nature human behaviour

LETTERS

PUBLISHED: 30 JUNE 2017 | VOLUME: 1 | ARTICLE NUMBER: 0131

Prosocial apathy for helping others when effort is required

Patricia L. Lockwood[1]*, Mathilde Hamonet[1], Samuel H. Zhang[1], Anya Ratnavel[1], Florentine U. Salmony[1], Masud Husain[1,2†] and Matthew A. J. Apps[1†]

Prosocial acts—those that are costly to ourselves but benefit others—are a central component of human coexistence[1–3]. While the financial and moral costs of prosocial behaviours are well understood[4–6], everyday prosocial acts do not typically come at such costs. Instead, they require effort. Here, using computational modelling of an effort-based task, we show that people are prosocially apathetic. They are less willing to choose to initiate highly effortful acts that benefit others compared with those benefitting themselves. Moreover, even when choosing to initiate effortful prosocial acts, people exhibit superficiality, exerting less force into the actions that benefit others than those that benefit themselves. These findings were replicated, and were present whether the other person was anonymous or not, and when choices were made to earn rewards or avoid losses. Importantly, the least prosocially motivated people had higher subclinical levels of psychopathy and social apathy. Thus, although people sometimes 'help out', they are less willing to benefit others and are sometimes 'superficially prosocial', which may characterize everyday prosociality and its disruption in social disorders.

Our willingness to be prosocial—to perform acts that benefit others—is a central component of human social and moral behaviour. Prosocial behaviours are argued to be a key driver of bonding within groups and to facilitate social cohesion[1–5,7]. Acts of prosociality are often reduced in those with criminal levels of psychopathy, in healthy individuals lower in empathy and after brain damage[4,7–13]. Prosociality has commonly been investigated in terms of people's willingness to benefit others when these acts come at a personal cost[2,4,6,14]. Typically, these costs have been financial, such as when donating to other individuals or charities[6], or moral, such as when deciding whether to profit from others' harm[4,5]. Yet, in day-to-day life, financial and moral costs are often negligible. Instead, prosocial acts require motivation to exert effort. Here, we use computational modelling to precisely characterize how effort costs influence motivation to benefit ourselves and other people.

Theoretical accounts of motivation highlight that there are at least two key components that define willingness to exert effort. First, before an action is performed, its costs and benefits are evaluated[15–20]. When the rewards are high, or the effort costs are low, the value of a behaviour is high and people are motivated to choose to perform that action. However, when the rewards are low or the effort required is high, the subjective value of the reward is reduced. Recent accounts suggest that such 'devaluation' or 'discounting' of rewards by effort can be quantified using computational modelling approaches[17,21–27]. Using such models, it is possible to precisely characterize the form that the devaluation takes, and the extent to which each individual devalues rewards by effort costs. Second, once chosen, actions must be energized to such a degree that they are sufficiently forceful[28–30] to obtain the desired outcome. While this framework has been used extensively to provide quantifiable measures of motivation to reward oneself, surprisingly, such models have not been applied to understand how we motivate actions that benefit others.

Are people prosocial or are they more selfish when it comes to putting in effort for others? To our knowledge, there are no studies that have quantified either the willingness or the force exerted when people have to put in physical effort for others' benefit. However, the results of studies using economic games are often consistent with a largely selfish view of human behaviour[6,31]. While individuals are willing to cooperate and share financial resources with others, they still value their own financial gains above those of other people[6,31,32]. In contrast, studies of moral decision-making have painted a more positive picture of human behaviour. When individuals trade off a financial cost to themselves against a number of painful shocks delivered to them or to an anonymous stranger, they appear to be 'hyperaltruistic'[4,5]. They are willing to pay more money to avoid painful shocks being delivered to another person than they are to themselves. Thus, if prosocial effort costs are like moral choices, people might be more willing to exert effort to benefit others over themselves and therefore should be hypermotivated to benefit others. In contrast, if effort costs are treated similarly to financial costs, people may be prosocially apathetic and far less willing to work to benefit others.

Levels of prosociality and motivation can vary considerably between people. Variability in apathy, the willingness to exert goal-directed behaviours, is thought to be key to understanding variability in motivation[33–38]. Multiple lines of evidence suggest that, as well as being common in several neurological and psychiatric disorders, apathy is also present to varying degrees in the general population[17,34–38]. This dimensional approach to understanding psychiatric and clinical disorders is consistent with the research domain criteria approach to uncovering the mechanisms of psychopathology in health and disease[39]. Notably, recent research has shown that people with high levels of the behavioural apathy show a reduced willingness to exert effort to benefit themselves[17,40], whereas those who are more socially apathetic report less engagement in social behaviours[34]. Here, we use the term 'prosocial apathy' to refer to reduced motivation to put in effort for others' benefit relative to one's own benefit.

Outside the domain of motivation, moral and financial prosocial choices often vary considerably between individuals[4,7,41,42]. For example, levels of psychopathic traits are negatively correlated with

[1]Department of Experimental Psychology, University of Oxford, Oxford OX1 3PH, UK. [2]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK. †These authors contributed equally to this work. *e-mail: patricia.lockwood@psy.ox.ac.uk

trait measures of prosociality[43] and self-reported enjoyment of prosocial interactions[44]. However, the significance of such trait-level associations for behaviour is unclear. For these reasons, we also tested whether those high in psychopathic traits or social apathy are less willing to exert effort to benefit others.

To examine whether people are hypermotivated or apathetic when choosing or energizing acts that benefit other individuals, we designed an effort-based decision-making task based on rodent, primate and human research[15,45,46]. Across two studies, participants decided whether or not they were willing to exert an offered level of effort for rewards, and subsequently executed the chosen effort (Fig. 1). To precisely quantify effort, we used different levels of grip force, with levels of effort scaled as a percentage of each individual's maximum voluntary contraction (MVC). Participants made choices between a baseline of no effort (0% of the MVC), a low-reward option (1 credit) and a variable 'offer', which was higher in reward (2–10 credits) but also higher in effort (30–70%). We also examined the choices made and force exerted when participants avoided losing rewards (2–10 credits) if they chose the 'offer'.

Once they had made their choice, participants were required to squeeze a handle to the required degree of force. Crucially, in half of the trials they made choices to earn payment for themselves (the 'self' condition) and in the other half of the trials they had the option of putting in effort to reward another person (the 'other' condition). Using computational modelling approaches, we quantified how willing people were to choose to exert effort to benefit themselves versus others, and used a mathematical function to characterize the discounting of reward. We also examined how energized people's actions were by measuring the force exerted at each level of effort. We could therefore test our key questions of whether effort is costlier when benefitting others or ourselves and whether trait levels of apathy and psychopathy explain individual differences in prosocial apathy.

Using analyses of variance (ANOVAs) with regard to the proportion of acceptance of the 'variable offer', we first examined whether willingness to choose to put in effort to obtain rewards differs by the agency (self or anonymous other) of the beneficiary. In study 1, there was a significant agent × effort interaction ($F_{4,188}=11.21$, $P<0.001$; Fig. 2a), as well as main effects of effort and agent (effort: $F_{4,188}=128.01$, $P<0.001$; agent: $F_{1,47}=54.93$, $P<0.001$). Thus, as the offer increased in effort, people were less willing to choose it. However, this effect was further exaggerated when another person was the beneficiary, with a greater reduction in willingness to exert higher levels of effort for the other person. Therefore, when the costs are low, people are willing to choose to benefit both themselves and others. However, when the costs are high, people are less willing to put in effort to help others than themselves. In investigating whether people are less incentivized by rewards for others, we found a significant main effect of reward ($F_{4,188}=52.05$, $P<0.001$) and agent ($F_{1,47}=54.93$, $P<0.001$) and an agent × reward interaction ($F_{4,188}=6.34$, $P<0.001$). Thus, as the rewards increased, people were more incentivized for self versus other (Fig. 2a).

In study 2, we aimed to replicate the findings of study 1 and also to examine whether prosocial motivation could be increased by changing the nature of the other person and the rewarding incentives. Previous studies suggest that prosocial behaviour can be affected by whether it is public or private[43,47,48]. For example, the introduction of an audience increases levels of prosociality[43,47,48], while anonymity reduces it[49]. Therefore, we examined whether or not face-to-face contact with the person who would receive the outcomes increases the effort exerted for others. In addition, we included an avoid loss condition, as studies of moral harm have often used negative valence to the benefit on offer (for example, losing money to avoid harm[4,5]). Moreover, previous research has suggested that self-motivation can be increased when the aim is to avoid the loss of rewards, rather than to earn them[50,51]. Thus, study 2 allowed us to examine three questions. (1) Do the effects of effort replicate when minimal contact is made with another person whom the participant is led to believe is the 'other' person? (2) Is there a loss-aversion effect that results in higher levels of prosocial motivation? (3) Is there still an increased willingness to exert higher levels of effort when avoiding losses?.

First, examining the win session only (Fig. 2b), we replicated the findings of study 1 for effort (agent × effort interaction: $F_{4,176}=18.02$, $P<0.001$), but not for reward (agent × reward interaction: $F_{4,176}=2.22$, $P=0.069$). This suggests that even when participants met the person

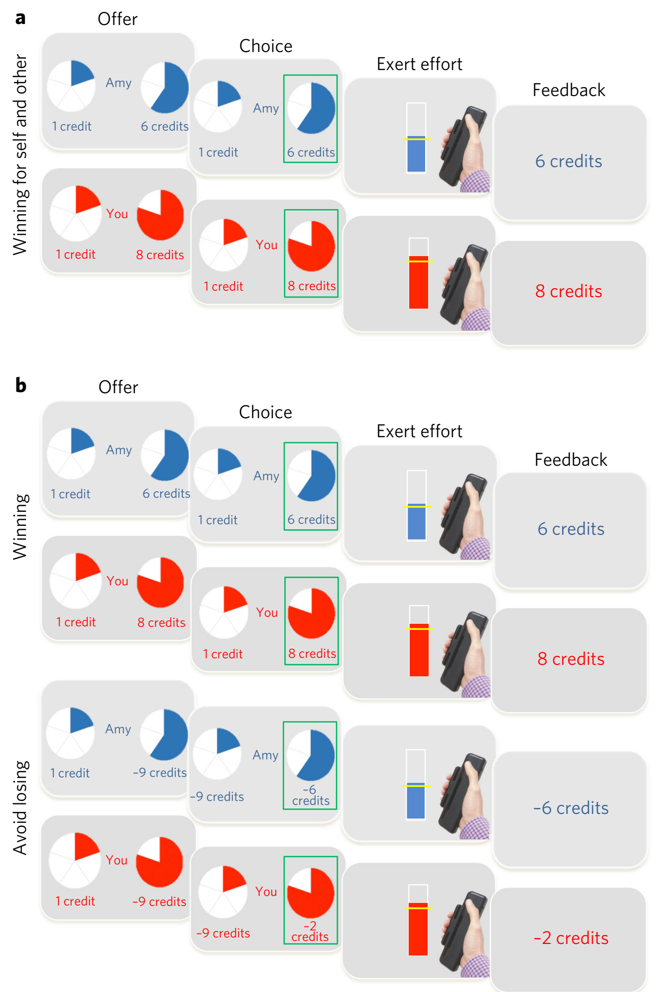

Figure 1 | Prosocial motivation measure for self versus other. a, In study 1, participants were assigned to one of two different roles completely anonymously. They made choices between a fixed low-effort, low-reward option (shown on the left) and a variable higher-effort (30, 40, 50, 60 or 70% of their MVC) but higher-reward option (2–10 credits). Once a choice was made the chosen force was exerted on a handheld dynamometer in the participant's dominant hand (the exert effort). Only if the required level was reached for 1 s out of a 3 s window would the offered reward be obtained; otherwise, 0 credits would be delivered. In half of the trials, the credits on offer were for themselves (self, red). In the other half, the credits were for the other participant (other, blue). **b**, In study 2, participants met face to face with a confederate (whom they believed to be another participant taking part in the same study). Participants took part in two sessions in counterbalanced order. In one session, they made choices to win money for the other person and themselves, while in the other session, they made choices to avoid losing money for the other person and themselves. For both studies, the name of the other participant (in this example 'Amy') was displayed during the trials when choosing for the other person.

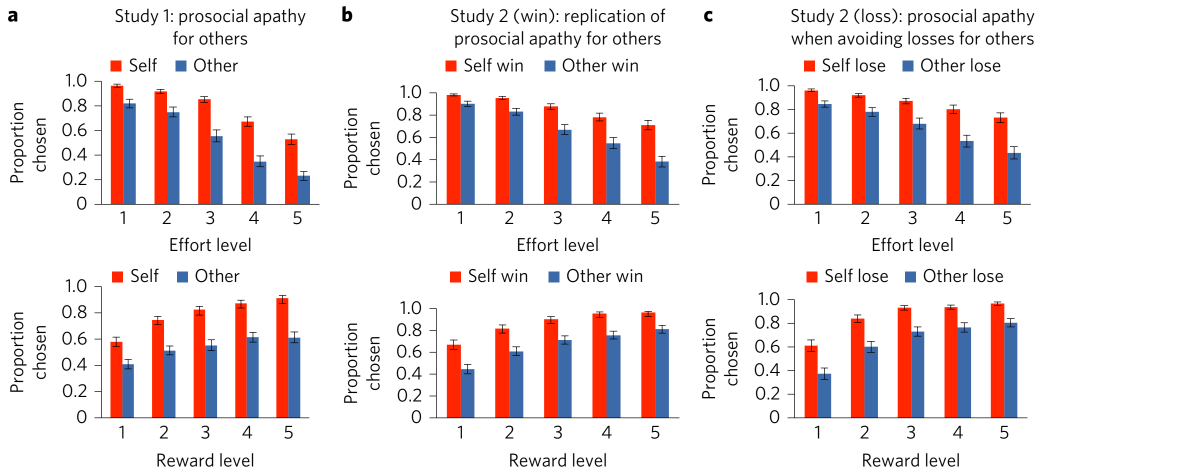

Figure 2 | Prosocial apathy when deciding to exert effort to reward others. a, Proportion of higher-effort and -reward options chosen over the baseline option (lower effort and reward) plotted against effort (top panel) and reward levels (bottom panels) in study 1 ($n=48$). Effort levels 1–5 correspond to 30, 40, 50, 60 and 70% of a participant's MVC. Reward levels 1–5 correspond to 2, 4, 6, 8 and 10 credits. Participants chose the higher-effort and -reward option more frequently for self (red bars) than other (blue bars), with this difference increasing with effort level. They were also less reward-sensitive for other people than themselves. **b**, In study 2 ($n=45$), participants again chose the higher-effort and -reward option more frequently for self than other, with this effect increasing with effort level (top panel). The participants were also less reward-sensitive for others compared with self (bottom panel). **c**, When making choices to avoid losing for self or other in study 2, participants chose the higher-effort and lower-loss option more frequently for self than other (top) and were less loss-sensitive for self than other. In all panels, the error bars show s.e.m.

who would be receiving the outcomes, their prosocial motivation did not increase in terms of the effort they were willing to expend. Moreover, in this analysis, effort levels were collapsed across reward and the variable offer was always associated with an effort cost. Thus, the interaction between agent and effort across both studies suggests that it is a reduced willingness to exert effort for others that consistently drives reduced choices of 'offers' for others and not only a reduced sensitivity to the rewards that others will receive.

Analysis of behaviour from the losing session (Fig. 2c) showed that framing the task as avoiding losses for self and other did not change motivation. Again, participants were less willing to choose to put in effort to avoid losing money for the other person compared with themselves at increasing levels of effort (agent × effort interaction: $F_{4,176}=13.59$, $P<0.001$). They also demonstrated significantly lower loss sensitivity for others compared with themselves (agent × reward interaction: $F_{4,176}=3.10$, $P<0.017$). These findings were also replicated when including winning and losing in the same ANOVA model (see Supplementary Results). Comparing choice behaviour between studies 1 and 2, we did not find a significant sample × agent × effort interaction ($F_{4,364}=2.13$, $P=0.08$).

Next, to precisely quantify the subjective influence of effort on rewards for self and other, we created a range of computational models that characterized how rewards were being devalued by the amount of effort. This approach allowed us to quantify people's motivation to put in effort to reward others using a single parameter that meaningfully characterized how motivation was influenced by the balance between effort and reward. The model also allowed us to use this parameter to test for associations between discounting rates and self-report measures, which can provide more interpretable results than using individual difference measures as covariates in ANOVAs that contain multiple factors with multiple levels.

Models were fitted to behavioural data using the 'softmax' function (see Methods and Supplementary Methods). Each model contained idiosyncratic parameters characterizing the degree to which a reward was devalued by effort (K), and noise parameters characterizing the stochasticity of choices (β). Two features were varied to create the model space. First, we varied the mathematical function that characterizes the form of the discounting (that is, whether rewards are devalued linearly, hyperbolically or parabolically by physical effort[22,52]). Second, we compared models that tested whether people devalue or discount rewards by effort to the same degree on self and other trials, or instead used separate discounting rates. We therefore created two classes of model that had either the same parameters to characterize discounting (K) on self and other trials (models 1–6) or separate ones (models 7–12; Fig. 3). Within these models, we tested a further two classes of model that characterized whether separate parameters for levels of noise (β, 'softmax') (models 4–6 and 10–12) or single parameters for noise (models 1–3 and 7–9) best explained behaviour.

The winning model in study 1 was a parabolic model in which separate parameters characterized the devaluation of rewards for self and other trials, but had a single noise parameter (model 7; Fig. 3a). This model was best able to explain participants' choices when correcting for the number of parameters in the models. The Bayesian information criterion (BIC) scores of this model were very close to a parabolic model that had separate discounting parameters for self and other but also separate parameters for noise (BIC two discount, one noise parameter: 5591.69; BIC separate K_{self} and K_{other}, one noise parameter: 5574.28; Fig. 3a,d). However, the choice behaviour of most participants (32 out of 48) was explained better by the model with a single noise parameter. Thus, devaluation of rewards by effort for self and other can be characterized using equation (1) as follows:

$$\mathrm{SV}=R-KE^2$$

$$K=\begin{cases} K_{\text{self}} & \text{if self trial} \\ K_{\text{other}} & \text{if other trial} \end{cases}$$

In this model, SV is the subjective net-value of a variable offer of given effort (E) and reward (R). The extent to which rewards are subjectively discounted is dependent on the discount parameter (K), which is different on self and other trials. A high K value indicates that participants are discounting rewards by the effort to a higher degree.

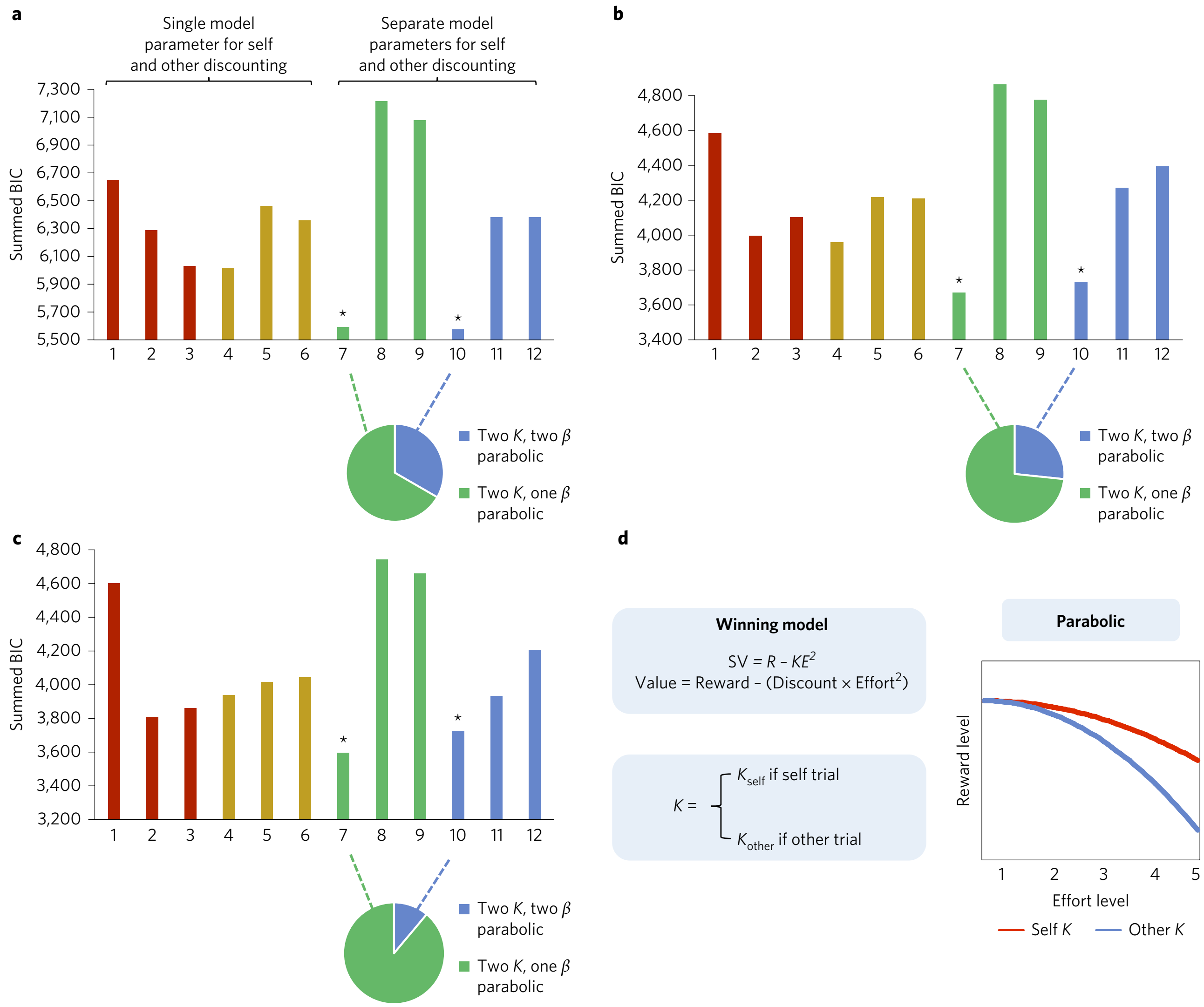

Figure 3 | A model comparison robustly shows across two studies that a model with separate discount parameters for self and other best explains behaviour. **a**, Model comparison results from study 1 with the *x* axis depicting the model number and the *y* axis depicting the sum of the BIC score. Models with parabolic, linear and hyperbolic discounting functions with either single (models 1–6) or separate discount (K) parameters (models 7–12) for self and other and/or single (red and green) or separate noise (β) parameters (yellow and blue) for self and other. A parabolic model with separate parameters for self and other discounting, but a single noise parameter, best explained behaviour in the majority of subjects (model 7 in all studies), which was determined by this model having the lowest summed BIC score in combination with explaining behaviour in most of the participants. The pie chart shows the proportion of participants that the winning model explains behaviour for (green) compared with the same model with separate noise parameters (blue). **b**, Model comparison results from study 2 (win) replicated the winning model from study 1. **c**, Model comparison results from study 2-loss also replicated the winning model from study 1 and study 2 (win). In panels **a**–**c** the asterisks show the two models with the lowest BIC scores. **d**, Mathematical formula for the winning parabolic model. The graph shows the discount parameters for the median K scores for self and other at reward level 4 for illustrative purposes.

The winning model (model 7) was replicated with the winning session data from study 2 (BIC = 3671.72; Fig. 3b) and in most participants (33 out of 45 people) compared with the separate noise model (model 10). Furthermore, the same model was again replicated as the best fitting model to the avoiding loss session in study 2 (BIC = 3597.85; Fig. 3c), and in most participants (40 out of 45 participants). We also ran model comparisons with the combined winning and losing data from study 2 to test whether separate win and loss parameters were also needed to explain participants' behaviour. This additional family of models also showed that separate parameters were needed to explain self and other discounting, but also that separate parameters were needed to explain the discounting of rewards and losses (see Supplementary Fig. 1).

To support the univariate analyses outlined at the beginning of the results, we also compared the discounting parameters from the model for self and other. The discounting parameters were higher for other compared with self in both studies and also for the winning data (one participant was excluded from study 2 (win) due to a very poor model fit resulting in a K value greater than 3 s.d. above the mean; Wilcoxon signed-rank test: study 1 $K_{other} > K_{self}$, critical z value, $z = 5.60$, $P < 0.001$; replicated in study 2 (win): $z = 5.01$, $P < 0.001$, and in study 2 (loss): $z = 4.41$, $P < 0.001$). However, the parameters were also correlated, albeit only weakly, in study 1, suggesting that there is some baseline level of motivation that affects discounting for both self and other (study 1: Spearman rank correlation coefficient, $r = 0.26$, $P = 0.08$ (marginal); study 2 (win): $r = 0.39$, $P = 0.008$; study 2 (loss): $r = 0.35$, $P = 0.017$). We also compared discount rates for self and other across the two samples. Intriguingly, we found that participants discounted rewards by effort for others to a lesser degree in study 2 compared with study 1 ($z = -2.83$,

$P=0.005$), but there was no greater discounting of rewards for self ($z=-1.49$, $P=0.14$). This suggests that meeting another person face to face did, in fact, increase prosocial motivation, but not to the same level as self-motivation. Using ANOVAs, we did not find a significant interaction between agent, effort and sample, suggesting that the model may have provided a more sensitive measure of the balance between reward and effort on participant choices. In summary, these findings support the notion that people are less willing to choose to put in effort to obtain rewards for others than they are to obtain rewards for themselves. That is, they are more apathetic when being prosocial than when benefiting themselves.

Existing studies suggest that there may be substantial variability in the extent to which people are prosocial[4,7,13,41,42]. Consequently, we investigated whether individual differences in self-reported motivation could explain some of the variance in behaviour. To ensure adequate power for these analyses, we pooled together self-reported and behavioural data from study 1 and the winning session of study 2 ($n=92$; one participant was excluded for an extremely poor model fit providing discount (K) values greater than 3 s.d. above the mean). Our first set of questions related to whether self-reported levels of motivation would correlate with willingness to put in effort on the task. We used the apathy–motivation index (AMI), a self-reported measure of levels of motivation with strong psychometric properties[34] (see Supplementary Methods).

The AMI divides apathy into three different domains: behavioural activation (self) motivation, emotional sensitivity and social (other) motivation. Our main hypothesis was that individuals who were less willing to put in effort on the 'other' trials would show higher levels of social apathy on the AMI. Consistent with this view, there was a significant positive correlation between levels of social apathy on the AMI and 'other' K values from the model ($r=0.22$, $P=0.03$). Previous studies have shown that willingness to put in effort to obtain rewards for ourselves correlates with behavioural activation[17,40]. In line with these studies, we found that people who had higher discounting for 'self' (higher 'self' K values) also had higher levels of behavioural activation apathy on the AMI ($r=0.34$, $P=0.002$). In contrast, there was no correlation between the 'other' K parameter and the behavioural activation subscale. Thus, variation in willingness to exert effort for others in the task appears to relate specifically to everyday self-reported social apathy, whereas willingness to put in effort to benefit ourselves relates to behavioural apathy.

Another key individual difference that may affect willingness to help others is our level of psychopathic traits. Psychopathy is a disorder characterized by a constellation of cognitive and behavioural atypicalities, including callousness, shallow affect, lack of guilt, antisocial behaviour and impulsivity[53–56]. In the typical population, psychopathic traits can be reliably measured, with these traits existing on a continuum and often mirroring associations with related constructs that are found in clinical samples[57–59]. Individuals high in psychopathic traits have consistently been found to be less willing to engage in prosocial behaviours[43,44]. In contrast, they may also engage in behaviours associated with impulsivity and antisociality[53,60,61]. Psychopathic traits are therefore prototypically linked to selfish behaviour and a reduced willingness to help others relative to oneself. On this basis, we predicted that psychopathic traits would be associated with a lower willingness to put in effort for rewards for others relative to oneself. As hypothesized, the difference in the discounting parameters between motivation for other and self ($K_{other} - K_{self}$) was correlated with the total psychopathy scores ($r=0.24$, $P=0.02$; K_{other}: $r=0.05$, $P=0.66$; K_{self}: $r=-0.15$, $P=0.15$). Follow-up analyses to determine if this effect was associated with specific aspects of psychopathy showed that there was a correlation with affective–interpersonal (core) psychopathic traits ($r=0.23$, $P=0.03$; K_{other}: $r=0.10$, $P=0.37$; K_{self}: $r=-0.10$, $P=0.35$), but not lifestyle antisocial traits ($r=-0.09$, $P=0.40$; K_{other}: $r=-0.12$, $P=0.28$; K_{self}: $r=-0.19$, $P=0.08$; see Supplementary Methods). Thus, those higher in psychopathic traits devalue rewards by effort to a greater degree when another person is the beneficiary, relative to when they are putting in effort to help themselves. These findings suggest that there might be benefits of using our paradigm to understand prosocial deficits in clinical disorders.

To assess whether participants energized their actions to the same degree when choosing to exert effort for others, we ran linear regressions to predict the area under the curve (see Methods) of the force applied in each trial. Only trials for which participants chose the variable offer were included in this analysis (since these were the only trials for which they chose to put in effort). In study 1, we found an agent × effort interaction, indicating that as the amount of effort required increased there was a greater difference in the total amount of force exerted (average $\beta=-0.04$, agent × effort interaction: $P<0.001$). Reduced levels of force exerted at the higher levels of effort for others, relative to self, drove this effect (Fig. 4a). There were also main effects of effort, agent and reward (effort: $\beta=0.86$, $P<0.001$; agent: $\beta=-0.05$, $P<0.001$; reward: $\beta=0.04$, $P<0.001$). Thus, in study 1, the force exerted increased as the effort required increased, but this happened to a lesser degree when effort was exerted to benefit another person.

These findings were replicated in study 2 for winning (agent: $\beta=-0.08$, $P<0.001$; effort: $\beta=0.85$, $P<0.001$; reward: $\beta=0.05$, $P<0.001$; agent × effort interaction: $\beta=-0.03$, $P<0.001$). However, for losses there was a main effect of agent on the force exerted ($\beta=-0.06$, $P<0.001$), but no significant interaction between agent and effort ($\beta=-0.02$, $P=0.14$). Given the effects of effort across studies 1 and 2, and the interactions between effort and agent when rewards were won in both studies, these results consistently show that people apply less grip force when exerting effort to reward others compared with themselves. Moreover, when benefitting others, at higher levels of effort, prosocial acts are energized to a lesser degree when rewards are monetary (see Supplementary Fig. 2).

There is considerable debate in philosophy, economics and psychology as to whether humans are truly prosocial; that is, whether they are willing to suffer costs to themselves to benefit others[2,3,6,14,62]. While empirical studies have quantified how willing people are to donate to charity or to individuals, or to cooperate with others[4–7,63,64], a crucial factor of everyday prosocial acts is that they come at the cost of quantums of effort that must be exerted. Here, we derived a computational model, through Bayesian model comparison approaches, that characterizes how motivated people are to exert effort to benefit themselves versus others. On two different metrics of effort, which were set to people's own levels of strength, we found that people are, in fact, prosocially apathetic. Although they will perform prosocial acts, they choose to do so less frequently than acts that benefit themselves (Fig. 2). Moreover, even after choosing to do so, they put in less force (Fig. 4). These effects were replicated across two studies, in which either rewards could be won or losses avoided, and were present whether the other person was completely anonymous or not.

To characterize the subjective influence of effort on rewards for self and other, we examined a range of computational models that describe how rewards are devalued by the amount of effort required to obtain them. The winning model was a parabolic discounting model with separate parameters for devaluation of rewards for self and other person trials and a single noise parameter (Fig. 3). This was robust across studies, regardless of whether people had to exert effort to gain reward or to avoid losing it. Moreover, there was a significant correlation between levels of self-reported social apathy and the model discount (K) values for how people devalued reward for another person. The difference in the discounting parameters between motivation for self and other also correlated with affective–interpersonal psychopathic traits, consistent with

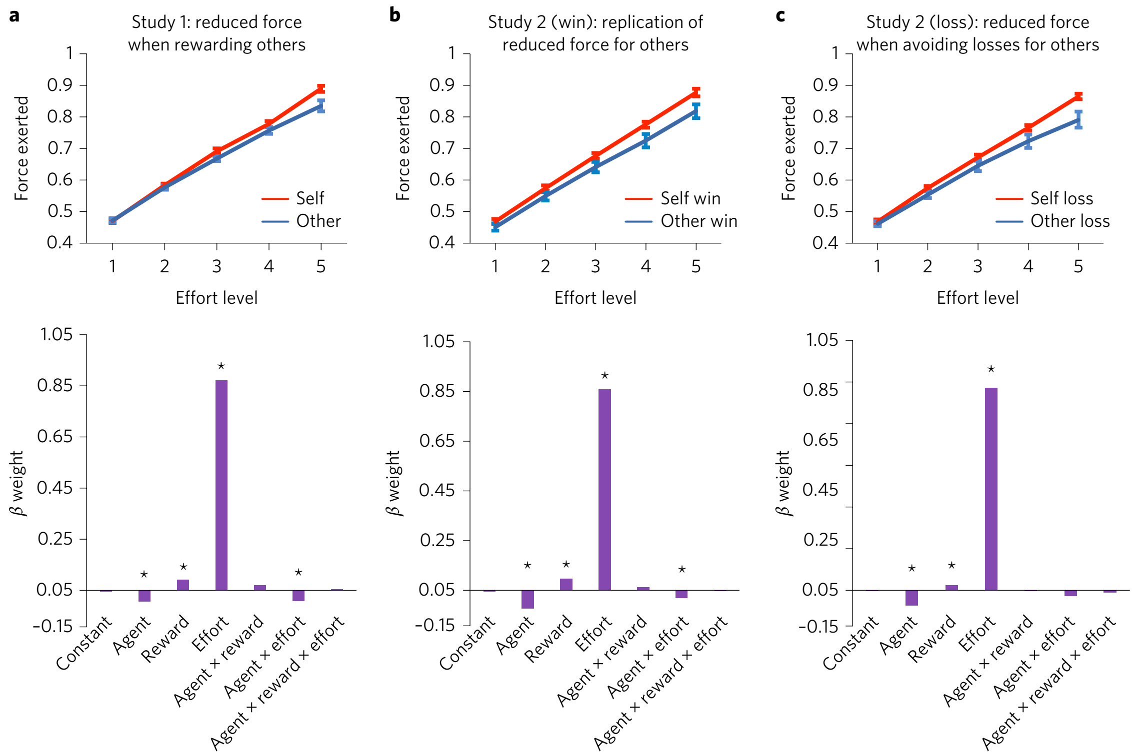

Figure 4 | Reduced force when exerting effort to help others. When choosing the experimental option for other, participants applied less grip force than when applying force to reward themselves. **a**, Force exerted as the total area under the curve during the 'effort period' on each trial (top panel). Participants applied less grip force when rewarding the other person compared with self at the higher levels of effort. In a linear regression, agent, reward, effort and the agent × effort interaction were all significant predictors of force exerted (all $P < 0.001$). **b**, Reduced force when rewarding others was replicated in study 2 (top). As in study 1, agent, reward, effort and the agent × effort interaction were significant predictors of force exerted (all $P < 0.001$). **c**, Participants also applied less force when avoiding losses for others compared with self (top). Agent, reward and effort were again all significant predictors of force exerted (all $P < 0.001$). The asterisks show significant β weight ($P < 0.001$) and the error bars show s.e.m. Together, these results show that people are superficially prosocial, even after choosing to exert effort, participants apply less grip force to reward other people than they apply to reward themselves.

greater discounting of rewards for others compared with self in people who have higher levels of psychopathy.

Overall, these results extend recent research examining the boundaries of prosocial behaviour. Economic theories posit that humans are inherently self-interested and predict that a rational agent would not exhibit any prosocial behaviours unless there were a benefit to themselves[6,32]. However, research into moral dilemmas has shown that people can be hyperaltruistic and desire to prevent harm to others more than they desire to prevent harm to themselves[4,5]. Prosocial motivation therefore does not appear to completely accord with classical economic theories or research examining prosocial behaviour in terms of moral decisions. Instead, people are willing to put in effort to help others, suggesting that they can be altruistic, but they are less motivated to do so than for themselves, which we refer to as prosocial apathy.

A plethora of research has now shown that while we can process information about others, when the same information pertains to ourselves, it is processed with greater fidelity and speed[65]. People therefore show a 'self-bias' for many sources of information. In the prosocial domain, this self-bias is also observed when people learn to perform actions that help others, with individuals being faster to learn which actions will benefit themselves[7]. Our results suggest that there may be a similar self-bias for motivation to exert effort. People are willing to exert effort for others and to energize prosocial acts, but to a lesser degree than they will for themselves.

One important benefit of our design is that we had two metrics of prosocial motivation: both choices and the measurement of force. We were therefore able to identify a previously unobserved, but intuitive, phenomenon: people show 'superficial prosociality'. They indicate they are willing to put in effort to help others, but when it comes to performing a prosocial act, they do not put in effort to the same degree as if they were doing it for their own benefit. Thus, strikingly, we find that even when people show seeming willingness to benefit others, they are not in fact as motivated to benefit another person, particularly when the effort required is high (Fig. 4). Moreover, although we found consistent results across both metrics, there was no *a priori* reason to expect this to be the case.

It could be argued that the participants in our study were simply complying with a social norm—that prosocial acts should not be directed towards strangers. However, such a norm would logically operate at all levels of reward and effort. That is, people should always choose not to benefit another, regardless of whether the other person will get a small or large reward. We found consistent interactions between effort and agency in people's decisions, but less consistent reward by agency effects. At low levels of effort, people were relatively similar in their willingness to choose to exert effort to benefit self or other, compared with at higher levels of effort, when the disparity in motivation between self and other was greater. Similar interaction effects were also present in participants' energization of actions that benefitted self or other. Thus, their willingness to exert effort for others depended on the costs of acting, suggesting that they were being influenced by the effort costs and not simply using a heuristic that benefiting strangers is undesirable.

There is emerging evidence that the systems in the brain that might be crucial for computing the subjective value of exerting effort, and guiding our choices of whether it is worth allocating resources, are partially distinct from those that are engaged in mobilizing the resources that are required to actually execute the actions[28,29,66]. It is therefore striking that we found evidence of reduced motivation to benefit others relative to self in both the choices and the energization of actions. This result also throws into question whether the hyperaltruistic behaviours observed for morally driven prosocial acts[4,5] may also be due to those studies only examining choice behaviours. Perhaps, if the metrics of people's actual energization of acts that prevent harm to self versus other were examined, their superficiality might lead to a reduction in our assessment of how altruistic people are.

Prosocial acts can be driven by the reputation of another person or by the possibility that prosocial acts will be reciprocated[2,14,43,47–49]. To maintain experimental control, we tried to limit the effects of reputation and reciprocity by keeping the other participant anonymous in study 1, and by informing the participant that the other person would not know how much money had been earned or lost in study 2. We also did not observe any interactions between participant gender and the willingness to exert effort for self or other, suggesting similar levels of prosocial motivation in males and females. These results show that regardless of context, people are less motivated to incur costs to benefit others. However, we found increased motivation for benefitting others in study 2 compared with study 1, suggesting that meeting face to face with another person may increase prosocial motivation, even if it still does not lead to the same levels of motivation as when benefitting ourselves. Future studies that manipulate social context, such as the reputation of the other person or the reciprocity of acts, or that examine the allocation of effort between individuals may be able to determine whether prosocial motivation can be raised to an equal or higher level than self-motivation. By examining these effects in ecologically valid situations, the generalizability of our results across the spectrum of social situations could be determined.

An alternative account of prosocial behaviour has come from the notion that people experience a 'warm glow' or 'vicarious reward' from acts that benefit others. This hypothesis is motivated by the fact that the areas of the brain that process rewards that others receive overlap with those that process one's own rewards[67–71]. This could potentially account for one element of the results presented here; specifically, that participants were motivated to perform acts that benefit others at all. However, it cannot account for why people show less motivation to benefit others. Intriguingly, there is growing evidence that there are areas of the brain that play important roles in processing information specifically about costs and benefits that pertain to others and not to ourselves[69,72,73]. The anterior cingulate gyrus and the subgenual anterior cingulate cortex have been shown to process this information only when processing the costs and benefits for others, and not ourselves, and when learning to be prosocial[7,16,69,72,74–76]. These regions also show variability in processing such information that correlates with trait levels of empathy—a crucial driver of prosocial behaviour[7,62,75]. Moreover, there is growing evidence that the exertion of effort to cooperate or compete with others is driven by separable neural circuits[77]. This would suggest that processing information about the benefits for others might be encoded in a separate system, which may motivate prosocial acts, and its variability drives individual differences in prosocial motivation.

Although, overall, the participants were prosocially apathetic, we also provide evidence that this tendency varies between people[4,7,17,40,41,42]. While self-reported levels of behavioural activation apathy were associated with an increased devaluation of rewards for self, thus replicating previous findings[17,40], social apathy related specifically to how motivated people were to put in effort for others. Moreover, apathy in the self or social domain only related to willingness to exert effort in the self or other domain, respectively. This suggests that a key component of social apathy might be the lack of willingness to engage in effortful social acts, which can be characterized by a specific computational measure. Previous studies have also suggested that self-reported levels of prosociality are lower among those with high psychopathic traits[43,44]. Here, we show that lower prosocial motivation on the behavioural level is also apparent in those with higher levels of affective–interpersonal psychopathic traits. Such findings hold promise for using this measure in future studies to precisely quantify levels of prosocial motivation in individuals with clinical levels of apathy and psychopathy.

Using computational modelling on an effort-based decision-making task, we have shown that people show a self-bias in motivation. They are less willing to choose to put in effort for others and, subsequently, they energize prosocial acts less than similarly effortful self-benefitting acts. We also show that people can be superficially prosocial, which means they are willing to put in effort to help others, but energize the same actions to a lesser degree than they would do if they were the beneficiary. These effects, replicated across two studies, were present both when the other person was fully anonymous and when they had met face to face, and also occurred regardless of whether the aim was to win more money or to avoid losses. Such effects were correlated with individual differences in apathy and psychopathy, highlighting the potential for this framework to enlighten our understanding of prosocial behaviour and its links to disorders of social cognition.

Methods

Study 1. *Participants.* Fifty-three participants aged 19–35 years old (mean age: 24.4, s.d.: 4.0, 31 females) took part. Participants were recruited through university databases. The sample size was based on a study by Crockett *et al.*[4], with additional participants to account for potential exclusions after testing (see below). All participants provided written informed consent and the study was approved by the Oxford University Medical Sciences Inter Divisional Research Ethics Committee. Exclusion criteria included previous or current neurological or psychiatric disorders (as reported by the participants) and non-normal or non-corrected-to-normal vision. Five participants were excluded from the study because they reported a disbelief in the deception. Participants were instructed that the money they would receive at the end of the experiment, as well as the earnings of the confederate participant, would depend on their task performance and would vary between £10 and 15.

Design. Participants completed 5 blocks of 30 trials (150 total) comprised of 75 decisions for themselves and 75 decisions for the other person. Breaks were included to avoid participant fatigue. Each trial involved a choice between a baseline option that consisted of gaining 1 credit for no effort or an alternative experimental 'offer' that varied in the level of effort (30, 40, 50, 60 or 70% MVC) and the level of reward (2, 4, 6, 8 or 10 credits; Fig. 1).

Apparatus. Stimulus presentation was programmed on a PC using MATLAB version 2011b (MathWorks; http://www.mathworks.com/) and Psychtoolbox version 3 (http://psychtoolbox.org/). Force was recorded using a hand-held TSD121B-MRI (BIOPAC Systems). The PC screen provided subjects with real-time visual feedback on the force being exerted.

Procedure. Study 1 began with role assignment. To ensure that participants believed that their choices were truly resulting in outcomes for another person, they were told that there was a second participant taking part in the experiment, but they did not see the other participant (who was in fact a confederate). This followed the procedure described by Crockett *et al.*[4,5]. Participants were told that selecting a ball from a box would randomly assign them to the different roles. Participants were handed a black glove and told not to speak so that the identity of either participant could not be uncovered. A second experimenter arrived in the room, bringing the confederate participant with them. The confederate participant was handed a second glove, but they remained behind the other side of the door at all times without ever being seen by the participant. The participants were asked to place their hands in front of the door and to wave to one another to ensure it was clear that there was another person there. The experimenter then tossed a coin to decide who would pick from the box first. Each participant selected a ball and was told which role in the experiment they were assigned to.

Task procedure. During the task procedure, participants were asked to grip a handheld dynometer with as much force as possible to determine their MVC. This ensured that although individuals differ in their strength, the effort levels used in the experiment would be relative to their MVC. This measurement was then used as a subject-specific threshold for the levels of effort required to obtain rewards in the main task. In the experimental task, participants made decisions between a baseline low effort (0% of the MVC) option that gained 1 credit and a variable offer in which more credits (2, 4, 6, 8 or 10 credits) were available but which also required more force (30, 40, 50, 60 or 70% of the MVC—represented by segments in a pie chart). The effort and reward levels were varied independently over the trials, with each effort–reward combination sampled three times for each agent. There were 150 trials in total, with 75 self trials, in which participants chose between the offer and the baseline for themselves, and 75 other trials, in which they made these decisions for the other person. To obtain the rewards in each trial, the participants had to apply a force that exceeded the required level for a total of 1 s in a 3 s window. Failure to do so resulted in 0 credits being delivered. The baseline condition was 1 credit, to ensure that there was a clear incentive for participants to choose the baseline if the value was not considered worth it, rather than them choosing the offer and then not exerting any effort at all. If a choice was not selected, 0 credits were delivered. All trials, regardless of the choice made (or if no response was made) lasted for the same duration. This ensured that choices were not influenced by the discounting effects of temporal delay rather than effort (for example, ref. [78]). Indeed, the failure rates were very low in the main experiment (<3% in study 1; <3.5% in study 2 (win); and <3.2% in study 2—loss), indicating that subjects were almost always able to achieve the required amount of force. The fact that the failure rates were so low also helps to rule out the potential effects of risk aversion (for example, ref. [79]) that may interact with effort discounting, as participants had a very high probability of receiving the rewards from the options they chose.

Prior to the decision-making task, participants experienced each effort level three times across 18 trials. They also learnt to associate each level of effort with the elements in a pie chart. They were instructed that if only one element of the pie chart was shown then 0% force was required and that this was the baseline offer, equivalent to a 'rest'. However, they still had to grip the dynometer in their hand. During the training session, only 1 credit was on offer and participants were instructed that this credit would not count towards their payment. In the training session, participants did not choose whether to opt out of exerting effort.

Study 2. *Participants.* Fifty-eight participants (mean age: 25.7, s.d.: 6.1, 31 females) took part in two sessions. The sample size was based on study 1. Participants were recruited through university databases. All participants provided written informed consent and the study was approved by the Oxford University Medical Sciences Inter Divisional Research Ethics Committee. Thirteen participants were excluded from the study because they showed a strong disbelief in the deception. Exclusion criteria included previous or current neurological or psychiatric disorders (as reported by the participants) and non-normal or non-corrected-to-normal vision. Participants were instructed that the money they would receive at the end of the experiment, as well as the earnings of the confederate participant, would depend on their task performance and would vary between £15 and 25.

Design. There were two additions to the design of study 1. First, although participants performed a session where the aim was to win points, as in study 1, they also included a second session, where the aim was to avoid losing points. Second, to avoid participant fatigue, they performed these sessions on separate days (at least one day apart) in a counterbalanced order across participants. In the loss session, participants made decisions between a baseline option of −9 credits (equivalent to winning 1 point) for 0% MVC or a variable offer where the offers would be losing 0, −2, −4, −6 or −8 credits for themselves or the other participant. If participants chose to accept the experimental option but did not obtain the reward, they lost 10 credits. The incentive structure was therefore identical, but the framing was distinct.

Procedure. In study 2, we used a role assignment procedure where participants met another person face to face. This allowed us to examine whether increasing the familiarity and removing the anonymous nature of the other players made people less prosocially apathetic. At the beginning of the experiment, the participants were introduced to another participant who was in fact a confederate of the experiment. The two participants were instructed that they would be asked to perform different tasks and that only one of them would be in charge of determining the extra amount of money that both would earn (Fig. 1c).

Studies 1 and 2. *Statistical analyses of behavioural data..* Analyses of behavioural data were performed in SPSS 22 (IBM). We examined differences between conditions in choice behaviour for effort and reward and self versus other using separate repeated measures ANOVAs. For the comparison between self and other discounting (K) rates from the model comparison, we used Wilcoxon signed-rank tests as the data were not normally distributed. For participants who never chose the variable offer, their discount (K) values were set to 0, as they did not discount rewards by effort (four participants in study 2). We also examined bivariate associations between K_{other}, K_{self} and K_{other}–K_{self} using Spearman's non-parametric correlation coefficient, also due to non-normality of the data. For the correlations between psychometric scores and computational parameters, we only performed planned comparisons that were *a priori* and hypothesis driven. On the basis of previous findings, we hypothesized that discount rates only for others would be associated with social apathy[34], and that discount rates only for self would be associated with behavioural apathy[17,40]. For correlations with psychopathic traits, we assessed the correlation with the total score and report the correlations with the two subscales as exploratory follow-up analyses. All comparisons were corrected for the false-discover rate using the Benjamini and Hochberg false-discover rate procedure[80]. To test whether there were differences between the force exerted for other and self, we calculated the area under the curve of the voluntary contraction trace recorded from a hand-held TSD121B-MRI (BIOPAC Systems Inc.) using the function 'trapz' in MATLAB 2011b (MathWorks). This standard function computes the integral of Y with respect to X using the trapezoidal method. We standardized this value by the maximum level of force exerted to take into account baseline differences in participants' MVCs. We then used linear regression analyses to predict the amount of force exerted by reward level, effort level and agent, and interactions between these variables. For all the analyses, the alpha level was set to $P < 0.05$ two-tailed.

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability. The computer code that supports the findings of this study are available from the corresponding author upon reasonable request.

Received 12 May 2017; accepted 23 May 2017;
published 30 June 2017

References

1. Fehr, E. & Camerer, C. F. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* **11,** 419–427 (2007).
2. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425,** 785–791 (2003).
3. Fehr, E. & Rockenbach, B. Human altruism: economic, neural, and evolutionary perspectives. *Curr. Opin. Neurobiol.* **14,** 784–790 (2004).
4. Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P. & Dolan, R. J. Harm to others outweighs harm to self in moral decision making. *Proc. Natl Acad. Sci. USA* **111,** 17320–17325 (2014).
5. Crockett, M. J. *et al.* Dissociable effects of serotonin and dopamine on the valuation of harm in moral decision making. *Curr. Biol.* **25,** 1852–1859 (2015).
6. Engel, C. Dictator games: a meta study. *Exp. Econ.* **14,** 583–610 (2011).
7. Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E. & Roiser, J. P. Neurocomputational mechanisms of prosocial learning and links to empathy. *Proc. Natl Acad. Sci. USA* **113,** 9763–9768 (2016).
8. Anderson, N. E. & Kiehl, K. A. The psychopath magnetized: insights from brain imaging. *Trends Cogn. Sci.* **16,** 52–60 (2012).
9. Batson, C. D. in *The Handbook of Social Psychology* 4th edn (McGraw-Hill, 1998).
10. Blair, J., Mitchell, D & Blair, K. in *The Psychopath: Emotion and the Brain* (Blackwell Publishing, 2005).
11. Blair, R. J. R., Jones, L., Clark, F. & Smith, M. Is the psychopath 'morally insane'? *Personal. Individ. Differ.* **19,** 741–752 (1995).
12. Krajbich, I., Adolphs, R., Tranel, D., Denburg, N. L. & Camerer, C. F. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J. Neurosci.* **29,** 2188–2192 (2009).
13. Lockwood, P. L., Seara-Cardoso, A. & Viding, E. Emotion regulation moderates the association between empathy and prosocial behavior. *PLoS ONE* **9,** e96555 (2014).
14. Fehr, E. & Fischbacher, U. Social norms and human cooperation. *Trends Cogn. Sci.* **8,** 185–190 (2004).
15. Apps, M. A. J., Grima, L. L., Manohar, S. & Husain, M. The role of cognitive effort in subjective reward devaluation and risky decision-making. *Sci. Rep.* **5,** 16880 (2015).
16. Apps, M. A. & Ramnani, N. The anterior cingulate gyrus signals the net value of others' rewards. *J. Neurosci.* **34,** 6190–6200 (2014).
17. Bonnelle, V. *et al.* Characterization of reward and effort mechanisms in apathy. *J. Physiol. Paris* **109,** 16–26 (2014).
18. Kurniawan, I. T. *et al.* Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action. *J. Neurophysiol.* **104,** 313–321 (2010).
19. Vassena, E. *et al.* Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS ONE* **9,** e91008 (2014).

20. Inzlicht, M., Bartholow, B. D. & Hirsh, J. B. Emotional foundations of cognitive control. *Trends Cogn. Sci.* **19,** 126–132 (2015).
21. Bonnelle, V., Manohar, S., Behrens, T. & Husain, M. Individual differences in premotor brain systems underlie behavioral apathy. *Cereb. Cortex* **26,** 807–819 (2016).
22. Hartmann, M. N., Hager, O. M., Tobler, P. N. & Kaiser, S. Parabolic discounting of monetary rewards by physical effort. *Behav. Processes* **100,** 192–196 (2013).
23. Klein-Flügge, M. C., Kennerley, S. W., Friston, K. & Bestmann, S. Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort–reward trade-off. *J. Neurosci.* **36,** 10002–10015 (2016).
24. Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D. & Bestmann, S. Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLoS Comput. Biol.* **11,** e1004116 (2015).
25. Skvortsova, V., Palminteri, S. & Pessiglione, M. Learning to minimize efforts versus maximizing rewards: computational principles and neural correlates. *J. Neurosci.* **34,** 15621 (2014).
26. Westbrook, A. & Braver, T. S. Cognitive effort: a neuroeconomic approach. *Cogn. Affect. Behav. Neurosci.* **15,** 395–415 (2015).
27. Westbrook, A. & Braver, T. S. Dopamine does double duty in motivating cognitive effort. *Neuron* **89,** 695–710 (2016).
28. Ang, Y.-S., Manohar, S. & Apps, M. A. J. Commentary: noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *Front. Behav. Neurosci.* **9,** 310 (2015).
29. Varazzani, C., San-Galli, A., Gilardeau, S. & Bouret, S. Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *J. Neurosci.* **35,** 7866–7877 (2015).
30. Manohar, S. G. *et al.* Reward pays the cost of noise reduction in motor and cognitive control. *Curr. Biol.* **25,** 1707–1716 (2015).
31. Winking, J. & Mizer, N. Natural-field dictator game shows no altruistic giving. *Evol. Hum. Behav.* **34,** 288–293 (2013).
32. Camerer, C. *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton Univ. Press, 2003).
33. Marin, R. S. Apathy: concept, syndrome, neural mechanisms, and treatment. *Semin. Clin. Neuropsychiatry* **1,** 304–314 (1996).
34. Ang, Y.-S., Lockwood, P., Apps, M. A. J., Muhammed, K. & Husain, M. Distinct subtypes of apathy revealed by the apathy motivation index. *PLoS ONE* **12,** e0169938 (2017).
35. Brodaty, H., Altendorf, A., Withall, A. & Sachdev, P. Do people become more apathetic as they grow older? A longitudinal study in healthy individuals. *Int. Psychogeriatr.* **22,** 426–436 (2010).
36. Clarke, D. E., Ko, J. Y., Lyketsos, C., Rebok, G. W. & Eaton, W. W. Apathy and cognitive and functional decline in community-dwelling older adults: results from the Baltimore ECA longitudinal study. *Int. Psychogeriatr.* **22,** 819–829 (2010).
37. Vansteenkiste, M., Lens, W., De Witte, S., De Witte, H. & Deci, E. L. The 'why' and 'why not' of job search behaviour: their relation to searching, unemployment experience, and well-being. *Eur. J. Soc. Psychol.* **34,** 345–363 (2004).
38. Vansteenkiste, V., Lens, W., Witte, H. & Feather, N. Understanding unemployed people's job search behaviour, unemployment experience and well-being: a comparison of expectancy-value theory and self-determination theory. *Br. J. Soc. Psychol.* **44,** 269–287 (2005).
39. Insel, T. *et al.* Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* **167,** 748–751 (2010).
40. Hartmann, M. N. *et al.* Apathy but not diminished expression in schizophrenia is associated with discounting of monetary rewards by physical effort. *Schizophr. Bull.* **41,** 503–512 (2015).
41. FeldmanHall, O., Dalgleish, T., Evans, D. & Mobbs, D. Empathic concern drives costly altruism. *NeuroImage* **105,** 347–356 (2015).
42. Zaki, J. & Ochsner, K. N. The neuroscience of empathy: progress, pitfalls and promise. *Nat. Neurosci.* **15,** 675–680 (2012).
43. White, B. A. Who cares when nobody is watching? Psychopathic traits and empathy in prosocial behaviors. *Personal. Individ. Differ.* **56,** 116–121 (2014).
44. Foulkes, L., McCrory, E. J., Neumann, C. S. & Viding, E. Inverted social reward: associations between psychopathic traits and self-report and experimental measures of social reward. *PLoS ONE* **9,** e106000 (2014).
45. Salamone, J. D., Yohn, S. E., López-Cruz, L., San Miguel, N. & Correa, M. Activational and effort-related aspects of motivation: neural mechanisms and implications for psychopathology. *Brain J. Neurol.* **139,** 1325–1347 (2016).
46. Kennerley, S. W., Behrens, T. E. J. & Wallis, J. D. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* **14,** 1581–1589 (2011).
47. Kurzban, R., DeScioli, P. & O'Brien, E. Audience effects on moralistic punishment. *Evol. Hum. Behav.* **28,** 75–84 (2007).
48. Piazza, J. & Bering, J. M. Concerns about reputation via gossip promote generous allocations in an economic game. *Evol. Hum. Behav.* **29,** 172–178 (2008).
49. Haley, K. J. & Fessler, D. M. T. Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.* **26,** 245–256 (2005).
50. Tversky, A. & Kahneman, D. Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* **5,** 297–323 (1992).
51. Tversky, A. & Kahneman, D. Loss aversion in riskless choice: a reference-dependent model. *Q. J. Econ.* **106,** 1039–1061 (1991).
52. Chong, T. T.-J. *et al.* Neurocomputational mechanisms underlying valuation of effort costs. *PLoS Biol.* **15,** e1002598 (2017).
53. Blair, R. J. R. The neurobiology of psychopathic traits in youths. *Nat. Rev. Neurosci.* **14,** 786–799 (2013).
54. Cleckley, H. M. *The Mask of* Sanity: *An Attempt to Clarify Some Issues about the So-Called Psychopathic Personality* (Mosby, 1964).
55. Hare, R. D. & Neumann, C. N. in *Handbook of Psychopathy* (ed. Patrick, C. J.) 58–88 (Guilford Press, 2006).
56. Hare, R. D. *The Hare Psychopathy Checklist-Revised: PLC-R* (Multi-Health Systems, 1999).
57. Paulhus, D., Neumann, C. & Hare, R. *Manual for the Self-Report Psychopathy Scale* (Multi-Health Systems, 2009).
58. Lockwood, P. L., Bird, G., Bridge, M. & Viding, E. Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Front. Hum. Neurosci.* **7,** 760 (2013).
59. Seara-Cardoso, A., Neumann, C., Roiser, J., Mccrory, E. & Viding, E. Investigating associations between empathy, morality and psychopathic personality traits in the general population. *Personal. Individ. Differ.* **52,** 67–71 (2012).
60. Porter, S. & Woodworth, M. in *Handbook of Psychopathy* (ed Patrick, C. J.) 481–494 (Guilford Press, 2006).
61. Blair, R. J. Psychopathic traits from an RDoC perspective. *Curr. Opin. Neurobiol.* **30,** 79–84 (2015).
62. Decety, J., Bartal, I. B.-A., Uzefovsky, F. & Knafo-Noam, A. Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **371,** 20150077 (2016).
63. Ruff, C. C. & Fehr, E. The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15,** 549–562 (2014).
64. Lee, D. & Seo, H. Neural basis of strategic decision making. *Trends Neurosci.* **39,** 40–48 (2016).
65. Sui, J. & Humphreys, G. W. The integrative self: how self-reference integrates perception and memory. *Trends Cogn. Sci.* **19,** 719–728 (2015).
66. Floresco, S. B. Noradrenaline and dopamine: sharing the workload. *Trends Neurosci.* **38,** 465–467 (2015).
67. Andreoni, J. & Miller, J. Giving according to GARP: an experimental test of the consistency of preferences for altruism. *Econometrica* **70,** 737–753 (2002).
68. Braams, B. R. *et al.* Reward-related neural responses are dependent on the beneficiary. *Soc. Cogn. Affect. Neurosci.* **9,** 1030–1037 (2014).
69. Lockwood, P. L. The anatomy of empathy: vicarious experience and disorders of social cognition. *Behav. Brain Res.* **311,** 255–266 (2016).
70. Mobbs, D. *et al.* A key role for similarity in vicarious reward. *Science* **324,** 900 (2009).
71. Morelli, S. A., Sacchet, M. D. & Zaki, J. Common and distinct neural correlates of personal and vicarious reward: a quantitative meta-analysis. *NeuroImage* **112,** 244–253 (2015).
72. Apps, M. A. J., Rushworth, M. F. S. & Chang, S. W. C. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron* **90,** 692–707 (2016).
73. Lamm, C. & Majdandžić, J. The role of shared neural activations, mirror neurons, and morality in empathy—a critical comment. *Neurosci. Res.* **90,** 15–24 (2015).
74. Chang, S. W. C., Gariépy, J.-F. & Platt, M. L. Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* **16,** 243–250 (2013).
75. Lockwood, P. L., Apps, M. A. J., Roiser, J. P. & Viding, E. Encoding of vicarious reward prediction in anterior cingulate cortex and relationship with trait empathy. *J. Neurosci.* **35,** 13720–13727 (2015).
76. Balsters, J. H. *et al.* Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain* **140,** 235–246 (2016).
77. Le Bouc, R. & Pessiglione, M. Imaging social motivation: distinct brain mechanisms drive effort production during collaboration versus competition. *J. Neurosci.* **33,** 15894–15902 (2013).
78. Green, L. & Myerson, J. A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* **130,** 769–792 (2004).
79. Kahneman, D. & Tversky, A. Prospect theory: an analysis of decision under risk. *Econometrica* **47,** 263–292 (1979).
80. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57,** 289–300 (1995).

Acknowledgements

We thank C. Neumann for his help and advice with regard to the self-report psychopathy scale. We also thank all members of the Cognitive Neurology Research Group for their assistance as experimental confederates. This work was supported by a Biotechnology and Biological Sciences Research Council Anniversary Future Leader Fellowship (BB/M013596/1) to M.A.J.A. and a Wellcome Trust Principal Fellowship and the National Institute for Health Research Oxford Biomedical Centre to M.Husain. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

P.L.L., M.A.J.A. and M.Husain designed the study, P.L.L., M.Hamonet, S.H.Z., A.R, F.U.S. and M.A.J.A. collected the data, P.L.L. and M.A.J.A. analysed the data, and P.L.L., M.A.J.A. and M.Husain wrote the paper.

Additional information

Supplementary information is available for this paper.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to P.L.L.

How to cite this article: Lockwood, P. L. *et al.* Prosocial apathy for helping others when effort is required. *Nat. Hum. Behav.* **1,** 0131 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Competing interests

The authors declare no competing interests.