collection of psychological mechanisms necessary to support social exchange (3); yet, the underlying neural representations of these mechanisms remain murky. In almost all social exchanges, one must detect and accurately track which social agent (who) gets credit for an outcome. Should credit for an outcome be assigned to one's own actions or those of one's partner? Perhaps such assignments are more a matter of degree—assigning the degree-of-credit to some shared outcome. Understanding such agent-specific mechanisms is important, because the assignment of social agency (7–13) appears to break down in a range of mental illnesses (14–16).

Social agency computations are also a prerequisite for generating models of others's mental states. This latter capacity, called theory-of-mind, is highly developed in humans and has been shown to activate a consistent set of brain regions in neuroimaging experiments (17–20). Recent work has complemented these theory-of-mind experiments by using interactive economic games as ecologically realistic models for human exchange (21–31). These experiments have elicited not only brain responses in previously described

theory-of-mind networks (27–29), but also have elicited formerly unreported activations along the cingulate cortex that correlate with the revelation of a social partner's decision (29). Although evoked during an economic exchange with another human, these cingulate activations did not modulate as a function of the fairness of the exchange, nor did they occur in exchanges with computer partners (28).

This lack of sensitivity to measures of outcome suggests that these responses do not encode some metrical aspect of the trade; instead, they are consistent with the social agency computation described above. To test this possibility directly, we scanned both

brains of two subjects interacting in a 10-round trust game (32, 33). This game has been used to identify neural correlates of reputation building and reciprocity (25). Because the trust game is a multiround economic exchange, it allows us to estimate brain responses in both subjects to multiple revelations of a partner's choice and multiple instances of one's own choice.

The trust game is illustrated in Fig. 1A. In each round, one player (investor) invests some amount I (investment phase) that is tripled and sent to the other player (trustee) who decides to repay some fraction f of the tripled amount (repayment phase) (33). Players maintain their roles for 10 consecutive rounds, and information about outcomes within each phase is presented simultaneously to both subjects. We parameterized social context within the trust game by implementing two separate versions of the task in different subject cohorts: personal ($n = 104$ subjects) and impersonal [$n = 96$; previously reported in (25)]. In the personal version, subjects met before the task, were instructed together, saw a picture of their partner during each round of the game, and met their partner afterward, where they were paid in front of each other. In the impersonal version, subjects never met, had no chance of subsequent encounter, and received no information about one another.

Given the previously reported activations in the anterior and posterior portions of the medial cingulate during a social exchange (28), a detailed analysis of the cingulate cortex in each pair of subjects was per-

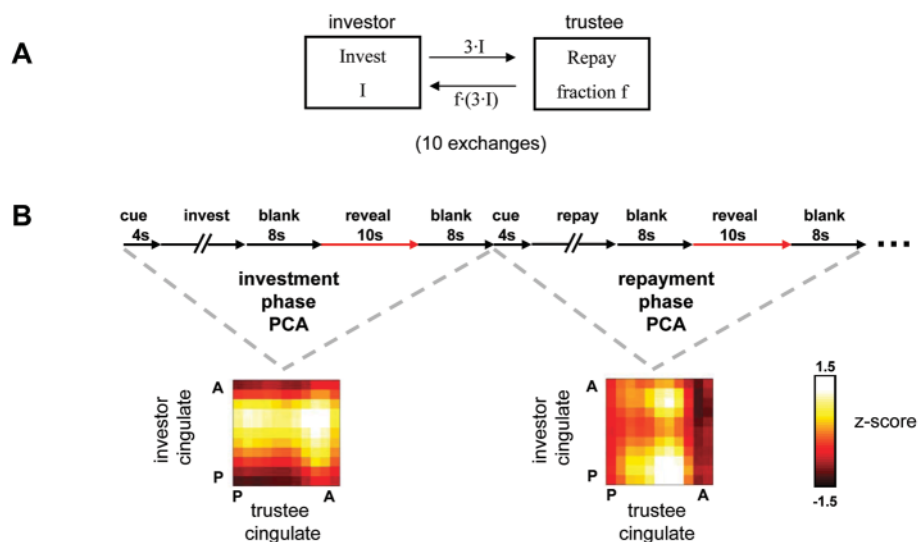


Fig. 1. Cross-cingulate correlations reveal complementary activity patterns across investment and repayment phase of game. **(A)** 10-round trust game. In each round the investor contributes some amount I between 0 monetary units and 20 monetary units, which is tripled ($3 \times I$) and sent to the trustee who then repays some fraction f of the tripled amount ($f \times 3 \times I$). **(B)** Cross-cingulate principal component analysis (PCA) revealed distinct, but complementary patterns when applied to the cross-correlations between cingulate cortices of investor and trustee (34). A similar result was found using independent component analysis on the same data (fig. S5).

¹Human Neuroimaging Laboratory, Department of Neuroscience, ²Department of Psychiatry, Center for Theoretical Neuroscience, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA. ³Social Cognitive Neuroscience Laboratory, California Institute of Technology, Division of Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, CA 91125, USA.

*These authors contributed equally to this work.

formed. We segmented the medial cingulate and the surrounding paracingulate cortex into separate spatial domains (33), computed cross-cingulate and cross-paracingulate correlation matrices for different lags in each phase of the task (investment phase and repayment phase), and carried out temporal principal component analysis (PCA) on the resulting three-dimensional correlation matrix (Fig. 1B) (34, 35). Analysis yielded complementary spatial patterns for cingulate cortices (Fig. 1B)—that is, patterns of activation in one phase were transposed across role when analysis was performed for the other phase. Similar results were found using independent component analysis (ICA) on the cross-cingulate correlation matrix (fig. S5) (33).

The cross-cingulate analysis led us to examine the hemodynamic time series in each cingulate segment. This region-of-interest analysis revealed three distinct response types (Fig. 2A). The first followed the submission of a subject's own decision (unimodal "own"-dominated response); the second followed the visual presentation of a partner's decision (unimodal "other"-dominated response). This is a remarkable finding, because visual presentation of the subject's own decision elicited little response in the cingulate cortex. The third response type was bimodal, yielding approximately equal responses after submission of one's own decision and revelation of the partner's decision. However, the peak amplitude of these distinct response types was not uniform across the anterior-posterior axis of the cingulate. Instead, they displayed a systematic spatial variation that was complementary across the basic response types ("own" and "other"). Specifically, the submission of one's own decision elicited maximal activation in middle cingulate regions (Fig. 2A, segment G), whereas viewing the revelation of a partner's decision yielded maximal activation in anterior and posterior cingulate (an example of an anterior response is shown in Fig. 2A, segment K). This result was in stark contrast to the results of the paracingulate analysis, which indicated that, although the dorsal anterior cingulate cortex was highly activated during the experiment, there was no spatial selectivity for either stimulus. In fact, the dorsal anterior cingulate cortex responded strongly to the submission of decisions and the revelation of partner choices, and it was the only paracingulate region significantly activated by either (fig. S6) (33).

The distinct response types and the systematic spatial variation of peak amplitudes across the anterior-posterior axis disappeared completely in motor control ($n = 15$; Fig. 2C) and sensory control experiments ($n = 17$; Fig. 2E) not involving exchange with another agent (33). In the motor control, subjects

reiterated the motor responses of randomly selected investors. We applied the same region-of-interest analysis to the control data (Fig. 2 and fig. S2). Statistical comparison of responses in each of the cingulate domains showed that responses differed significantly between the normal trust task and the control tasks (33). In particular, no significant re-

sponse was present in the middle cingulate (Fig. 2C), ruling out the possibility that middle cingulate activation in the trust game was the result of motor activity produced by button tapping. In the sensory control, partner reveal screens from the trust game were viewed passively by a separate cohort of subjects ($n = 17$). Because partner reveal screens

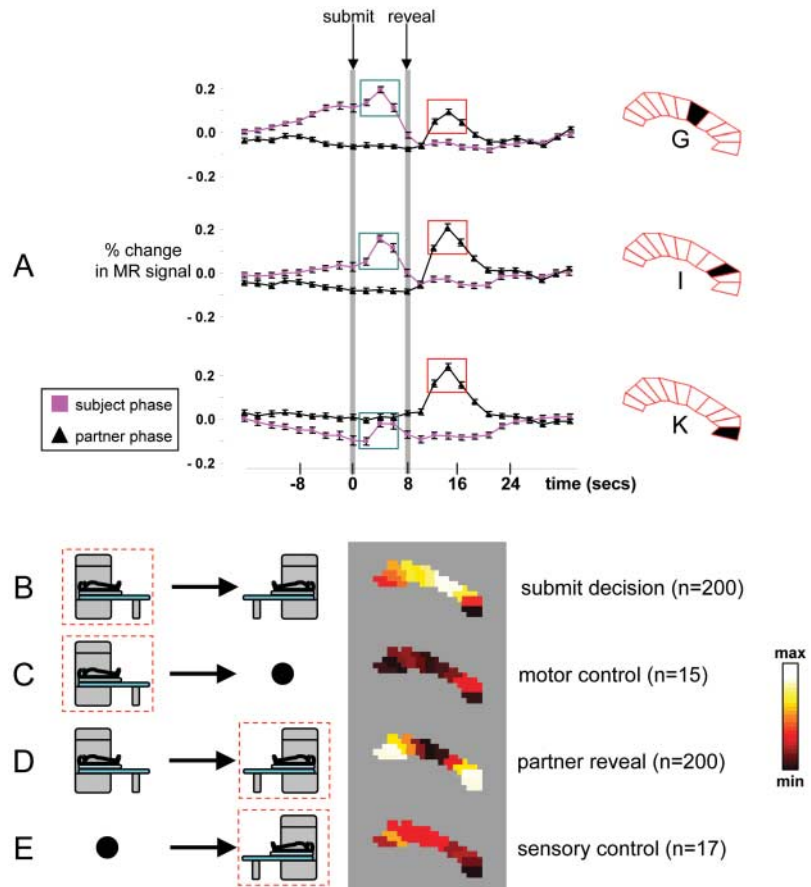
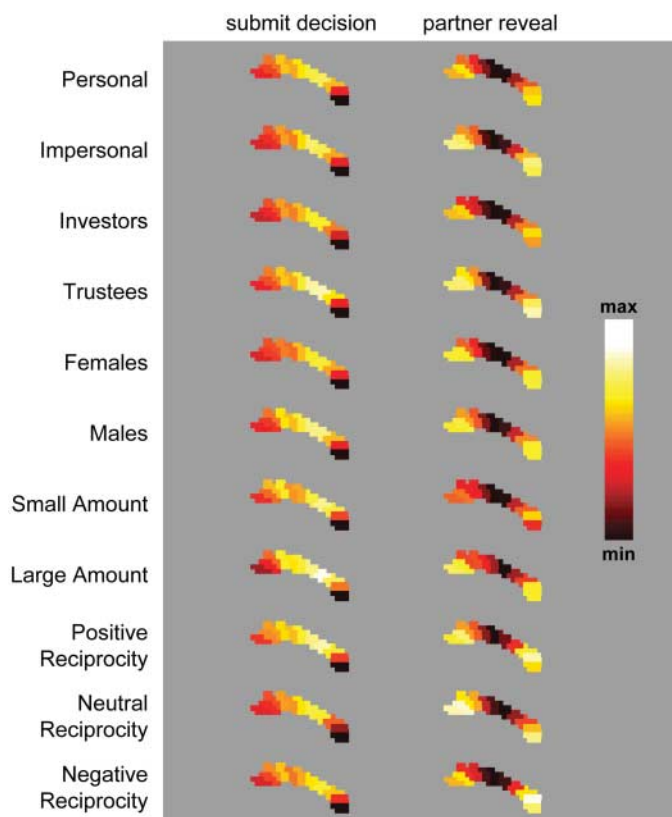


Fig. 2. Agent-specific responses and their pattern disappear outside of economic exchange. **(A)** Calculation of response pattern diagrams. Traces are the average magnetic resonance (MR) signal during subject decision phases (magenta lines) and during partner decision phases (black lines); error bars represent the standard error of the mean ($n = 200$ subjects). To compute the magnitude of responses to submitting a decision, MR values were selected from the time of peak response and the peak's two flanking points (teal boxes). These values, when averaged, represent the responsiveness of a segment to the submission of the decision. This measure was performed for all segments, and a pseudo color image was produced, as depicted in Figs. 2 and 3. For responses to partner reveal screens, MR values corresponding to the peak activity after screen onset and the peak's two flanking points (red boxes) were averaged and compiled into a similar pseudo color map. **(B)** The average response to submitting a decision is shown for subjects playing the linked trust game ($n = 200$), and a predominance of the middle cingulate is apparent. **(C)** Average response profile to submitting decisions in the unlinked motor control experiment ($n = 15$). No significant differentiation was observed across the cingulate of subjects in this task, but response levels in the middle cingulate were significantly different than those in the linked trust game ($P = 0.00001$). **(D)** Subjects from the linked trust experiment ($n = 200$) demonstrate the average response to viewing a social partner's decision. The predominance of responses in the anterior and posterior poles of the cingulate is apparent in this group. **(E)** Average response to viewing screens in the unlinked visual control experiment ($n = 17$). No significant differentiation was observed across cingulate domains, but responses in both anterior and posterior regions were significantly different than those in the linked trust game ($P < 0.01$). Maximum activation in (B) and (C) is 0.21% change in MR signal; maximum activation in (D) and (E) is 0.12%; minimum activation for each is 0.00%.

in the trust game had novel content and had been generated by an external agent, we could not use the original data set to separate responses to social or novel stimuli. Thus, subjects in the sensory control task were informed that their compensation depended on money shown under the “gave” label on the screen, but were not told about the social task from which this screen was derived (fig. S1). This manipulation was performed so that a screen’s content still held novel and valuable information, but was devoid of social interaction. In each of the 11 cingulate domains, BOLD responses after each of 10 outcome screens did not resemble those obtained during the analogous presentation in the linked experiment (Fig. 2E and fig. S3). There was no systematic spatial variation in response amplitudes across the cingulate gyrus. To verify that motor control and sensory control effects were not due to the smaller number of subjects involved in the control manipulations, we confirmed this finding by using random samples of the existing 200-subject database (fig. S3).

Fig. 3. Cingulate pattern of “me” and “not me” remains constant across a range of variables. All responses to decision submission are shown in the left column, whereas those responses to partner reveal screens are shown in the right column. With the exception of the reciprocity and amount diagrams, all responses were averaged across rounds before compilation. Rows labeled “Personal” and “Impersonal” separate activity across social context: the personal ($n = 104$ individuals) and impersonal ($n = 96$ individuals) tasks. Rows labeled “Investors” and “Trustees” demonstrate the consistency of the responses across the two different roles ($n = 100$ for each). Rows labeled “Males” and “Females” demonstrate that these responses do not differ across gender ($n = 100$ for each). Rows labeled “Small amount” and “Large amount” show that these patterns do not depend upon the amount of resource sent or received by the player (upper 25% versus lower 25% of payments; $n = 454$ and 161, respectively). Finally, the rows labeled “Positive,” “Neutral,” and “Negative” reciprocity depict responses across different valences of a behavioral variable of already known interest (25). These diagrams correspond to average BOLD responses to positive (values > 0.1 ; $n = 377$ choices), neutral ($-0.1 \leq \text{values} \leq 0.1$; $n = 865$ choices), and negative (values < -0.1 ; $n = 458$ choices) values of the reciprocity index. Left column maximum is 0.25% change in MR signal; right column maximum is 0.16%; minimum activation for both is 0.00%.



The results provide strong support for three new findings: (i) agent-specific response types localized on the medial bank of cingulate cortex, (ii) a systematic spatial variation of each response type across the anterior-posterior axis of cingulate cortex, and (iii) a dependence of both signals on the presence of a responding agent. Despite the relative simplicity of this economic exchange game, other variable(s) related to this task may have been the underlying cause of the different response types, the spatial variation across the cingulate, and the difference in response to visual revelation of one’s own decision and one’s partner’s decision. However, the different response types and their systematic but complementary spatial variation across the cingulate did not change as a function of a range of dimensions (Fig. 3). The most dramatic dimensions tested in Fig. 3 are reciprocity and social context (personal versus impersonal). In previously published work, reciprocity, expressed as degree of tit-for-tat behavior across rounds, acted as a powerful behavioral signal to one’s partner and elicited strong, measurable neural correlates (25). Yet,

as illustrated in Fig. 3 (bottom three rows), differences in reciprocity had no effect on the response types or on their spatial variation along the cingulate. The same result held for the difference in social context (personal, $n = 104$; impersonal, $n = 96$), where prior exposure to one’s partner, the sight of their picture in each round, and the knowledge of an imminent encounter afterward had no effect. Likewise, no differences were observed when comparing subject role (investor or trustee), sex of subject, or amount of money sent or received.

Using an iterated economic exchange task, we found two distinct response types along the cingulate cortex consistent with agent-specific responses that signal “me” and “not me.” Rather than residing in strictly demarcated functional zones, these complementary responses types exhibited smooth transitions across the entire medial bank of the cingulate gyrus. It is difficult to probe the extent to which a subject is considering outcomes for oneself or a social partner; individuals in a social exchange must necessarily model the actions of both agents as decisions are made and revealed. Despite this obstacle, the pattern of activation observed in these data was clearly sensitive to which participant was responsible for a given action. The response types and their variation through the tissue space disappeared in control experiments where money sent, actions taken, and money received were matched to those experienced during the normal multiround exchange (Fig. 2 and fig. S3) (35). These controls provide strong evidence that the response types were due to neither motor and premotor responses nor to sensory responses to outcome screens.

One question deserves separate consideration: Did the reveal screens generate simple surprise or novelty responses along cingulate that were not related to the social element of the exchange? Although this reasonable interpretation is possible, the control experiments suggest otherwise. The response pattern along the cingulate disappeared in the control experiments where subjects received stimuli that were visually identical to those in the trust game and were composed of novel, reward-related information. This manipulation used novel stimuli with economically meaningful content to probe the reveal response and showed neither an “other” response anywhere along the cingulate nor the spatial variation so prominent in the linked trust task. We take these data as strong support that the responses observed in the linked trust game were not the mere result of surprising content.

The response types and their spatial variation along the cingulate were remarkably stable to a range of manipulations. They survived the personal and impersonal treatments, did not change as a function of the reciprocity (a variable previously shown to be the major behavioral signal in this game) (25), and were not

changed as a function of sex, role, amount sent, or amount received. Lastly, these signals were clear even in individual subject pairs, as shown in fig. S4 for a single interacting dyad.

The observed lack of change as a function of reciprocity is extremely important because it reduces the likelihood of two alternate interpretations of these data. The average behavior in this game is initial cooperation followed by tit-for-tat moves, a strategy conjectured to be optimal in a reciprocal interaction (3, 6). To play such a tit-for-tat strategy, a player's brain must compute the expected next move of their partner and compare this to the actual outcome. Consequently, large deviations in reciprocity would also carry large prediction error signals, a signal type known to show up near or around dorsal anterior cingulate cortex (dACC) (36, 37). Two possibilities arise. The error signals could activate dACC because they reflect directly an error response. Alternately, large deviations in reciprocity represent a signal with a large amount of uncertainty and might engage an output conflict response typical for this brain region (36–44). However, neither of these interpretations would anticipate an important feature of the data actually observed. There was no difference in response types or their spatial variation as a function of positive, negative, or neutral reciprocity. One would at least expect both alternate explanations to show responses that differentiated neutral reciprocity from the other two categories (positive and negative). One possibility is that our current analysis missed the error signals altogether for some unidentified reason. However, by using this same behavioral task, we have previously identified such error-related signals elsewhere in the brain and have shown these regions to be sensitive to reciprocity (25). Consequently, our capacity to detect these error signals elsewhere makes it less likely that we simply missed error signals in cingulate related to strong deviations in reciprocity. However, it remains a possibility that some unprobed behavioral dimension generated an error signal along cingulate cortex during this task.

In a two-person social exchange, it is crucial for each agent to know how to credit an outcome. Failure to assign this credit accurately will compromise an agent's capacity to decide on an appropriate level of cooperation with the partner—a mistake that could prove extremely costly when averaged over multiple encounters (1–6). Consequently, we suspect that these data derive from a neural mechanism dedicated to distinguishing “me” outcomes from “not me” outcomes. The systematic spatial progression of responses suggests to us that this social agency variable may be arrayed as a map; however, the current experiment cannot adequately test this provocative possibility. It is important, therefore, to note that the assignment of credit (or agency) within a social interaction necessarily implicates a variety of cognitive and emotional

mechanisms. Thus, although agency parsimoniously characterizes the activations seen with these data, it may not necessarily be congruent to the underlying functions represented along the cingulate.

Extant data support a multifunctional role for the cingulate cortex, particularly in light of the extreme diversity of information that impinges on this region. Cingulate and paracingulate cortices have long been hypothesized as sites of integration of information sources that include cognitive, emotional, and interoceptive signals. Consequently, a range of functions has been ascribed to cingulate cortex (38–51), and there are disagreements over the exact variables processed and represented in these regions. However, it is reasonably clear that cingulate and paracingulate cortices contribute to normal social cognition and adaptive decision-making (17–19). The results of this paper add the important possibility that many other variables in the social domain may be arranged in such a systematic fashion through the spatial domain, a phenotype that could be disturbed in afflictions where the capacity to distinguish “me” from “not me” is impaired (14–16, 52–57).

References and Notes

- W. D. Hamilton, *J. Theor. Biol.* **7**, 1 (1964).
- W. D. Hamilton, *J. Theor. Biol.* **7**, 17 (1964).
- R. L. Trivers, *Q. Rev. Biol.* **46**, 35 (1971).
- J. Maynard Smith, G. R. Price, *Nature* **146**, 15 (1973).
- R. Axelrod, W. D. Hamilton, *Science* **211**, 1390 (1981).
- M. A. Nowak, K. Sigmund, *Nature* **355**, 250 (1992).
- U. Frith, C. D. Frith, *Curr. Dir. Psych. Sci.* **10**, 151 (2001).
- W. M. Kelley et al., *J. Cogn. Neurosci.* **14**, 785 (2002).
- M. D. Lieberman, J. H. Pfeifer, in *Cognitive Neuroscience of Emotional and Social Behavior*, A. Easton, N. Emery, Eds. (Psychology Press, Philadelphia, 2005), pp. 195–235.
- K. N. Ochsner et al., *J. Cogn. Neurosci.* **16**, 1746 (2004).
- C. A. Seger, M. Stone, J. P. Keenan, *Neuropsychologia* **42**, 1168 (2004).
- K. Vogeley et al., *Neuroimage* **14**, 170 (2001).
- K. Vogeley, G. R. Fink, *Trends Cogn. Sci.* **7**, 38 (2003).
- S. Baron-Cohen, M. K. Belmonte, *Annu. Rev. Neurosci.* **28**, 109 (2005).
- C. D. Frith, U. Frith, *Science* **286**, 1692 (1999).
- M. Brune, *Schizophr. Bull.* **31**, 21 (2005).
- H. L. Gallagher et al., *Neuropsychologia* **38**, 11 (2000).
- E. Brunet, Y. Sarfati, M. Hardy-Bayle, J. Decety, *Neuroimage* **11**, 157 (2000).
- B. Wicker, D. I. Perrett, S. Baron-Cohen, J. Decety, *Neuropsychologia* **41**, 139 (2003).
- J. Decety, P. L. Jackson, J. A. Sommerville, T. Chaminade, A. N. Meltzoff, *Neuroimage* **23**, 744 (2004).
- M. R. Delgado, R. H. Frank, E. A. Phelps, *Nat. Neurosci.* **8**, 1611 (2005).
- D. J. de Quervain et al., *Science* **305**, 1254 (2004).
- N. I. Eisenberger, M. D. Lieberman, K. D. Williams, *Science* **302**, 290 (2003).
- J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, J. D. Cohen, *Science* **293**, 2105 (2001).
- B. King-Casas et al., *Science* **308**, 78 (2005).
- K. McCabe, D. Houser, L. Ryan, V. Smith, T. Trouard, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11832 (2001).
- J. K. Rilling et al., *Neuron* **35**, 395 (2002).
- J. K. Rilling, A. G. Sanfey, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Neuroimage* **22**, 1694 (2004).
- A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Science* **300**, 1755 (2003).
- M. Bhatt, C. Camerer, *Games Ec. Behav.* **52**, 424 (2005).
- T. Singer et al., *Nature* **439**, 466 (2006).
- P. R. Montague et al., *Neuroimage* **16**, 1159 (2002).
- Materials and methods are available as supporting material on Science Online.
- A. Cichocki, S. Amari, *Adaptive Blind Signal and Image Processing* (Wiley & Sons Inc., New York, 2003).
- A. Hyvarinen, J. Karhunen, E. Oja, *Independent Component Analysis* (Wiley & Sons Inc., New York, 2001).
- C. B. Holroyd, M. G. H. Coles, *Psychol. Rev.* **109**, 679 (2002).
- C. B. Holroyd, S. Nieuwenhuis, N. Yeung, J. D. Cohen, *Neuroreport* **14**, 2481 (2003).
- C. S. Carter et al., *Science* **280**, 747 (1998).
- H. D. Critchley et al., *Brain* **126**, 2139 (2003).
- W. J. Gehring, R. T. Knight, *Nat. Neurosci.* **3**, 516 (2000).
- M. P. Milham, M. T. Banich, E. D. Claus, N. J. Cohen, *Neuroimage* **18**, 483 (2003).
- D. H. Weissman, B. Giesbrecht, A. W. Song, G. R. Mangun, M. G. Woldorff, *Neuroimage* **19**, 1361 (2003).
- M. Botvinick, L. E. Nystrom, K. Fissell, C. S. Carter, J. D. Cohen, *Nature* **402**, 179 (1999).
- C. S. Carter et al., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1944 (2000).
- A. R. Damasio et al., *Nat. Neurosci.* **3**, 1049 (2000).
- P. Rainville, G. H. Duncan, D. D. Price, B. Carrier, M. C. Bushnell, *Science* **277**, 968 (2000).
- K. N. Ochsner et al., *Neuropsychologia* **39**, 219 (2001).
- D. D. Dougherty et al., *Biol. Psychiatry* **46**, 466 (1999).
- K. L. Phan, I. Liberzon, R. C. Welsh, J. C. Britton, S. F. Taylor, *Neuropsychopharmacology* **28**, 1344 (2003).
- G. Bush, P. Luu, M. I. Posner, *Trends Cogn. Sci.* **4**, 215 (2000).
- F. A. Nielsen, D. Balslev, L. K. Hansen, *Neuroimage* **27**, 520 (2005).
- N. Georgieff, M. Jeannerod, *Conscious. Cogn.* **7**, 465 (1998).
- M. D. Lieberman, J. M. Jarcho, A. B. Satpute, *J. Pers. Soc. Psychol.* **87**, 421 (2004).
- K. N. Ochsner et al., *J. Cogn. Neurosci.* **16**, 1746 (2004).
- C. A. Seger, M. Stone, J. P. Keena, *Neuropsychologia* **42**, 1168 (2004).
- L. C. Johns et al., *Psychol. Med.* **31**, 705 (2001).
- J. M. Allman, K. K. Watson, N. A. Tetreault, A. Y. Hakeem, *Trends Cogn. Sci.* **9**, 367 (2005).
- This work was supported by the Center for Theoretical Neuroscience at Baylor College of Medicine (P.R.M.), National Institute on Drug Abuse (NIDA) grant DA11723 (P.R.M.), National Institute of Neurological Disorders and Stroke grant NS045790 (P.R.M.), National Institute of Mental Health grant MH52797 (P.R.M.), NIDA grant DA14883 (G. Berns), the Angel Williamson Imaging Center (P.R.M.), The Kane Family Foundation (P.R.M.), The David and Lucile Packard Foundation (S.R.Q. and C.F.C.), and The Gordon and Betty Moore Foundation (S.R.Q. and C.F.C.). P.R.M. also acknowledges support from the Institute for Advanced Study (Princeton, NJ) during completion of this work. We thank the Hyperscan Development Team at Baylor College of Medicine for Network Experiment Management Object (NEMO) software implementation (www.hnl.bcm.tmc.edu/nemo) and G. Berns for early discussions and efforts leading to the development of hyperscanning. We also thank C. Bracero, A. Harvey, S. Flaherty, J. McGee, K. Pfeiffer, R. Pruitt, and S. Gleason for technical assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1047/DC1
Materials and Methods
Figs. S1 to S6
References

31 January 2006; accepted 24 April 2006
10.1126/science.1125596