Neurocomputational mechanisms of prosocial learning and links to empathy

Patricia L. Lockwood[a,b,1], Matthew A. J. Apps[b], Vincent Valton[c], Essi Viding[a,2], and Jonathan P. Roiser[c,2]

[a]Division of Psychology and Language Sciences, University College London, London WC1H 6BT, United Kingdom; [b]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom; and [c]Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, United Kingdom

Edited by Ernst Fehr, University of Zurich, Zurich, Switzerland, and accepted by Editorial Board Member Susan T. Fiske June 26, 2016 (received for review February 26, 2016)

Reinforcement learning theory powerfully characterizes how we learn to benefit ourselves. In this theory, prediction errors—the difference between a predicted and actual outcome of a choice—drive learning. However, we do not operate in a social vacuum. To behave prosocially we must learn the consequences of our actions for other people. Empathy, the ability to vicariously experience and understand the affect of others, is hypothesized to be a critical facilitator of prosocial behaviors, but the link between empathy and prosocial behavior is still unclear. During functional magnetic resonance imaging (fMRI) participants chose between different stimuli that were probabilistically associated with rewards for themselves (self), another person (prosocial), or no one (control). Using computational modeling, we show that people can learn to obtain rewards for others but do so more slowly than when learning to obtain rewards for themselves. fMRI revealed that activity in a posterior portion of the subgenual anterior cingulate cortex/basal forebrain (sgACC) drives learning only when we are acting in a prosocial context and signals a prosocial prediction error conforming to classical principles of reinforcement learning theory. However, there is also substantial variability in the neural and behavioral efficiency of prosocial learning, which is predicted by trait empathy. More empathic people learn more quickly when benefitting others, and their sgACC response is the most selective for prosocial learning. We thus reveal a computational mechanism driving prosocial learning in humans. This framework could provide insights into atypical prosocial behavior in those with disorders of social cognition.

reinforcement learning theory | prosocial behavior | empathy | reward | subgenual anterior cingulate cortex

Prosocial behaviors, namely, social behaviors or actions intended to benefit others, are a fundamental but poorly understood aspect of social interaction (1). To behave prosocially, animals need to learn about the consequences that their actions can have for others. In reinforcement learning theory (RLT), prediction errors (PEs)—differences between expected and actual outcomes—drive learning (2). RLT provides a powerful framework for understanding how animals learn to obtain rewards for themselves (3). However, the processes by which animals learn to make choices that benefit others are unknown. Here we use RLT to characterize prosocial learning, combining functional magnetic resonance imaging (fMRI) and detailed computational modeling of behavior.

Studies using economic games, moral judgments, or charity donation tasks have consistently reported activity in the ventral striatum, posterior regions of the subgenual cingulate cortex/basal forebrain (hereinafter referred to as sgACC), dorsal anterior cingulate cortex (dACC), and dorsolateral prefrontal cortex (DLPFC) during prosocial behavior (4–7). Each of these regions receives input from midbrain dopaminergic neurons (8), and these cortical regions all project to the ventral striatum (9–11). There is substantive evidence that dopamine neurons projecting to this circuit code PEs for rewards delivered to animals (and humans) themselves (3).

The ventral striatum, sgACC, dACC, and DLPFC are also implicated in processing information about rewards others will receive (12–16), PEs when interacting with others (13, 17–19), and prosocial behavior (e.g., refs. 4–7 for reviews). Therefore, information processing in these regions may conform to RLT principles during social interactions. However, no prior work has examined how we learn to make choices that benefit others, a fundamental aspect of behaving prosocially. Do any of these areas signal a unique prosocial PE specifically when learning to benefit another? Or is learning to benefit another encoded in regions that signal PEs regardless of the beneficiary?

Moreover, although humans have a remarkable inclination to engage in prosocial behaviors, there are also substantial individual differences (1, 20–22). Empathy, the capacity to vicariously experience and understand the affect of others (23–27), has been hypothesized to be a critical motivator of prosocial behaviors (25–28). Previous studies have consistently shown that empathy can modulate neural responses to viewing others' pain (29) and viewing desirable outcomes (rewards) that will be delivered to others (15, 30). Moreover, although empathy can be broken down into separable components associated with different social behaviors and traits (23, 31, 32), studies have suggested that both cognitive and affective aspects of empathic processing may motivate prosocial behaviors (33). Despite this body of research, the mechanistic link between empathy and prosocial learning remains unknown. If empathy is indeed linked to prosocial behavior, we might predict that empathy and prosocial learning would be associated, with those higher in empathy learning more quickly to benefit others.

Participants ($n = 31$) performed a reinforcement learning task during fMRI. On each trial participants were required to choose between one of two symbols. One symbol was associated with a high probability (75%) and one was associated with a low probability

Significance

Prosocial behaviors are essential for social bonding and cohesion, but the mechanisms that underpin these behaviors are still poorly understood. Using computational modeling and neuroimaging, we show that people can learn to benefit others and that this learning is underpinned by reinforcement learning signals in the subgenual anterior cingulate cortex (sgACC). However, there is substantial individual variability in people's ability for prosocial learning. More empathic people learn faster and have more selective responses in the sgACC when benefitting others. Our results thus reveal a computational mechanism driving prosocial learning in humans and why empathy and prosocial behavior may be linked. This framework could help to explain reduced empathy and prosocial behavior in people with disorders of social cognition.

Author contributions: P.L.L., M.A.J.A., E.V., and J.P.R. designed research; P.L.L. performed research; P.L.L., V.V., and E.V. analyzed data; and P.L.L., M.A.J.A., V.V., E.V., and J.P.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. E.F. is a Guest Editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: patricia.lockwood@psy.ox.ac.uk.

[2]E.V. and J.P.R. contributed equally to this work.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1603198113/-/DCSupplemental.

(25%) of a reward. These contingencies were not instructed but had to be learned through trial and error. Critically, participants performed this task for themselves (self), for another person (prosocial), or in a control condition with no beneficiary (no one) (Fig. 1*A*). The no one condition was crucial to account for previous studies showing fictive reward or fictive prediction error brain responses, which occur when rewards are not delivered to ourselves in entirely nonsocial situations (34–36). This condition also allowed us to test for regions that showed relative specificity for processing social information, that is, regions that did not respond to self or nonself information (see also ref. 37 for a recent discussion).

Using an RLT framework, we then conducted detailed computational modeling of trial-by-trial variation of behavior, supported by Bayesian model comparison, to examine whether people were able to learn to benefit others at the same rate that they learned to benefit themselves. We examined whether activity in brain areas previously implicated in coding PEs for ourselves or in prosocial behavior signaled PEs regardless of the beneficiary that received the outcome or whether any of these regions exclusively reflected a prosocial PE when learning to benefit another [ventral striatum; sgACC (Brodmann Area (BA) 25/s24); dACC (BA24); and DLPFC (BA9/46d); *Experimental Procedures* and *SI Experimental Procedures*]. Moreover, we hypothesized that if empathy motivates prosocial behavior, then the rate at which people can learn to obtain rewards for others and the neural signatures that underpin prosocial learning should vary with trait levels of empathy.

Results

Behavioral Differences in Learning to Obtain Rewards for Self, Another Person, or No One. Participants were able to learn to obtain rewards for themselves, the other person, and no one, performing significantly above chance in all conditions (all t values > 9.1, all p values < 0.001, all degrees of freedom = 30; Fig. 1*B* and Fig. S1). Bayesian model comparison revealed that participants' choices were best characterized by a model with separate learning rates and choice variability parameters in each condition [winning model evidence (ΔBIC) > 600; see also *SI Experimental Procedures* and Fig. S2]. Comparing the learning rate parameters between conditions revealed a main effect of learning condition [$F(2,60) = 11.47$, $P < 0.001$]. Participants learned more slowly if they were obtaining rewards for another person (prosocial) ($d = 0.87$, $P < 0.001$) or no one ($d = 0.53$, $P = 0.01$) than if they were obtaining rewards for themselves (Fig. 1*C*). There was no difference in learning rate between the prosocial and no one conditions ($d = 0.25$, $P = 0.18$). Choice variability [main effect of condition: $F(2,60) = 7.87$, $P < 0.001$] could not explain these results because participants had similar consistency scores when choosing for themselves and the other person ($d = 0.24$, $P = 0.20$) but were more random when choosing for no one compared with themselves ($d = 0.46$, $P = 0.017$) and the other person ($d = 0.58$, $P = 0.003$) (Fig. 1*E*). Together, these findings suggest that people have a self-bias in their learning, learning more quickly about rewards for themselves compared with for another person or no one. However, people are similarly variable when choosing for themselves and others and most variable when no beneficiary will receive the reward.

Identifying Common and Distinct Coding of Prosocial Prediction Errors Using Functional Imaging. Using concurrently collected fMRI data, we next examined activity corresponding to the magnitude of PEs, time-locked to choice outcome and modeled independently of choice-related activity (*Experimental Procedures*). To identify common coding of self, prosocial, and no one PEs we first used a stringent conjunction-null analysis (38) (for main effects, see Tables S1 and S2; for categorical analyses of the outcome, see Table S3). Only responses bilaterally in ventral striatum, a region consistently shown to encode rewards delivered to self and others (39, 40), covaried with PEs in all three conditions (MNI coordinates [$x = 10, y = 15, z = -9$], $Z = 4.09$, $k = 91$, $P = 0.006$ voxel-level small-volume family-wise error corrected (SVC-FWE), and [$x = -12, y = 10, z = -11$], $Z = 3.72$, $k = 78$, $P = 0.023$ SVC-FWE; Fig. 2 *A–C*). Responses in each learning condition were significantly greater than 0 (all $Z > 3.72$, $P < 0.05$ SVC-FWE; Fig. S3). The ventral striatum therefore signaled PEs regardless of the beneficiary.

We then identified regions that responded to prosocial PEs exclusively by contrasting the prosocial condition against the combined self and no one conditions. The sgACC was the only region to specifically respond to prosocial PEs ([$x = -2, y = 4, z = -15$], $Z = 3.83$, $k = 148$, $P = 0.019$ SVC-FWE; Fig. 3 *A–C*), and only parameter estimates for prosocial PEs were greater than 0 ($Z = 4.95$, $P < 0.001$, SVC-FWE). The sgACC therefore uniquely signaled PEs when learning to benefit another.

We also tested for regions that showed greater responses to self/no one than prosocial PEs. Both left ([$x = -36, y = 18, z = 43$], $Z = 4.47$, $k = 62$, $P = 0.006$ SVC-FWE) and right ([$x = 32, y = 15, z = 39$], $Z = 4.36$, $k = 27$, $P < 0.020$ SVC-FWE) DLPFC showed this pattern. We did not observe significant responses in the dACC for any

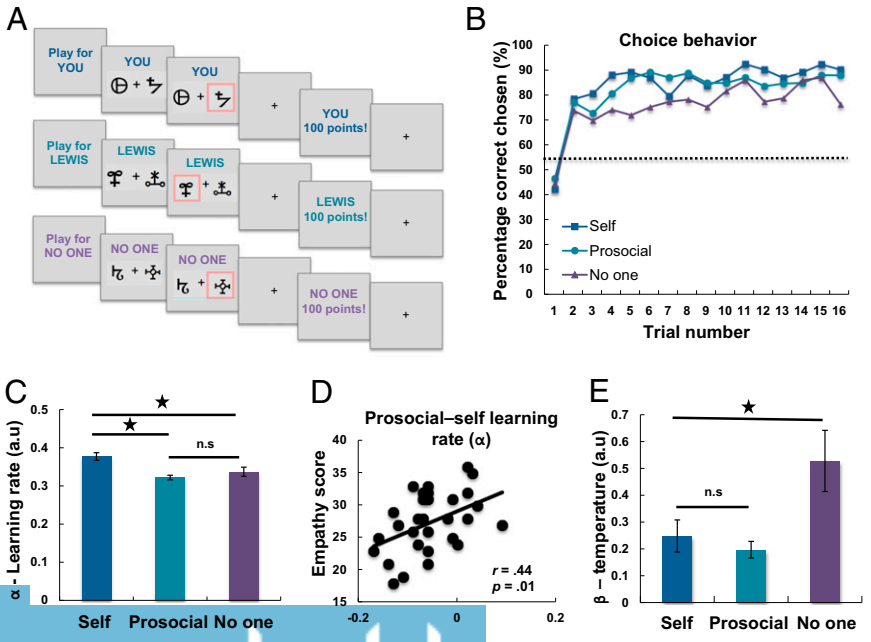

Fig. 1. Behavioral task and data. (*A*) Participants performed a reinforcement learning task in which they had to learn the probability that abstract symbols were rewarded. At the beginning of each block, participants were told whom they were playing for, either themselves, for the other participant, or in a condition where no one received the outcome. (*B*) Group-level learning curves showing choice behavior in the three learning conditions. Trials are averaged over the three blocks (48 trials total per condition, 16 trials per block) for the self, prosocial, and no one conditions. Dotted line shows chance level. (*C*) Comparison of learning rates (α) from the computational model. Participants had a significantly higher learning rate when learning in the self compared with the prosocial and no one conditions. (*D*) Individual differences in empathy (online simulation) modulated the prosocial vs. self learning rate difference, with those higher in empathy having a more similar learning rate between the prosocial and self conditions. (*E*) Participants were less consistent (higher β) when choosing for no one compared with choosing in the self and prosocial conditions. Asterisks represent significant differences ($P < 0.05$).

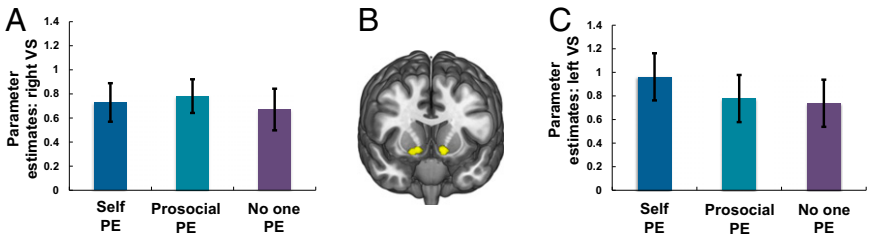

Fig. 2. fMRI data, common coding of prediction errors. (*A* and *C*) Ventral striatum responses to PEs regardless of the agent the outcome was received by. (*B*) Overlay of ventral striatum response. All images displayed at $P < 0.001$ uncorrected. Peak voxels all survive $P < 0.05$ FWE-SVC (see Fig. S3).

contrast. All significant results remained when comparing the self condition only to the prosocial condition (see Table S4).

Mechanistic Links Between Empathy and Prosocial Learning. Next we tested whether individuals higher in empathy learned at a similar rate to obtain rewards for others compared with themselves and whether variability in empathy modulated neural responses to prosocial PEs. Consistent with our predictions, we found that the online simulation subscale of the Questionnaire of Cognitive and Affective Empathy (41), a validated and psychometrically rigorous measure of trait empathy that probes the tendency to imagine how other people will feel (*SI Experimental Procedures*), was positively associated with the learning rate for the prosocial condition, relative to the self (to control for individual differences in learning per se) condition ($r = 0.44$, $P = 0.01$, 95% CI = 0.18, 0.66; Fig. 1*D* and Table S5). We then tested whether the online simulation subscale also predicted neural responses to prosocial PEs. Online simulation was also positively associated with prosocial (compared with self) PE responses in the sgACC ($r = 0.39$, $P = 0.03$, 95% CI = 0.13, 0.60; Fig. 3*D*), with those higher in online simulation showing greater sgACC specificity to prosocial relative to self PEs. Together these findings suggest that there are both behavioral and neural links between empathy and prosocial behavior.

Discussion

RLT has provided important insights into how we learn about rewarding information for ourselves. Here we used the RLT framework to characterize prosocial learning and its underlying computational and neural basis. We found that the ventral striatum commonly coded PEs in all conditions, responding to PEs for self, another person, and no one. In contrast, the sgACC coded PE signals specifically when learning to benefit another. We also observed substantial variability in prosocial learning with differences in both neural and behavioral responses predicted by trait empathy.

Our findings advance theoretical accounts of the neural basis of social behavior by finding evidence for both common coding and socially specific regions of the brain underpinning prosocial behavior (4–7, 42). PE signals in the ventral striatum, which is extensively connected with the sgACC (9), were evident regardless of the context of learning. This finding cannot be easily accommodated within current theories of ventral striatum contributions to learning, which suggest that this region is engaged when learning to obtain beneficial outcomes for oneself (3). Moreover, this finding adds to and extends existing studies of the role of the ventral striatum in social behavior. Consistent with previous research, we find ventral striatum responses to rewards delivered to both self and other (39, 40). However, we also find that these signals are evident even when no one receives a rewarding outcome. The ventral striatum may therefore be important for learning in many contexts even when a reward is not obtained or consumed by anyone. The profile of response in the ventral striatum (VS) differs from the sgACC, which shows specificity for signaling PEs when learning to benefit another. Thus, although both regions may play a role in driving prosocial behavior, the function of ventral striatum may be more domain general than that of sgACC, which we speculate may compute a prediction error specifically for outcomes delivered to others.

The sgACC region we identified as responding exclusively to prosocial PEs overlapped with a septal-anterior hypothalamic area that is part of the basal forebrain (43). Recent studies of sgACC function have found signals in this region relevant for social cognition and behavior including credit assignment (44), prosocial and moral behavior (6, 16, 45, 46), the experience of positive affect (47), trust (48), social emotion (49), and vicarious reward (50, 51). Moreover, there is evidence that the sgACC may signal PEs for self-reward but only when learning occurs at a specific level of abstraction beyond basic stimulus–response association (52). One possible explanation for this convergence is that similar abstract learning mechanisms may drive how we learn to benefit others and that both are underpinned by RLT principles. Further studies could compare tasks that manipulate both the social and the hierarchical context of a learning environment to directly test the parallels between these types of learning. In addition, future research should aim to dissociate the functions of the heterogeneous cortical and subcortical structures of the sgACC complex, using high-resolution fMRI, to understand which of these subregions contribute to social cognition and prosocial learning.

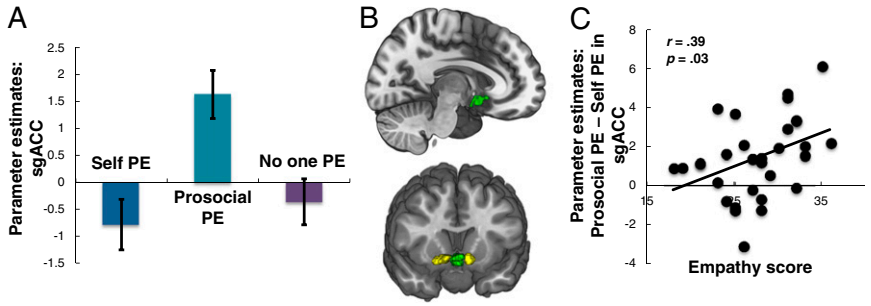

Fig. 3. fMRI data, distinct coding of prosocial prediction errors. (*A*) Subgenual anterior cingulate cortex (sgACC) responses to PEs when learning to benefit another person. (*B*) Overlay of sgACC response (green) and VS response (yellow). Image displayed at $P < 0.001$ uncorrected, and all peak voxels survive $P < 0.05$ FWE-SVC. Crucially, the sgACC region in which activity covaried with prosocial PEs did not overlap with the ventral striatum clusters. (*C*) sgACC response was modulated by individual differences in empathy (online simulation).

We did not observe responses in anterior insula or dACC in any of our contrasts. This may seem surprising given that these regions have previously been implicated in empathy and/or social behavior (15, 29, 30, 37). Although it is difficult to interpret a null finding, because there can be a number of reasons that a particular neural response is not observed, we note that in this experiment, participants were performing their task in a different reference frame (see refs. 53 and 54 for recent discussions of the role of references frames in studies of social cognition) compared with other studies measuring empathic/vicarious processing. In our task, participants were making choices for another person, not only observing events that happened to others. In other words, in our task, decisions were made in a self-action reference frame. Studies that compare different reference frames in the same paradigm could help shed light on the functional roles of specific brain areas during empathy and prosocial behavior.

An important aim of our study was to explain what might drive variability in prosocial learning (1, 20–22). We identify evidence of a mechanism linking variability in empathy to variability in prosocial behavior, with highly empathic individuals having an increased learning rate, and stronger sgACC PE signals, for other people's rewards. Many accounts of empathy have argued for a crucial role of empathy in the development of prosocial and moral behavior and the inhibition of aggression (25, 27, 55–58). Our demonstration that empathy is associated with a higher rate of learning about actions that result in beneficial outcomes for other people as well as the neural drivers of prosocial learning suggests a computational link by which empathy could influence the development of prosocial and moral behaviors. Enhanced PE signaling and faster learning when benefitting others also provide a potential explanation for reports of individuals higher in empathy being more motivated to behave in a prosocial manner compared with those lower in empathy (21). Moreover, our results support an emerging view that PE signals may be crucial for learning how to interact, and empathize, with others (59). Twin data indicate substantial heritability of prosociality across development (60, 61). Longitudinal developmental investigations, particularly ones that are able to tease apart genetic and environmental contributions to brain function (e.g., twin studies), would be helpful in determining the degree to which the specificity of sgACC activity during prosocial learning reflects an endophenotype for prosocial behavior.

An influential theory within the literature on empathy is that we empathize with others by a process of (embodied) simulation (62). This view is largely driven by studies that show a degree of overlap in the neural responses to self and other pain, particularly in anterior insula and dACC (reviewed in refs. 29 and 30), and for pleasant touch (63). Other studies have also supported a simulationist view of empathy by exploiting the placebo analgesia effect, showing that placebo analgesia changes self pain as well as vicarious pain (64, 65). Although these studies are consistent with a simulationist account, it remains possible that there exist additional processes that do not operate through self–other overlap that participate in the experience of empathy (15, 30). For example, studies have suggested that other neurocognitive processes, in addition to simulation, may be important when processing vicarious information, particularly in the domain of positive affect (see ref. 51 for a metaanalysis and ref. 30 for review). In the present study we observed that PE responses in the sgACC were present only when learning to benefit another person and not when learning to benefit oneself or no one (including the latter to control for fictive PE signals), suggesting that a socially specific signal is important for prosocial learning in this context. We also observed that those who self-report that they simulate most readily also have the most specific signals related to prosocial learning in the sgACC. This points to the intriguing possibility that simulation at the level of self-report may not necessarily be encoded in brain areas that respond to both self and other during learning. It should be noted that given potential gender differences in empathy and prosocial behavior (e.g., ref. 66) our sample in this study was composed only of males. Future studies would benefit from also examining prosocial learning in females.

Prosocial behaviors are fundamental for promoting social bonds and cohesion (1, 55, 67) and are disrupted in a number of psychiatric and neurological disorders (23, 68–70). Using the framework of RLT to understand how we learn to make decisions that benefit other people could offer insights into why these disorders are associated with atypical prosocial behavior and empathy. Taken together, our findings reveal a computational link between prosocial learning and empathy in humans and therefore pave the way to characterize atypical prosocial interactions in those with disorders of social cognition and behavior.

Experimental Procedures

Participants. Thirty-four right-handed healthy males (age 19–32, M = 22.7, SD = 3.0) were recruited through university participant databases. Exclusion criteria included previous or current neurological or psychiatric disorder, nonnormal or noncorrected to normal vision, nonnative English language, and previous or current study of psychology. Three participants were excluded from the analysis [two due to performance at chance level (~50%) in all learning conditions and one due to neurological abnormalities evident on the MRI scan], leaving a final sample of 31. With 31 subjects we had 80% power to detect a medium effect size of $d = 0.52$ at alpha = 0.05 (two-tailed), an effect size smaller than typically reported in this field, indicating sufficient power. All participants gave written informed consent, and the study was approved by the University College London Research Ethics Committee.

Experimental Task. We examined BOLD signals that scaled parametrically with the size of a PE at the time of an outcome delivered to self, another person (here a confederate–prosocial condition), or no one. Participants performed a probabilistic reinforcement learning task where they were required to learn the probability that each of two symbols would be rewarded. One symbol of each pair was associated with a high probability (75%) and one was associated with a low probability (25%) of reward. Participants performed this task in three different learning contexts: self, prosocial, and no one. Participants were instructed that when they were playing for themselves they would receive any money they won. Crucially, when they were playing for the confederate, that participant would receive the money (see *SI Experimental Procedures* for full instructions given to participants). When they were playing for no one, the points they saw would not be converted into any additional payment, either for themselves or the other participant. Participants were informed that the other participant was not aware that they were performing a task where they could earn extra money and that any money they won would be given to the other participant anonymously (i.e., it would be placed in a sealed envelope and the two participants would leave the scanning center at different times).

Self blocks began with the instruction "play for you" and had the word "you" written above all choice symbols and outcomes. Prosocial blocks had the name of the confederate participant written above them. No one blocks had the words "no one" written above elements in a trial. This ensured that participants were explicitly aware whether the decisions they made resulted in outcomes for themselves, for the other participant, or for no one (for trial structure and order, see *SI Experimental Procedures* and Fig. 1*A*). Participants practiced one block (16 trials) of the task in a separate session ~7 d before the scanning session to familiarize them with the experimental task. During this practice they were instructed that the outcomes would not be converted into any payment.

Procedure. Participants were paired with one of two age- and gender-matched confederates whom they believed were naïve participants and had never met before the experiment. The confederates were trained in acting as naïve participants during a pilot experiment. Participants attended two sessions. The first session was attended only by the experimental participant and involved practicing the experimental task and completing questionnaires. This was done due to scheduling considerations and so that participants could practice the learning task on their own without the confederate present. The second session (<7 d later) was attended by both the experimental participant and the confederate. The participant and confederate were taken together to the MRI center and filled in consent forms together in the same room. The confederate was then led into a behavioral testing room and instructed to complete questionnaires, with the experimental participant within earshot of this interaction to increase belief in the deception. The experimental participant was taken to the scanning room and reminded of the instructions for the task, with the confederate participant unable to overhear this interaction to ensure that the experimental participants' choices remained anonymous. Participants were told that they would view a pair of symbols on each trial and that they should select one of them. They would receive points for some of their choices that would be converted into money at the end of the experiment, such that the more points they received, the more extra money they would earn. They were instructed that the two symbols would

not be the same in terms of how often they gave points, and with some symbols they were more likely to win points than with other symbols. Whether the symbols appeared on the left or right did not affect their meaning.

Participants were instructed that they would receive extra payment based on the outcomes they received during the experimental task, but in fact, all participants were paid the same amount due to ethical restrictions (total £30, representing an additional £7 to the standard participant payment for the required time commitment). They also believed that the confederate participant could earn an extra payment based on the choices the experimental participant made during the task. A set of standardized questions completed after the scan confirmed that no participant had become suspicious about the deception during the experiment (see *SI Experimental Procedures*).

Computational Modeling of Behavioral Data. Learning behavior in the self, prosocial, and no one conditions was modeled using a reinforcement learning (RL) algorithm (2), which has been extensively used to examine the behavioral and neural basis of arbitrary visuomotor associations in both self and social contexts (13, 17–19). The RL model assumes that the associative value of an action (or stimulus) changes when new information reveals that the actual outcome of a decision is different from the expected outcome (2). Thus, on each trial t, an action a has an expected associative value $Q_t(a)$ that is updated by the mismatch between experienced and expected outcome on the current trial (the PE). At their most simple, RL algorithms state that expectations of future reward for action a, $Q_{t+1}(a)$, should be a function of current expectations $Q_t(a)$ and the discrepancy between the actual reward that has just been experienced on this trial r_t (coded as 1 or 0 for reward or no reward, respectively) and the expected reward for this trial t, $Q_t(a)$. The degree to which this discrepancy updates the expectation is scaled by the learning rate α (bounded between 0 and 1), such that

$$Q_{t+1}(a) = Q_t(a) + \alpha \times \underbrace{[r_t - Q_t(a)]}_{\text{Prediction Error}}.$$

The learning rate α controls the extent to which the current expected value is updated by new information. Consequently, a low learning rate will minimize the influence of the prediction error and the amount that the value is updated. The probability that a subject chooses action a on trial t, given the expected values of the available actions $Q_t(a)$, is given by the softmax link function

$$p_t[a|Q_t(a)] = \frac{e^{[Q_t(a)/\beta]}}{\sum_{a'} e^{[Q_t(a')/\beta]}}.$$

The temperature parameter β controls the amount of exploration or noisiness for that participant (i.e., extent to which the subject decides to choose the most rewarding option vs. exploring potentially more rewarding actions). The softmax link function estimates the trial-by-trial probability of each action by weighting the ratio of expected values by the temperature parameter. In this framework, a high temperature parameter β would lead to similar action probabilities irrespective of the expected value of each action (resulting in random behavior). A low β would lead to consistent behavior, where the action with the higher expected value is invariably selected on each trial. In the full model, separate α and β parameters were estimated for each of the self, prosocial, and no one conditions because this provided the most parsimonious explanation for the behavioral data (see *SI Experimental Procedures* for details of model fitting and model comparison and Fig. S2).

Statistical Analysis of Behavioral Data. Analyses of behavioral data were performed in SPSS 22 (IBM Corp.). We examined differences between conditions in the learning rate and temperature parameters at the group level using separate repeated measures analyses of variance (ANOVAs), with three levels (self, prosocial, and no one). Separate ANOVAs were conducted because the learning rate and temperature parameters represent different units of measurement. We examined bivariate associations between the prosocial–self difference in learning rate and temperature and empathy questionnaire subscales using the Pearson correlation coefficient. Effect sizes (Cohen's d) were calculated by dividing the mean difference between conditions by the SD of the difference (71). Confidence intervals for correlation coefficients were estimated using SPSS 22 Bootstrap procedure.

fMRI Acquisition and Analysis. A Siemens Avanto 1.5-T MRI scanner was used to acquire a 5.5-min three-dimensional T1-weighted structural scan and 424 multislice T2*-weighted echo planar volumes with blood oxygenation-level–dependent (BOLD) contrast. The structural scan was acquired using a magnetization prepared rapid gradient echo (MPRAGE) sequence with 176 slices, slice thickness = 1 mm, gap between slices = 0.5 mm, TR = 2730 ms, TE = 3.57 ms, field of view = 256 mm × 256 mm^2, matrix size = 256 × 256, and voxel size = 1 × 1 × 1 mm resolution. The functional imaging sequence was acquired in an ascending manner, at an oblique angle (~30°) to the AC–PC line to decrease the effect of susceptibility artifact in the orbitofrontal cortex (72), and had the following acquisition parameters: 424 volumes, 1 mm gap, echo time = 50 ms, repetition time = 2,975 ms, flip angle = 90°, field of view = 192 mm, and matrix size = 64 × 64.

Imaging data were analyzed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Data preprocessing followed a standard sequence: the first four volumes and last volume were discarded. Images were then realigned and coregistered to the participant's own anatomical image. The anatomical image was processed using a unified segmentation procedure combining segmentation, bias correction, and spatial normalization to the MNI template using the new segment procedure (73); the same normalization parameters were then used to normalize the EPI images. The voxel size was resampled to 1.5 × 1.5 × 1.5mm. Last, a Gaussian kernel of 8 mm FWHM was applied to spatially smooth the images. Before the study, example first-level design matrices were checked to ensure that estimable GLMs could be performed with independence between the parametric regressors (chosen value and PE in the three conditions), with correlations coefficients of $r <$ 0.25. This allowed us to look at PE-related responses independent of chosen value.

Eight event types were used to construct regressors in which event timings were convolved with SPM's canonical hemodynamic response function. The three conditions at the time of the cues and three conditions at the time of the outcome were modeled as separate regressors using stick functions. Each of these regressors was associated with a parametric modulator taken from the computational model. At the time of the cue this was the chosen value, and at the time of the outcome, the PE. The PEs were estimated using each subject's own alpha and beta from each condition. The instruction cue at the beginning of each block was also modeled in a single regressor as a stick function. In some participants, an eighth regressor modeled all missed trials, on which participants did not select one of the two symbols in the response window. For those participants where there was visible head motion in a particular scan (scans with >1 mm or 1° movement relative to the next were examined visually) an extra regressor was included corresponding to each scan. These images were removed and replaced with an image created by interpolating the two adjacent images to prevent distortion of the between-subjects mask (four participants, less than 1% of total time series). Six head motion parameters modeled the residual effects of head motion as covariates of no interest. Data were high-pass filtered at 128 s to remove low-frequency drifts, and the statistical model included an AR (1) autoregressive function to account for autocorrelations intrinsic to the fMRI time series. Our primary analysis focused on the PEs at outcome (for response to chosen value, see Table S1).

Contrast images from the first level were input into a second-level flexible-factorial design with three levels (self PE, prosocial PE, and no one PE). Main effects are reported at $P < 0.05$, family-wise error (FWE) corrected at the voxel level across the whole brain, or $P < 0.05$ small volume corrected (SVC) at the voxel level in regions where we had a strong a priori hypothesis (see below).

ROI Selection and fMRI Contrasts. The a priori regions of interest (ROIs) were defined anatomically using masks taken from an appropriate atlas [bilateral VS, sgACC, dACC, bilateral DLPFC; toolboxes: Harvard-Oxford Atlas, regions 46v and 9 (74), Anatomy Toolbox regions s24 and 25 (75), and region 24 from ref. 76 (corresponding to the gyral portion of the anterior cingulate cortex); see Fig. S4 and *SI Experimental Procedures* for further details of ROI selection). We additionally applied a false discovery rate correction (FDR) (77), rather than Bonferroni correction [which may be overly conservative given that our ROIs were not entirely independent from one another, because they are functionally and anatomically connected (9–11)], for the number of ROIs. All ROI comparisons remained significant ($P < 0.05$) when controlling for the number of comparisons using FDR.

ACKNOWLEDGMENTS. We thank Mr. Lewis Pollock and Mr. Jonathan Blott for acting as the confederates and assistance with data collection and Dr. Joe Devlin and Dr. Ana Seara-Cardoso for helpful discussions. This work was supported by a studentship from the Medical Research Council to P.L.L. and funding from the Birkbeck–University College London Centre for Neuroimaging. M.A.J.A. was supported by an anniversary future leaders fellowship from the Biotechnology and Biological Sciences Research Council (Award BB/M013596/1). E.V. is a Royal Society Wolfson Research Merit Award holder.

1. Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425(6960):785–791.
2. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
3. Schultz W (2013) Updating dopamine reward signals. *Curr Opin Neurobiol* 23(2):229–238.
4. Lee D, Seo H (2016) Neural basis of strategic decision making. *Trends Neurosci* 39(1): 40–48.
5. Rilling JK, Sanfey AG (2011) The neuroscience of social decision-making. *Annu Rev Psychol* 62:23–48.

6. Moll J, Schulkin J (2009) Social attachment and aversion in human moral cognition. *Neurosci Biobehav Rev* 33(3):456–465.
7. Ruff CC, Fehr E (2014) The neurobiology of rewards and values in social decision making. *Nat Rev Neurosci* 15(8):549–562.
8. Williams SM, Goldman-Rakic PS (1998) Widespread origin of the primate mesofrontal dopamine system. *Cereb Cortex* 8(4):321–345.
9. Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. *J Neurosci* 15(7 Pt 1):4851–4867.
10. Lynd-Balta E, Haber SN (1994) Primate striatonigral projections: A comparison of the sensorimotor-related striatum and the ventral striatum. *J Comp Neurol* 345(4):562–578.
11. Yeterian EH, Pandya DN (1991) Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. *J Comp Neurol* 312(1):43–67.
12. Apps MAJ, Ramnani N (2014) The anterior cingulate gyrus signals the net value of others' rewards. *J Neurosci* 34(18):6190–6200.
13. Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS (2008) Associative learning of social value. *Nature* 456(7219):245–249.
14. Chang SWC, Gariépy J-F, Platt ML (2013) Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci* 16(2):243–250.
15. Lockwood PL, Apps MAJ, Roiser JP, Viding E (2015) Encoding of vicarious reward prediction in anterior cingulate cortex and relationship with trait empathy. *J Neurosci* 35(40):13720–13727.
16. Moll J, et al. (2006) Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc Natl Acad Sci USA* 103(42):15623–15628.
17. Apps MAJ, Lesage E, Ramnani N (2015) Vicarious reinforcement learning signals when instructing others. *J Neurosci* 35(7):2904–2913.
18. Burke CJ, Tobler PN, Baddeley M, Schultz W (2010) Neural mechanisms of observational learning. *Proc Natl Acad Sci USA* 107(32):14431–14436.
19. Suzuki S, et al. (2012) Learning to simulate others' decisions. *Neuron* 74(6):1125–1137.
20. Crockett MJ, Kurth-Nelson Z, Siegel JZ, Dayan P, Dolan RJ (2014) Harm to others outweighs harm to self in moral decision making. *Proc Natl Acad Sci USA* 111(48):17320–17325.
21. Silk JB, House BR (2011) Evolutionary foundations of human prosocial sentiments. *Proc Natl Acad Sci USA* 108(Suppl 2):10910–10917.
22. Sul S, et al. (2015) Spatial gradient in value representation along the medial prefrontal cortex reflects individual differences in prosociality. *Proc Natl Acad Sci USA* 112(25):7851–7856.
23. Bird G, Viding E (2014) The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neurosci Biobehav Rev* 47:520–532.
24. Decety J, Jackson PL (2004) The functional architecture of human empathy. *Behav Cogn Neurosci Rev* 3(2):71–100.
25. de Vignemont F, Singer T (2006) The empathic brain: How, when and why? *Trends Cogn Sci* 10(10):435–441.
26. Hoffman ML (2008) Empathy and prosocial behavior. Handbo*o*k *o*f Emot*i*ons, eds Lewis M, Haviland-Jones JM, Barrett LF (Guilford Press, New York), Vol 3, pp 440–455.
27. Singer T, Lamm C (2009) The social neuroscience of empathy. *Ann N Y Acad Sci* 1156:81–96.
28. Ballesta S, Duhamel J-R (2015) Rudimentary empathy in macaques' social decision-making. *Proc Natl Acad Sci USA* 112(50):15516–15521.
29. Lamm C, Decety J, Singer T (2011) Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* 54(3):2492–2502.
30. Lockwood PL (2016) The anatomy of empathy: Vicarious experience and disorders of social cognition. *Behav Brain Res* 311:255–266.
31. Lockwood PL, Bird G, Bridge M, Viding E (2013) Dissecting empathy: High levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Front Hum Neurosci* 7:760.
32. Blair (2005) Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious Cogn* 14(4):698–718.
33. Lockwood PL, Seara-Cardoso A, Viding E (2014) Emotion regulation moderates the association between empathy and prosocial behavior. *PLoS One* 9(5):e96555.
34. Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci USA* 104(22):9493–9498.
35. Hayden BY, Pearson JM, Platt ML (2009) Fictive reward signals in the anterior cingulate cortex. *Science* 324(5929):948–950.
36. Gu X, Kirk U, Lohrenz TM, Montague PR (2014) Cognitive strategies regulate fictive, but not reward prediction error signals in a sequential investment task. *Hum Brain Mapp* 35(8):3738–3749.
37. Apps MAJ, Rushworth MFS, Chang SWC (2016) The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron* 90(4):692–707.
38. Nichols T, Brett M, Andersson J, Wager T, Poline J-B (2005) Valid conjunction inference with the minimum statistic. *Neuroimage* 25(3):653–660.
39. Báez-Mendoza R, Harris CJ, Schultz W (2013) Activity of striatal neurons reflects social action and own reward. *Proc Natl Acad Sci USA* 110(41):16634–16639.
40. Harbaugh WT, Mayr U, Burghart DR (2007) Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316(5831):1622–1625.
41. Reniers RLEP, Corcoran R, Drake R, Shryane NM, Völlm BA (2011) The QCAE: A questionnaire of cognitive and affective empathy. *J Pers Assess* 93(1):84–95.
42. Rilling JK, King-Casas B, Sanfey AG (2008) The neurobiology of social decision-making. *Curr Opin Neurobiol* 18(2):159–165.
43. Zaborszky L, et al. (2008) Stereotaxic probabilistic maps of the magnocellular cell groups in human basal forebrain. *Neuroimage* 42(3):1127–1141.
44. Akaishi R, Kolling N, Brown JW, Rushworth M (2016) Neural mechanisms of credit assignment in a multicue environment. *J Neurosci* 36(4):1096–1112.
45. Wiech K, et al. (2013) Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. *Cognition* 126(3):364–372.
46. Zahn R, de Oliveira-Souza R, Bramati I, Garrido G, Moll J (2009) Subgenual cingulate activity reflects individual differences in empathic concern. *Neurosci Lett* 457(2):107–110.
47. Rudebeck PH, et al. (2014) A role for primate subgenual cingulate cortex in sustaining autonomic arousal. *Proc Natl Acad Sci USA* 111(14):5391–5396.
48. Krueger F, et al. (2007) Neural correlates of trust. *Proc Natl Acad Sci USA* 104(50): 20084–20089.
49. Moll J, et al. (2011) Impairment of prosocial sentiments is associated with frontopolar and septal damage in frontotemporal dementia. *Neuroimage* 54(2):1735–1742.
50. Mobbs D, et al. (2009) A key role for similarity in vicarious reward. *Science* 324(5929):900.
51. Morelli SA, Sacchet MD, Zaki J (2015) Common and distinct neural correlates of personal and vicarious reward: A quantitative meta-analysis. *Neuroimage* 112:244–253.
52. Iglesias S, et al. (2013) Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 80(2):519–530.
53. Chang SWC (2013) Coordinate transformation approach to social interactions. *Front Neurosci* 7:147.
54. Hunt L, Behrens T (2011) Frames of reference in human social decision making. *Neural Basis of Motivational and Cognitive Control* (MIT Press, Cambridge, MA), pp 409–424.
55. Decety J (2011) The neuroevolution of empathy. *Ann N Y Acad Sci* 1231(1):35–45.
56. Decety J (2010) The neurodevelopment of empathy in humans. *Dev Neurosci* 32(4): 257–267.
57. Eisenberg N (2000) Emotion, regulation, and moral development. *Annu Rev Psychol* 51(1):665–697.
58. Eisenberg N, Miller PA (1987) The relation of empathy to prosocial and related behaviors. *Psychol Bull* 101(1):91–119.
59. Hein G, Engelmann JB, Vollberg MC, Tobler PN (2016) How learning shapes the empathic brain. *Proc Natl Acad Sci USA* 113(1):80–85.
60. Gregory AM, Light-Häusermann JH, Rijsdijk F, Eley TC (2009) Behavioral genetic analyses of prosocial behavior in adolescents. *Dev Sci* 12(1):165–174.
61. Knafo-Noam A, Uzefovsky F, Israel S, Davidov M, Zahn-Waxler C (2015) The prosocial personality and its facets: Genetic and environmental architecture of mother-reported behavior of 7-year-old twins. *Front Psychol* 6:112.
62. Bernhardt BC, Singer T (2012) The neural basis of empathy. *Annu Rev Neurosci* 35:1–23.
63. Lamm C, Silani G, Singer T (2015) Distinct neural networks underlying empathy for pleasant and unpleasant touch. *Cortex* 70:79–89.
64. Rütgen M, Seidel E-M, Riečanský I, Lamm C (2015) Reduction of empathy for pain by placebo analgesia suggests functional equivalence of empathy and first-hand emotion experience. *J Neurosci* 35(23):8938–8947.
65. Rütgen M, et al. (2015) Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proc Natl Acad Sci USA* 112(41): E5638–E5646.
66. Schulte-Rüther M, Markowitsch HJ, Shah NJ, Fink GR, Piefke M (2008) Gender differences in brain networks supporting empathy. *Neuroimage* 42(1):393–403.
67. Fehr E, Rockenbach B (2004) Human altruism: Economic, neural, and evolutionary perspectives. *Curr Opin Neurobiol* 14(6):784–790.
68. Anderson NE, Kiehl KA (2012) The psychopath magnetized: Insights from brain imaging. *Trends Cogn Sci* 16(1):52–60.
69. Lockwood PL, et al. (2013) Association of callous traits with reduced neural response to others' pain in children with conduct problems. *Curr Biol* 23(10):901–905.
70. Henry JD, von Hippel W, Molenberghs P, Lee T, Sachdev PS (2016) Clinical assessment of social cognitive function in neurological disorders. *Nat Rev Neurol* 12(1):28–39.
71. Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences* (Routledge, Hillsdale, NJ), 3rd Ed.
72. Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19(2 Pt 1):430–441.
73. Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26(3):839–851.
74. Sallet J, et al. (2013) The organization of dorsal frontal cortex in humans and macaques. *J Neurosci* 33(30):12255–12274.
75. Palomero-Gallagher N, et al. (2015) Functional organization of human subgenual cortical areas: Relationship between architectonical segregation and connectional heterogeneity. *Neuroimage* 115:177–190.
76. Beckmann M, Johansen-Berg H, Rushworth MFS (2009) Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci* 29(4): 1175–1190.
77. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 57(1):289–300.
78. Hogan R (1969) Development of an empathy scale. *J Consult Clin Psychol* 33(3): 307–316.
79. Davis M (1983) Measuring individual differences in empathy: Evidence for a multidimensional approach. *J Pers Soc Psychol* 44(1):113–126.
80. Mehrabian A, Epstein N (1972) A measure of emotional empathy. *J Pers* 40(4):525–543.
81. Baron-Cohen S, Wheelwright S (2004) The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J Autism Dev Disord* 34(2):163–175.
82. Daw ND (2011) Trial-by-trial data analysis using computational models. *Decision Making, Affect, and Learning Attention and Performance XXIII*, eds Delgado MR, Phelps EA, Robbins TW (Oxford Univ Press, Oxford), pp 3–38.
83. Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90(430):773–795.
84. Fletcher PC, et al. (2001) Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nat Neurosci* 4(10):1043–1048.
85. Waytz A, Zaki J, Mitchell JP (2012) Response of dorsomedial prefrontal cortex predicts altruistic behavior. *J Neurosci* 32(22):7646–7650.