Title: Neurocomputational mechanisms of prosocial learning and links to empathy

Patricia. L. Lockwood[1,2], Matthew A. J. Apps[2], Vincent Valton[3], Essi Viding[1*] Jonathan P. Roiser[3*]

[1]Division of Psychology and Language Sciences, University College London, London, WC1H 6BT United Kingdom

[2]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom, OX1 3UD.

[3]Institute of Cognitive Neuroscience, University College London, London, WC1N 3AZ, United Kingdom.

Corresponding author email: patricia.lockwood@psy.ox.ac.uk

Keywords: Reinforcement learning theory, prosocial behaviour, empathy, reward, subgenual cingulate cortex, ventral striatum, functional MRI

Classification: Social sciences. Psychological and cognitive sciences

Additional footnotes: * Equal senior contribution.

Abstract

Reinforcement learning theory powerfully characterises 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 hypothesised to be a critical facilitator of prosocial behaviours, but the link between empathy and prosocial behaviour is 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 modelling 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 behavioural 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 novel computational mechanism driving prosocial learning in humans. This framework could provide new insights into atypical prosocial behaviour in those with disorders of social cognition.

Significance statement

Prosocial behaviours are essential for social bonding and cohesion, but the mechanisms that underpin these behaviours 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 novel computational mechanism driving prosocial learning in humans and why empathy and prosocial behavior may be linked. This new framework could help to explain reduced empathy and prosocial behavior in people with disorders of social cognition.

\body

Introduction

Prosocial behaviours, namely social behaviours 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 characterise '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 (hereon in referred to as sgACC), dorsal anterior cingulate cortex (dACC) and dorsolateral prefrontal cortex (DLPFC) during prosocial behaviour (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. 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

that benefits another individual encoded in regions that signal PEs regardless of the beneficiary?

Although humans have a remarkable inclination to engage in prosocial behaviours there exist substantial individual differences (1, 20–22). Empathy, the capacity to vicariously experience and understand the affect of others (23–27), has been hypothesised to be a critical motivator of prosocial behaviours (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, whilst 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 behaviours (33). Despite this body of research, the mechanistic link between empathy and prosocial learning remains unknown. If empathy is indeed linked to prosocial behaviour 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 functional magnetic resonance imaging (fMRI). On each trial they were required to choose between one of two symbols. One symbol was associated with a high probability (75%) and one a low probability (25%) of a reward. These contingencies were not instructed but had to be learnt 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. 1a). 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 non-social 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 'non-self' information (see also (37) for a recent discussion).

Using an RLT framework we conducted detailed computational modelling of trial-by-trial variation of behaviour, 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 behaviour signalled 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) – see Experimental Procedures). Moreover, we hypothesised that if empathy motivates prosocial behaviour 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.

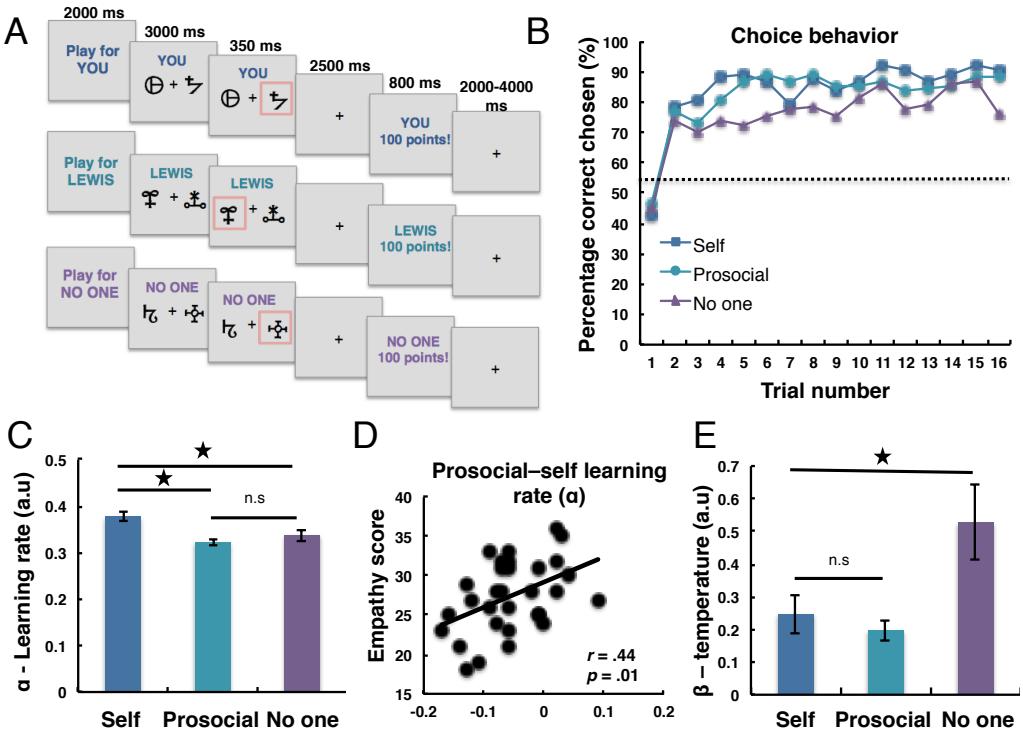

Figure 1: Behavioural task and data. **(a)** Participants performed a reinforcement-learning task in which they had to learn the probability that abstract symbols were rewarded. At the beginning of each block participants were told whom they were playing for, either themselves, the other participant or in a condition where no-one received the outcome. (b) Group level learning curves showing choice behaviour 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 to 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 choosing in the self and prosocial conditions. Asterisks represent significant differences (p<.05).

Results

Behavioural 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 ts>9.1, all ps<.001, all dfs=30; Fig. 1b). Bayesian model comparison revealed that participants' choices were best characterised by a model with separate learning rates and choice variability parameters in each condition (winning model evidence (ΔBIC)>600; see Online methods & Fig. S2. Comparing the learning rate parameters between conditions revealed a main effect of learning condition ($F(2,60)=11.47$, $p<.001$). Participants learnt more slowly if they were obtaining rewards for another person (prosocial) ($d=.87$, $p<.001$) or no-one ($d=.53$, $p=.01$) than if they were obtaining rewards for themselves (Fig. 1c). There was no difference in learning rate between the prosocial and no-one conditions ($d=.25$, $p=.18$). Choice variability (main effect of condition: $F(2,60)=7.87$, $p<.001$) could not explain these results, as participants had similar consistency scores when choosing for themselves and the other person ($d=.24$, $p=.20$), but were more random when choosing for no-one

compared to themselves (*d*=.46, *p*=.017) and the other person (*d*=.58, *p*=.003) (Fig. 1e). Together, these findings suggest that people have a 'self-bias' in their learning, learning more quickly about rewards for themselves compared to 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 modelled independently of choice-related activity (see 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&2). 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*=.006 voxel-level small-volume family-wise error corrected (SVC-FWE); and [x=-12, y=10, z=-11], Z=3.72, k=78, *p*=.023 SVC-FWE; Fig. 2a-c). Responses in each learning condition were significantly greater than 0 (all Z>3.72, p<.05 SVC-FWE; Fig.3). The ventral striatum therefore signalled PEs regardless of the beneficiary.

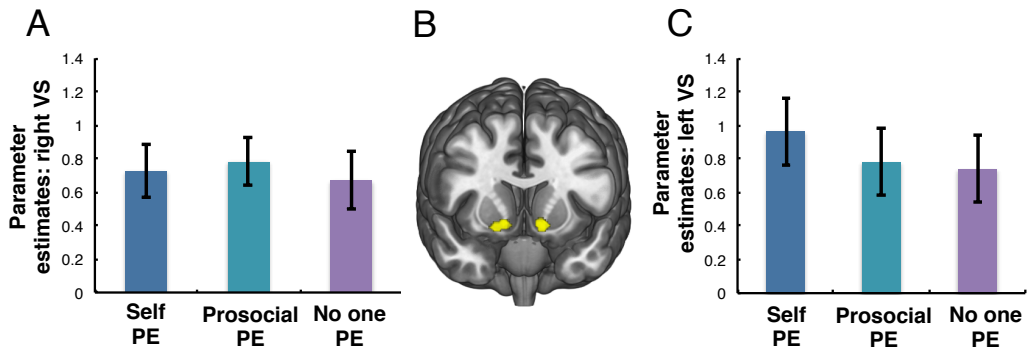

Figure 2: fMRI data, common coding of prediction errors. (a) Ventral striatum responses to PEs regardless of the agent the outcome was received by. **(b)** Overlay of ventral striatum response. All images displayed at p<.001 uncorrected. Peak voxels all survive p<.05 FWE-SVC (see Fig. S2).

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*=.019 SVC-FWE; Fig. 3a-c) and only parameter estimates for prosocial PEs were greater than 0 (Z=4.95, p<.001, SVC-FWE). The sgACC therefore uniquely signalled PEs when learning to benefit another.

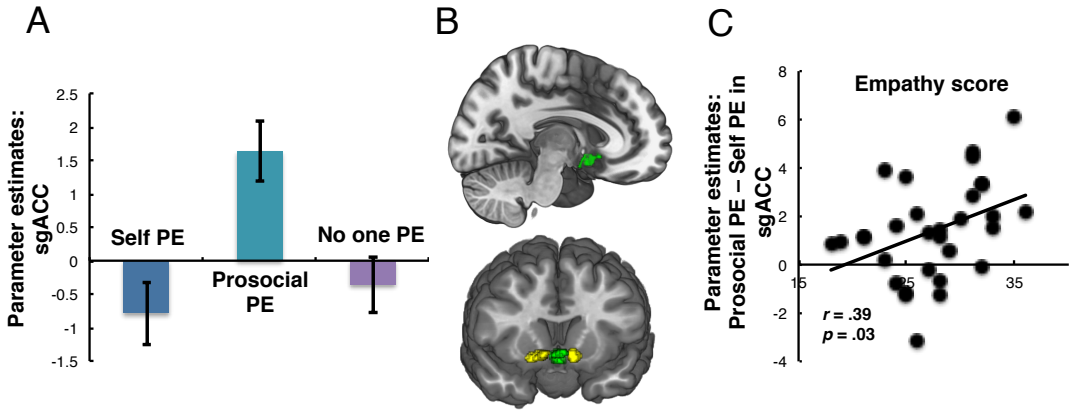

Figure 3: fMRI data, distinct coding of prosocial prediction errors. (a) subgenual anterior cingulate cortex (sgACC) responses to PEs when learning to benefit another person. **(b)** Overlay of sgACC response (green) and VS response (yellow). Image displayed at p<.001 uncorrected and all peak voxels survive p<.05 FWE-SVC. Crucially the sgACC region in which activity covaried with prosocial PEs did not overlap with the vental striatum clusters. **(c)** sgACC response was modulated by individual differences in empathy (online simulation).

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*=.006 SVC-FWE) and right ([x=32, y=15, z=39], Z=4.36, k=27, *p*<.020 SVC-FWE) DLPFC showed this pattern. We did not observe significant responses in the dACC for any contrast.

Mechanistic links between empathy and prosocial learning

Next we tested whether individuals higher in empathy learnt at a similar rate to obtain rewards for others compared to 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 (see Supplemental 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*=.44, *p*=.01, 95% CI = .18, .66 ; Fig. 1d & Table S4). We then tested whether the online simulation subscale also predicted neural responses to prosocial PEs. Online simulation was also positively associated with prosocial (compared to self) PE responses in the sgACC (*r*=.39, *p*=.03, 95% CI = .13, .60; Fig. 3d), with those higher in online simulation showing greater sgACC specificity to prosocial relative to self PEs. Together these findings suggest behavioural and neural links between empathy and prosocial behaviour.

Discussion

Reinforcement learning theory (RLT) has provided important insights into how we learn about rewarding information for ourselves. Here we used the RLT

framework to characterise 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 behavioural responses predicted by trait empathy.

Our findings advance theoretical accounts of the neural basis of social behaviour by finding evidence for both ‘common coding’ and ‘socially specific’ regions of the brain underpinning prosocial behaviour (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 behaviour. 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 VS differs from the sgACC, which shows specificity for signalling PEs when learning to benefit another. Thus while both regions may

291 play a role in driving prosocial behaviour, the function of ventral striatum may
292 be more domain general than that of sgACC, which we speculate may
293 compute a prediction error specifically for outcomes delivered to others.
294
295 Intriguingly, the sgACC region we identified as responding exclusively to
296 prosocial PEs overlapped with a septal-anterior hypothalamic area that is part
297 of the basal forebrain (43). Recent studies of sgACC function have found
298 signals in this region relevant for social cognition and behaviour including
299 credit assignment (44), prosocial and moral behaviour (6, 16, 45, 46), the
300 experience of positive affect (47), trust (48), social emotion (49) and vicarious
301 reward (50, 51). Moreover, there is evidence that the sgACC may signal PEs
302 for self-reward, but only when learning occurs at a specific level of abstraction
303 beyond basic stimulus-response association (52). One possible explanation
304 for this convergence is that similar abstract learning mechanisms may drive
305 how we learn to benefit others, and that both are underpinned by RLT
306 principles. Further studies could compare tasks that manipulate both the
307 social and the hierarchical context of a learning environment to directly test
308 the parallels between these types of learning. In addition, future research
309 should aim to dissociate the functions of the heterogeneous cortical and
310 subcortical structures of the sgACC complex, using high-resolution fMRI, in
311 order to understand which of these sub-regions contribute to social cognition
312 and prosocial learning.
313

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 behaviour (15, 29, 37). Whilst it is difficult to interpret a null finding, as there can be a number of reasons that a particular neural response was not observed, we note that in this experiment participants were performing their task in a different “reference frame” (see 53, 54 for recent discussions of the role of references fromes in studies of social cognition) as compared with other studies on empathy/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 the first evidence of a mechanism linking variability in empathy to variability in prosocial behaviour, 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 behaviour 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 behaviours. Enhanced signalling of PEs and faster learning when benefitting others also provides a potential explanation for previous reports of individuals higher in empathy being more motivated to behave in a prosocial manner compared to those lower in empathy (21). Moreover, our results support an emerging view that PE signals may be crucial for learning how to interact, and empathise, 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 behaviour.

An influential theory within the literature on empathy is that we empathise 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 ACC/MCC (reviewed in(29)), 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). While 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 (51) for a meta-analysis and (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 for 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 behaviour (e.g. 66) our sample in this study was comprised only of males. Future studies would benefit from also examining prosocial learning in females.

Prosocial behaviours 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 new insights into why these disorders are associated with atypical prosocial behaviour and empathy. Taken together, these findings reveal a novel computational link between prosocial learning and empathy in humans and

therefore pave the way to characterise atypical prosocial interactions in those with disorders of social cognition and behaviour.

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, non-normal or non-corrected to normal vision, non-native 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 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*. 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 centre 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 word '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 Supplemental Experimental Procedures and Fig 1a). Participants practiced one block (16 trials) of the task in a separate session ~7 days before the scanning session to familiarise 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 prior to 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 days later) was attended by both the experimental participant and the confederate. The participant and confederate were taken together to the MRI centre and filled in consent forms together in the same room. The confederate was then led into a behavioural testing room and instructed to complete questionnaires, with the experimental participant within earshot 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 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 standardised questions completed after the scan confirmed that no participant had become suspicious about the deception during the experiment.

Computational modelling of behavioural data

Learning behaviour in the self, prosocial and no-one conditions was modelled using a reinforcement learning (RL) algorithm (Sutton and Barto, 1998), which has been extensively used to examine the behavioural 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:

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

The learning rate '*α*' controls the extent to which the current expected value is updated by new information. Consequently, a low learning rate will minimise 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:

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

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 behaviour). A low '*β*' would lead to consistent behaviour, where the

500 action with the higher expected value is invariably selected on each trial. In
501 the full model, separate α and β parameters were estimated for each of the
502 self, prosocial and no-one conditions, as this provided the most parsimonious
503 explanation for the behavioural data (see supplemental experimental
504 procedures for details of model fitting and model comparison, and Fig. S2).

505 **Statistical analysis of behavioural data**

506 Analyses of behavioural data were performed in SPSS 22 (Armonk, New
507 York: IBM Corp). We examined differences between conditions in the learning
508 rate and temperature parameters at the group level using separate repeated
509 measures analyses of variance (ANOVAs), with three levels (self, prosocial
510 and no-one). Separate ANOVAs were conducted as the learning rate and
511 temperature parameters represent different units of measurement. We
512 examined bivariate associations between the prosocial-self difference in
513 learning rate and temperature and empathy questionnaire sub-scales using
514 the Pearson correlation coefficient. Effect sizes (Cohen's d) were calculated
515 by dividing the mean difference between conditions by the standard deviation
516 of the difference(71). Confidence intervals for correlation coefficients were
517 estimated using SPSS 22 Bootstrap procedure.

518 **Fmri acquisition and analysis**

519 A Siemens Avanto 1.5-T MRI scanner was used to acquire a 5.5-minute 3-
520 dimensional T1-weighted structural scan and 424 multislice T2*-weighted
521 echo planar volumes with blood oxygenation-level–dependent (BOLD)
522 contrast. The structural scan was acquired using a magnetization prepared

523 rapid gradient echo (MPRAGE) sequence with 176 slices; slice thickness=1
524 mm; gap between slices=0.5 mm; TR=2730 ms; TE=3.57 ms; field of
525 view=256 mm x 256mm2; matrix size=256 x 256; voxel size=1×1×1 mm
526 resolution. The functional imaging sequence was acquired in an ascending
527 manner, at an oblique angle (≈30˚) to the AC-PC line to decrease the impact
528 of susceptibility artefact in the orbitofrontal cortex (72) and had the following
529 acquisition parameters: 424 volumes, 1 mm gap; echo time=50 ms; repetition
530 time=2975 ms; flip angle=90°; field of view=192 mm; matrix size=64x64.
531
532 Imaging data were analysed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Data
533 preprocessing followed a standard sequence: the first 4 volumes and last
534 volume were discarded. Images were then realigned and co-registered to the
535 participant's own anatomical image. The anatomical image was processed
536 using a unified segmentation procedure combining segmentation, bias
537 correction, and spatial normalization to the MNI template using the New
538 Segment procedure (73); the same normalization parameters were then used
539 to normalize the EPI images. The voxel size was resampled to 1.5 x 1.5 x
540 1.5mm. Lastly, a Gaussian kernel of 8 mm FWHM was applied to spatially
541 smooth the images. Before the study, example first-level design matrices
542 were checked to ensure that estimable GLMs could be performed with
543 independence between the parametric regressors (chosen value and PE in
544 the three conditions), with correlations coefficients of $r<0.25$. This allowed us
545 to look at PE-related responses independent of chosen value.
546

547 Eight event-types were used to construct regressors in which event timings
548 were convolved with SPM's canonical haemodynamic response function. The
549 three conditions at the time of the cues and three conditions at the time of the
550 outcome were modelled as separate regressors using stick functions. Each of
551 these regressors was associated with a parametric modulator taken from the
552 computational model. At the time of the cue this was the chosen value, and at
553 the time of the outcome, the PE (calculated as in Equation 1 above). The PEs
554 were estimated using each subjects own alpha and beta from each condition.
555 The instruction cue at the beginning of each block was also modelled in a
556 single regressor as a stick function. In some participants, an eighth regressor
557 modelled all missed trials, on which participants did not select one of the two
558 symbols in the response window. For those participants where there was
559 visible head motion in a particular scan (scans with >1 mm or 1 degree
560 movement relative to the next were examined visually) an extra regressor was
561 included corresponding to each scan. These images were removed and
562 replaced with an image created by interpolating the two adjacent images in
563 order to prevent distortion of the between-subjects mask (4 participants, less
564 than 1% of total time series). Six head motion parameters modelled the
565 residual effects of head motion as covariates of no interest. Data were high-
566 pass filtered at 128 s to remove low-frequency drifts, and the statistical model
567 included an AR(1) autoregressive function to account for autocorrelations
568 intrinsic to the fMRI time-series. Our primary analysis focused on the PEs at
569 outcome (for response to chosen value see table S1).

570 Contrast images from the first level were input into a second-level flexible-
571 factorial design with three levels (self PE, prosocial PE, no-one PE). Main
572 effects are reported at p<.05, family-wise error (FWE) corrected at the voxel
573 level across the whole brain or p<.05 small volume corrected (SVC) at the
574 voxel level in regions where we had a strong *a priori* hypothesis (see below).
575

576 **ROI selection and fMRI contrasts**

577 The *a priori* regions-of interest (ROIs) were defined anatomically using masks
578 taken from an appropriate atlas (bilateral VS, sgACC, dACC, bilateral DLPFC;
579 toolboxes: Harvard-Oxford Atlas, regions 46v and 9 (74), Anatomy Toolbox
580 regions s24 and 25 (75), region 24 from (76), see Fig. S3 and supplemental
581 experimental procedures for further details of ROI selection). We additionally
582 applied a False Discovery Rate correction (FDR) (77), rather than Bonferroni
583 correction (which may be overly conservative given that our ROIs were not
584 entirely independent from one another as they are functionally and
585 anatomically connected (9,10,11))), for the number of ROIs. All ROI
586 comparisons remained significant (p<.05) when controlling for the number of
587 comparisons using FDR.
588

589 **Acknowledgements**

590 This work was supported by a studentship from the Medical Research Council
591 to Patricia L. Lockwood and funding from the Birkbeck–University College
592 London Centre for Neuroimaging. Matthew A. J. Apps was supported by an
593 anniversary future leaders fellowship from the BBSRC (award number

BB/M013596/1). Essi Viding is a Royal Society Wolfson Research Merit Award holder. The authors also thank Mr. Lewis Pollock and Mr. Jonathan Blott, for acting as the confederates and assistance with data collection.

References

1. Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425(6960):785–91.

2. Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction* (MIT press, Cambridge, Massachusetts).

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):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* 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 U S A* 107(32):14431–14436.

19. Suzuki S, et al. (2012) Learning to simulate others' decisions. *Neuron* 74(6):1125–37.

20. Crockett MJ, Kurth-nelson Z, Siegel JZ, Dayan P, Dolan RJ (2015) Harm to others outweighs harm to self in moral decision making. *Proc Natl Acad Sci* 112(4):201424572.

21. Silk JB, House BR (2011) Evolutionary foundations of human prosocial sentiments. *Proc Natl Acad Sci* 108:10910–10917.

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* 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):435–41.

26. Hoffman ML (2008) Empathy and prosocial behavior. *Handb Emot* 3:440–455.

27. Singer, 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* 112(50):15516–15521.

662 29. Lamm C, Decety J, Singer T (2011) Meta-analytic evidence for common and
663 distinct neural networks associated with directly experienced pain and
664 empathy for pain. *NeuroImage* 54(3):2492–2502.

665 30. Lockwood PL (2016) The anatomy of empathy: typical variability and
666 disorders of vicarious experience. *Behav Brain Res*.

667 31. Lockwood PL, Bird G, Bridge M, Viding E (2013) Dissecting empathy: high
668 levels of psychopathic and autistic traits are characterized by difficulties in
669 different social information processing domains. *Front Hum Neurosci* 7,1–6.

670 32. Blair (2005) Responding to the emotions of others: dissociating forms of
671 empathy through the study of typical and psychiatric populations. *Conscious*
672 *Cogn* 14(4):698–718.

673 33. Lockwood PL, Seara-Cardoso A, Viding E (2014) Emotion regulation
674 moderates the association between empathy and prosocial behavior. *PloS*
675 *One* 9(5):e96555.

676 34. Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of
677 fictive learning signals in a sequential investment task. *Proc Natl Acad Sci U*
678 *S A* 104:9493–9498.

679 35. Hayden BY, Pearson JM, Platt ML (2009) Fictive reward signals in the
680 anterior cingulate cortex. *Science* 324(5929):948–950.

681 36. Gu X, Kirk U, Lohrenz TM, Read Montague P (2013) Cognitive strategies
682 regulate fictive, but not reward prediction error signals in a sequential
683 investment task. *Hum Brain Mapp*

684 37. Apps MAJ, Rushworth MFS, Chang SWC (2016) The Anterior Cingulate Gyrus
685 and Social Cognition: Tracking the Motivation of Others. *Neuron* 90(4):692–
686 707.

687 38. Nichols T, Brett M, Andersson J, Wager T, Poline J-B (2005) Valid conjunction
688 inference with the minimum statistic. *NeuroImage* 25(3):653–660.

689 39. Báez-Mendoza R, Harris CJ, Schultz W (2013) Activity of striatal neurons
690 reflects social action and own reward. *Proc Natl Acad Sci U S A* 110(41).

691 40. Harbaugh WT, Mayr U, Burghart DR (2007) Neural responses to taxation and
692 voluntary giving reveal motives for charitable donations. *Science*
693 316(5831):1622–1625.

694 41. Reniers RLEP, Corcoran R, Drake R, Shryane NM, Völlm B a (2011) The QCAE:
695 a Questionnaire of Cognitive and Affective Empathy. *J Pers Assess* 93(1):84–
696 95.

697 42. Rilling JK, King-Casas B, Sanfey AG (2008) The neurobiology of social
698 decision-making. *Curr Opin Neurobiol* 18(2):159–65.

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 U S A* 111(14):5391–5396.

48. Krueger F, et al. (2007) Neural correlates of trust. *Proc Natl Acad Sci U S A* 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):900–900.

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).

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):257–67.

57. Eisenberg N (2000) Emotion, regulation, and moral development. *Annu Rev Psychol* 51(1):665–697.

58. Eisenberg N, Miller P a (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 U S A* 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* 12:165–174.

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 U S A* 112(41):E5638–5646.

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):784–90.

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 CB* 23(10):901–5.

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. 2nd edn. Hillsdale, New Jersey: L.

72. Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19(2):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*:289–300.

78. Hogan R (1969) Development of an empathy scale. *J Consult Clin Psychol* 33(3):307.

79. Davis M (1983) Measuring individual differences I empathy: Evidence for a multidimensional approach. *J Pers Soc Psychol* 44:113–126.

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. *Decis Mak Affect Learn Atten Perform XXIII* 23:1.

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):1043–8.

85. Waytz A, Zaki J, Mitchell JP (2012) Response of dorsomedial prefrontal cortex predicts altruistic behavior. *J Neurosci* 32(22):7646–50.