PNAS Nexus, 2023, 2, 1–15
https://doi.org/10.1093/pnasnexus/pgad199
Advance access publication 13 June 2023
Research Report

Neural responses underlying extraordinary altruists' generosity for socially distant others

Shawn A. Rhoads[a,*], Katherine O'Connell[b], Kathryn Berluti[a], Montana L. Ploe[a], Hannah S. Elizabeth[a], Paige Amormino[a], Joanna L. Li[a], Mary Ann Dutton[c], Ashley Skye VanMeter[b,d] and Abigail A. Marsh[a,b]

[a]Department of Psychology, Georgetown University, 3700 O St NW, Washington, DC 20057, USA
[b]Interdisciplinary Program in Neuroscience, Georgetown University, 3700 O St NW, Washington, DC 20057, USA
[c]Department of Psychiatry, Georgetown University, 3700 O St NW, Washington, DC 20057, USA
[d]Department of Neurology, Georgetown University, 3700 O St NW, Washington, DC 20057, USA
*To whom correspondence should be addressed: Email: sr1209@georgetown.edu
Edited By: Paulo Boggio

Abstract

Most people are much less generous toward strangers than close others, a bias termed social discounting. But people who engage in extraordinary real-world altruism, like altruistic kidney donors, show dramatically reduced social discounting. Why they do so is unclear. Some prior research suggests reduced social discounting requires effortfully overcoming selfishness via recruitment of the temporoparietal junction. Alternatively, reduced social discounting may reflect genuinely valuing strangers' welfare more due to how the subjective value of their outcomes is encoded in regions such as rostral anterior cingulate cortex (ACC) and amygdala. We tested both hypotheses in this pre-registered study. We also tested the hypothesis that a loving-kindness meditation (LKM) training intervention would cause typical adults' neural and behavioral patterns to resemble altruists. Altruists and matched controls ($N = 77$) completed a social discounting task during functional magnetic resonance imaging; 25 controls were randomized to complete LKM training. Neither behavioral nor imaging analyses supported the hypothesis that altruists' reduced social discounting reflects effortfully overcoming selfishness. Instead, group differences emerged in social value encoding regions, including rostral ACC and amygdala. Activation in these regions corresponded to the subjective valuation of others' welfare predicted by the social discounting model. LKM training did not result in more generous behavioral or neural patterns, but only greater perceived difficulty during social discounting. Our results indicate extraordinary altruists' generosity results from the way regions involved in social decision-making encode the subjective value of others' welfare. Interventions aimed at promoting generosity may thus succeed to the degree they can increase the subjective valuation of others' welfare.

Keywords: altruism, social discounting, social valuation, prosocial choice, neuroeconomics

Significance Statement

Most people are much less generous toward distant strangers than close others. But real-world extraordinary altruists (e.g. altruistic organ donors) choose to sacrifice for a stranger as most people would only for a close friend or relative. We found no evidence that this results from altruists effortfully overcoming selfish preferences. Instead, our results suggest altruists are more generous to strangers because of how they encode the subjective value of distant others' welfare in a constellation of social brain regions that include the amygdala and rostral anterior cingulate cortex. We sought to reproduce altruists' generous behavioral and neural patterns in controls using a meditation intervention but were unsuccessful.

Introduction

Why people sometimes choose to provide costly help to strangers—including anonymously giving them money, blood, or even organs—is a longstanding puzzle (1, 2). Anonymously helping socially distant others is inconsistent with established biological theories of altruism such as kin selected and reciprocal altruism (3, 4). Such actions also seem to contradict widespread beliefs among many scholars and the general public that humans are inherently selfish (5, 6). One possible resolution to this puzzle is that acting generously toward socially distant others requires successfully overriding selfish preferences. Some prior neuroimaging studies of intra-individual variation in generous decision-making support this possibility. Increased activation in regions such as the temporoparietal junction (TPJ) during generous decisions in social discounting tasks has been interpreted as reflecting this region modulating value signals in medial prefrontal cortex (PFC) to resolve conflicts between selfish and generous choices (7, 8). This interpretation is consistent with theories of prosocial restraint,

Competing Interest: The authors declare no competing interest.
Received: October 11, 2022. **Revised:** April 22, 2023. **Accepted:** June 2, 2023

according to which prosocial choices result from deliberately restraining selfish impulses (9, 10).

This interpretation suggests that those who engage in extraordinary, costly altruism to benefit anonymous strangers, such as altruistic organ donors, may be more generous both in the real world and in social discounting tasks (11, 12) because they are extraordinarily good at overriding their selfish preferences. And perhaps interventions with the potential to increase costly generosity, such as loving-kindness meditation (LKM) training (13, 14), do so by improving the ability to override selfish preferences.

But other findings from behavioral and comparative research are inconsistent with this hypothesis. Behavioral research suggests costly generosity (including real-world extraordinary altruism) tends to be fast and intuitive, consistent with it reflecting true prosocial preferences (15, 16). Prior work also finds real-world altruists are characterized by unselfish personality traits (11) and elevated empathic responding for others' distress or negative outcomes (17–19). This suggests that variation in costly altruism across individuals may result from processes that are distinct from those that support intra-individual variation in generous decision-making in typical adults. Rather than overriding selfish impulses when they make generous decisions, unusually altruistic individuals may instead encode the subjective value of others' outcomes differently in value-encoding brain regions, causing them to genuinely value others' welfare more. However, no prior research has tested this hypothesis.

We thus recruited a rare sample of extraordinary real-world altruists, all of whom had donated a kidney to a stranger. Operationalizing altruism as objectively measured, real-world costly generosity enables inferences to be drawn about the origins of altruism that may not be possible when using only laboratory measures, which cannot ethically assess highly risky or costly altruism and which may, like self-report measures, be contaminated by social desirability and norm-adherence motives and may not reliably correspond to real-world altruism (20). We compared altruists to demographically matched typical adults as they completed a social discounting task in which they had the opportunity to make selfish or costly, generous choices (21, 22) during functional magnetic resonance imaging (fMRI) scanning. This task has emerged as a particularly good model of altruistic helping (11, 23). Prior work consistently finds that among typical adults, generosity in this task progressively declines for more socially distant recipients (7, 12, 24–27), following a hyperbolic function:

$$v = \frac{V_0}{1 + kN},$$

where V_0 represents the intercept or undiscounted value of the reward, k represents degree of discounting (discounting rate), N represents social distance (for example, 1 representing the closest social relationship such as a spouse or child, and 100 representing a stranger), and v represents the amount willing to forgo for each social other. As N increases, the resources individuals are willing to forgo (v) typically decrease in a hyperbolic fashion governed by the parameter, k, which represents the rate by which the function decays across social distance.

The amount people are willing to sacrifice (v) is thought to reflect variation in the subjective value respondents place on others' welfare (21, 22). But, it is poorly understood how this value is neurally encoded during social discounting (one study considered this question, but the authors largely focused on the role of TPJ and found it was not associated with the subjective valuation of others' outcomes during generous choices (7)). And nothing is known about whether or how this subjective value is encoded differently in real-world altruists.

Two regions prior work has been found to be involved in encoding the subjective value of others' welfare are the amygdala and medial PFC, specifically rostral anterior cingulate cortex (ACC) (28, 29). Recent work in nonhuman animals finds that during generous decision-making, activation in rostral ACC is modulated by signals from the amygdala that encode the value of outcomes for others (30, 31). A growing body of research in both humans (32–36) and nonhuman animals (29–31, 37) supports the idea that these regions play a causal role in prosocial decisions that promote others' welfare (38). And extraordinary human altruism has been previously linked to increased volume and activation of the amygdala (19). But whether activation patterns in this region or rostral ACC correspond to generous decision-making in altruists is not yet known.

It also remains unknown whether typical adults' behavioral or neural patterns of generous decision-making can be made to more closely resemble those of real-world altruists. LKM (or *metta*) training is among the best-supported interventions for increasing care and generosity for distant others (13). This practice is designed to increase feelings of love and positivity for others at varying social distances (39). Prior work has linked LKM training with increased feelings of social connectedness (13, 40, 41), decreased intergroup bias (42), prosocial behavior in the laboratory (13, 14), and charitable donations (13, 43)—although a recent systematic review and meta-analysis found that the robustness of such effects is moderated by the type of control group (active control versus waiting list) and whether the meditation intervention teacher co-authored the published work (44). Thus, while prior work suggests LKM may cause reductions in social discounting comparable to those observed in naturally-occurring altruists, this has not previously been tested.

In a pre-registered study, we thus sought to identify the neural basis of social discounting in a sample of extraordinary altruists as well as a sample of controls, half of whom were randomized to undergo an eight-week LKM training intervention created by the world-renowned LKM trainer Sharon Salzberg (39). We tested the following hypotheses: (i) that costly altruism for strangers reflects effortful prosocial restraint and (ii) that it reflects altruists' true valuation of distant others' welfare. We also tested a third hypothesis, (iii) that typical adults' behavior and neural patterns of decision-making—in pre-registered regions such as amygdala, anterior cingulate, ventromedial PFC, insula, temporal gyrus, and TPJ—would converge with those of altruists following LKM training.

Results

Eighty participants completed three runs of a social discounting task during fMRI. This task was designed to incorporate features of previous discounting tasks (7). Three participants were excluded (one for excessive motion, with more than 15% of volumes greater than 0.5 mm framewise displacement, one because they asked to change over 20% of their choices after completing the scan, and one because they missed responding to more than half of all choices). The final sample thus included 77 participants (62.34% female; M age = 41.49 years, SD = 7.65 years). 26 participants were extraordinary altruists verified as having donated a kidney to a stranger. The remaining 51 were demographically matched typical adults who had never donated an organ. Twenty-five of these control participants were randomly selected (with stratification by age and gender) to complete an 8-week LKM training intervention prior to scanning. (See Supplementary

Table S1 for participant characteristics.) The meditation training program was custom-designed for the purposes of this research by world-renowned LKM training expert Sharon Salzberg and included 40 training sessions in which participants practiced LKM five times per week for 8 weeks prior to scanning (see Materials and methods). Participants randomized to LKM completed an average of 89.6% of training sessions (SD = 0.07%, range = 77.5–100%). Both wait-list controls and altruists were offered access to the meditation training program after the study's completion.

All participants then completed a social discounting task during fMRI. Prior to this task, and following established procedures, participants identified six individuals ranging in social closeness about whom they would make decisions during testing. These individuals included the person closest to them ($N = 1$), the person second-closest to them ($N = 2$) and so on, such that each participant identified an $N = 1$, 2, 5, 10, 20, and 50. Groups did not differ in their subjective ratings of nominated social others in terms of emotional closeness nor in terms of relationship types across social distances, suggesting that all groups similarly understood nomination instructions and similarly differentiate between ranked social targets (see Materials and methods). The task also included choices for $N = 100$, representing a stranger (7). In each trial of the social discounting fMRI task, participants had the option to choose to keep an amount of money for themselves or to share with one of these seven people. The task consisted of three independent runs of fMRI scanning, during which participants made 189 decisions separated into 21 randomized social distance blocks corresponding to one of these seven individuals (i.e. three blocks per social other). Within each block, participants made nine dichotomous choices to either keep resources for themselves (selfish option) or share them with the Nth person listed (generous option) (Fig. 1). Following MRI scanning, all participants rated how difficult each decision they had made was on a scale from 1 (very easy) to 7 (very difficult).

To avoid deception and meet standards of behavioral economics, participants' decisions resulted in real monetary outcomes. Participants were informed that one of their choices would be randomly chosen and implemented such that the participant would be paid 10% of the amount they selected to receive themselves, and, if a generous choice was selected, the other person in question also would be contacted and disbursed 10% of the amount the participant selected for them (see Materials and methods). If a participant chose to share with a stranger, a random name was selected from an online directory (United States), and the payment was disbursed to them.

The primary difference between our task design and the task previously used by Strombach et al. (7) is that we omitted the visual number line indicating social distance using spatial referents. Instead, participants saw a numeric indicator of social distance. We selected this design to minimize potential confounds due to correlations between the spatial positions of items on the number line and generous and selfish decisions.

We first aimed to confirm that altruists exhibit reduced social discounting relative to typical adults (11, 12). Results of a hyperbolic mixed-effects model confirmed that altruists are indeed more generous to socially distant others than controls (Table 1, Fig. 2), with altruists showing significantly shallower discounting slopes ($\log k$) relative to wait-list controls ($\beta = -1.64$, SE = 0.67, 95% CI = [−2.96, −0.33], $P = 0.015$, $n = 77$). However, we did not find that controls who completed LKM were any more generous towards socially distant others than wait-list controls.

We also ran a linear mixed-effects model predicting response times (log-transformed) as a function of group and decision type (Supplementary Table S2, Supplementary Fig. S1). Results indicated a main effect of decision type and altruist group. People were generally faster for generous than selfish choices on average ($\beta = 0.286$, SE = 0.011, 95% CI = [0.266, 0.307], $P < 0.001$, $n = 80$) and altruists were relatively slower than controls on average ($\beta = 0.185$, SE = 0.085, 95% CI = [0.015, 0.355], $P = 0.033$, $n = 80$). We also found a meditation group × decision type interaction: Among LKM controls, the difference in response times were less pronounced than wait-list controls on average ($\beta = -0.047$, SE = 0.024, 95% CI = [−0.093 −0.0004], $P = 0.048$, $n = 80$).

We next tested the reported difficulty of decision-making using a linear mixed-effects model to test how participants' subjective difficulty ratings varied as a function of social distance and group (Table 2, Fig. 3). Results showed that across groups, participants rated their decisions to be more difficult as social distance increased, but this effect was stronger among LKM controls than among either altruists or wait-list controls ($\beta = -0.004$, SE = 0.001, 95% CI = [−0.01, −0.003], $P < 0.001$, $n = 77$). Thus, LKM increased how subjectively difficult decisions to share with socially distant others were, even though the outcomes of prosocial decisions were not themselves affected. By contrast, altruists' reported difficulty of decisions as a function of social distance were not different from wait-list controls'. This contradicts the hypothesis that altruists' generosity for more distant others reflects greater effortful control. We also ran a logistic mixed-effects model to test the relationship between subjective difficulty and decision type (Supplementary Table S3, Supplementary Fig. S2). We found that generous decisions are easier on average across all groups (log odds = −0.517, SE = 0.040, 95% CI = [−0.596, −0.438], $P < 0.001$, $n = 77$). In addition, we found that the strength of this relationship was stronger for both altruists (log odds = −0.496, SE = 0.107, 95% CI = [−0.706, −0.286], $P < 0.001$, $n = 77$) and meditation controls (log odds = −0.390, SE = 0.095, 95% CI = [−0.576, 0.203], $P < 0.001$, $n = 77$) relative to wait-list controls.

We next assessed neural activation during generous versus selfish decision-making across social distances and across the three groups. These analyses tested the alternative hypotheses that extraordinary altruists' generosity toward distant others reflects effortful control versus greater valuation of distant others' welfare. For analyses of effortful control, we could only include the 62 participants who exhibited variation in generosity (i.e. those who did not choose the generous option for >99% of trials, as did 10 altruists, 4 wait-list controls, and 1 LKM control; Supplementary Table S4). We fit a first-level general linear model (GLM) to BOLD responses to estimate percent signal change maps for generous and selfish choices (see Materials and methods). Group-level analyses found no differences in activation patterns between LKM and wait-list controls, consistent with their statistically identical choice patterns. With the aim to include as many participants as possible, we additionally tested whether groups differed in neural activation for generous choices versus implicit baseline (fixation) among all participants with complete social discounting behavior ($N = 77$; Supplemental Table S4; see Materials and methods) and again did not find any group differences. We therefore combined both groups' data for subsequent analyses.

To test the hypothesis that altruistic decisions reflect greater effortful control, we repeated analyses performed by Strombach et al. (7), which assessed how activity varies as a function of the temptation to be selfish. To do this, for each participant and for each trial, we first computed a "selfish temptation value" by subtracting the subjective valuation placed on others' welfare from the value of a potential selfish reward for a given choice.

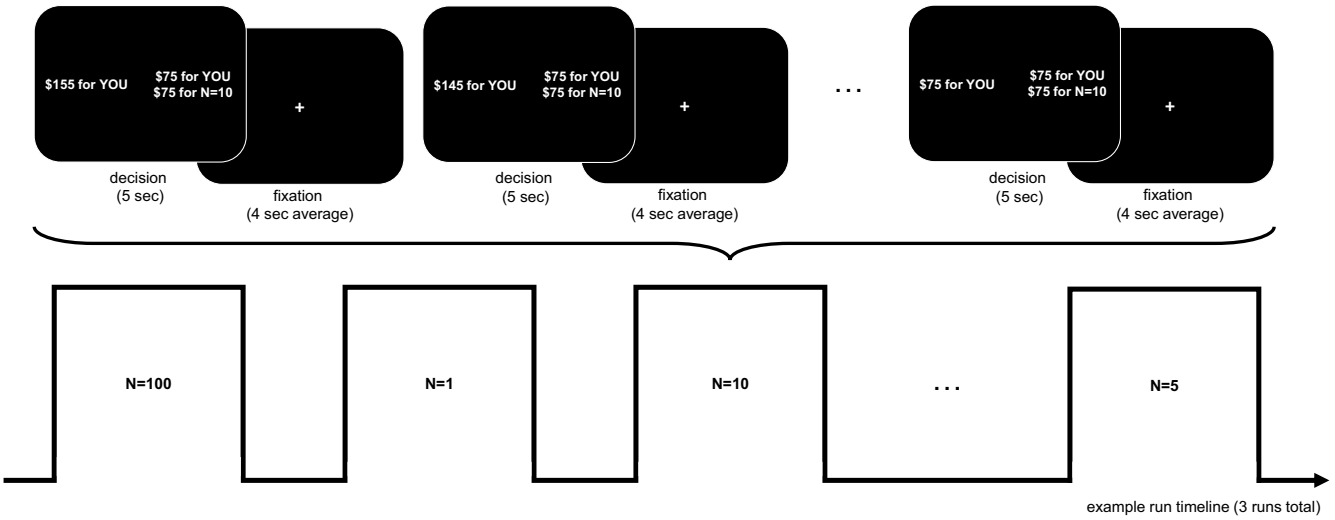

Fig. 1. Social discounting paradigm. *Note.* Prior to scanning, participants were provided with instructions that they will make 189 decisions across three independent runs involving six people from a participant's social network of varying social distances ($N = 1, 2, 5, 10, 20, 50$) plus one stranger ($N = 100$). The task consisted of 21 blocks with nine trials each, in which they made 9 dichotomous choices about keeping and/or sharing amounts of money with each. For each trial, participants were asked to indicate if they would prefer to keep an amount of money for themselves alone (selfish option) or to keep an amount of money for themselves and also share an amount of money with the *N*th person on their list (generous option). The selfish options ranged from keeping \$185 to \$65 in the order of decreasing \$10 increments, and the generous option always kept a set amount of money (either \$85, \$75, or \$65) and sharing that same amount with the *N*th person. Each trial lasted 5 s and the interval between trials was randomly jittered with an average of 4 s (jittered ±1 s). The ordering of the blocks was pseudo-randomized. Stimuli consist of text only in which participants chose between two options: "\$XX for YOU" or "\$XX for YOU and \$XX for N=##", where ## indicated the number of the social distance. Participants were informed that their decisions would result in real outcomes. One of each participant's choices was randomly chosen and implemented such that the participant was paid 10% of the amount they selected to receive themselves, and, if a generous choice was selected, the other person in question also was contacted and disbursed 10% of the amount the participant selected for them.

Table 1. Hyperbolic mixed-effects model results for group differences in social discounting.

N = 77	*b*	SE	95% CI	T	*P*
Level 1					
υ_1, mean amount to forgo for $N = 1$	76.06[a]	1.56	73.01, 79.12	48.70	<0.0001
log*k*, mean discounting rate	−4.52[a]	0.43	−5.35, −3.69	−10.67	<0.0001
Level 2					
LKM controls >Wait-list controls	0.13	0.60	−1.05, 1.31	0.21	0.830
Altruists >Wait-list controls	−1.65[b]	0.67	−2.96, −0.33	−2.45	0.015

Note. *b* represents unstandardized coefficients. 95% CI indicates lower/upper limits of a confidence interval. Groups are indicator variables relative to the reference group (wait-list controls).
[a]$P < 0.001$.
[b]$P < 0.05$.
LKM, loving-kindness meditation.

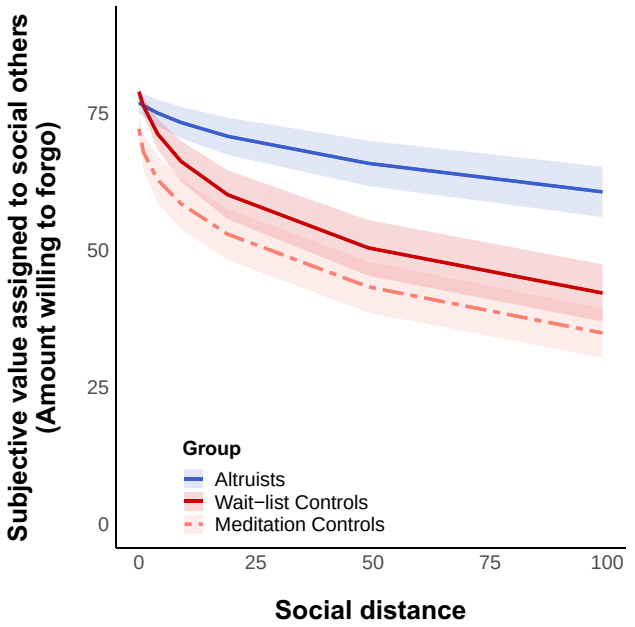

Fig. 2. Social discounting in altruists, meditation controls, and wait-list controls. *Note.* Group labels are listed in descending order of mean discounting slope across groups (log*k*). Ribbons reflect the 95% CIs around each group mean.

According to this calculation, the temptation to be selfish increases as the value of the potential self-reward increases relative to the amount a participant would sacrifice by making a generous decision (7). For example—consider a choice between \$125 and \$75/75 for $N = 20$—if the participant's model estimated amount willing to forgo for their $N = 20$ was \$65, then the modulator is \$125–\$65 = 60 (lower temptation bias). Considering a choice between \$125 and \$75/75 for $N = 50$: If the participant's model estimated amount willing to forgo for their $N = 50$ was \$25, then the modulator is \$125–\$25 = 100 (higher temptation bias). We thus fit a first-level GLM to BOLD responses that included two additional regressors that assessed the amplitude modulation of responses as a function of this temptation bias for generous and selfish choices.

In contrast to observations in their prior sample ($N = 23$) that activity in right TPJ increases as a function of the temptation to be selfish when a generous choice is made, we found no differences in TPJ either across the full sample or between groups. Only when small-volume correcting in a 6-mm sphere around the right

TPJ [Montreal Neurological Institute (MNI) coordinates: x = 51, y = −49, z = 34], we found an effect in this region—except our findings indicated that altruists exhibited the opposite effect as previously reported in ordinary adults (7) (Supplementary Fig. S3). This finding contradicts the hypothesis that altruists' generous choices reflect the exertion of greater effortful control to overcome a selfishness bias. Similarly, a psychophysiological interaction (PPI) analysis, also repeating analyses by Strombach, that assessed functional connectivity with this seed region in right TPJ also provided no evidence for a role of TPJ in modulating other regions during generous versus selfish choices. We found no regions that were significantly functionally correlated with right TPJ during generous versus selfish choices.

Table 2. Linear mixed-effects model results for group differences in difficulty.

N = 77	*b*	SE	95% CI	T	*P*
Intercept	6.64[a]	0.10	6.44, 6.84	65.28	<0.0001
LKM controls >Wait-list controls	0.07	0.14	−0.21, 0.36	0.51	0.61
Altruists >Wait-list controls	−0.08	0.15	−0.36, 0.21	−0.52	0.61
Social distance (N)	−0.002[b]	0.14	−0.003, −0.001	−3.26	0.001
LKM controls × N	−0.004[a]	0.001	−0.01, −0.003	−5.43	<0.0001
Altruists × N	−0.0001	0.001	−0.002, 0.001	−0.22	0.83

Note. *b* represents unstandardized coefficients. 95% CI indicates lower/upper limits of a confidence interval. Groups are indicator variables relative to the reference group (wait-list controls).
[a]$P < 0.001$.
[b]$P < 0.05$.
LKM, loving-kindness meditation.

Overall, results revealed a main effect of choice type (generous versus selfish choices), but no main effect of group. Across altruists and controls, generous choices in contrast to selfish choices recruited several regions previously linked to prosocial decision-making (36), including left medial PFC, right superior temporal gyrus, and left temporal pole. Selfish choices preferentially recruited regions that included left inferior frontal gyrus, bilateral presupplementary motor area, and bilateral posterior cingulate gyrus (Fig. 4; Supplementary Fig. S4; Supplementary Tables S5 and S6). But no whole-brain differences were observed across groups, suggesting that altruists and controls engage similar processes when making generous decisions.

We next tested the alternate hypothesis that altruists' heightened generosity may reflect how they encode the subjective value of others' welfare during decisions. Evidence in favor of this hypothesis would reveal divergent profiles of activation between altruists and controls corresponding to the subjective value that they assign to social others' welfare (v) (N = 77; Supplementary Table S1). To test this, we fit a first-level GLM to BOLD responses to estimate the extent to which the amplitude of responses was modulated by the model-fitted value v (amount participants were willing to forgo for each social other; see Materials and methods). Because a second-level group analysis again found no differences between LKM and wait-list controls, participants were combined. Group-level analyses yielded the only results to reveal distinct profiles of activation in altruists and controls. We found multiple regions in which neural activity encoded subjective valuation of others' welfare differently for altruists than controls. These regions notably included two regions that we predicted: bilateral amygdala and left rostral anterior cingulate gyrus (BA 32) (Figs 5 and 6; Supplementary Fig. S5; Supplementary Tables S7

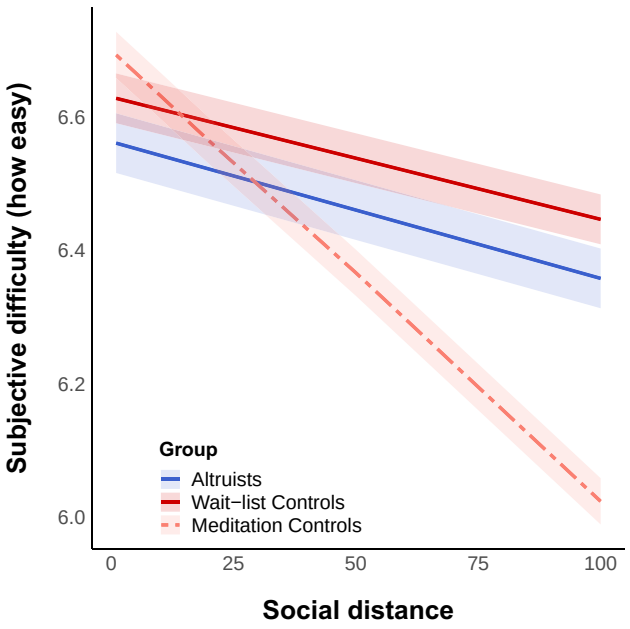

Fig. 3. Linear mixed-effects model results for group differences in decision difficulty as a function of social distance. *Note.* Group labels are listed in ascending order of mean slope across groups. Ribbons reflect the 95% CIs around each group mean. For ease of comparison between social discounting in Fig. 2 and difficulty ratings, we reversed-coded the difficulty ratings (i.e. subtracted each score from 7) such that a higher score indicated greater ease rather than greater difficulty for this plot.

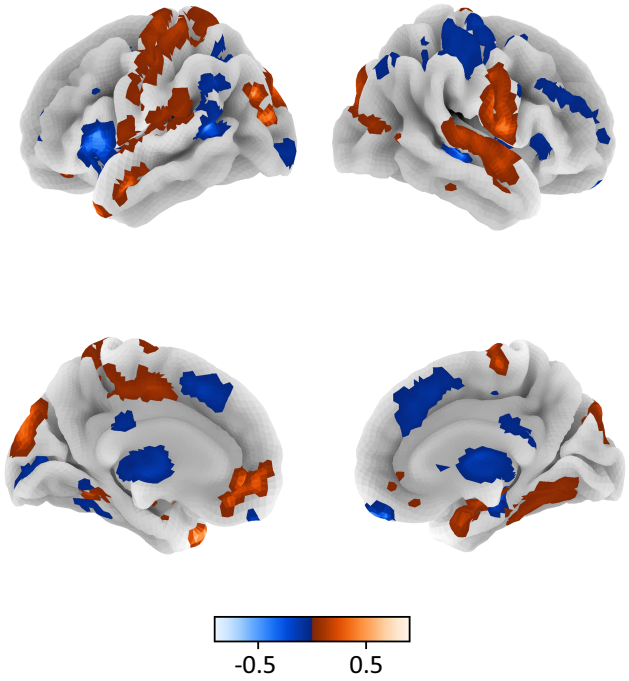

Fig. 4. Neural activation for generous versus selfish decisions across the full sample. *Note.* Colormap corresponds to the difference in neural activation (FDR-corrected $q < 0.05$; height threshold $t > 6.5$ for visualization purposes) between generous and selfish choices across the full sample. Surface corresponds to the following key: left lateral = top left; right lateral = top right; left medial = bottom left; right medial = bottom right.

and S8), regions in which neurons that encode the value of outcomes for others have been observed in nonhuman animal studies (29–31, 37). Mean magnitude differences were assessed in these regions by fitting the same GLM to BOLD responses including regressors of interest that corresponded to social other blocks (including nuisance regressors, but excluding the amplitude modulator of subjective valuation assigned to others' welfare). Results revealed that, relative to controls, altruists exhibited more uniform mean activity in left rostral ACC and right amygdala as the value assigned to social others increased (by contrast, in controls, activation increased in left rostral ACC and decreased in right amygdala as social value increased). Other regions in which altruists showed more uniform patterns of activity than controls as social value increased included left superior frontal gyrus (BA 8), left precuneus (BA 31), and right inferior occipital gyrus (BA 18). Importantly, these patterns parallel altruists' more uniform generosity across social distances (Supplementary Fig. S6).

Altruists' value-related activity also differed from controls in a nonuniform manner in several regions that were not predicted (Supplementary Fig. S6). These included left superior temporal sulcus (BA 21), right fusiform gyrus (BA 19), left anterior inferior temporal gyrus (BA 20), right superior medial frontal gyrus (BA 9/10; dorsomedial PFC), left inferior frontal gyrus (triangularis; BA 45), left middle frontal gyrus (BA 8), right inferior temporal gyrus (BA 20), parahippocampal gyrus, right inferior frontal gyrus (BA 44), right putamen, left supplementary motor area (BA 6), and right cerebellum (lobule IV–V). To address potential deviations from parametric assumptions due to the unequal sample sizes from pooling control participants, we employed a bootstrap resampling approach to test in which regions in which neural activation encoded subjective valuation of others' welfare differently between 5,000 bootstrap samples of altruists and controls. A large majority of identified regions using this approach coincided strongly with resulting maps from parametric statistics (Supplementary Fig. S7, Supplementary Table S9). We did not find value-related group differences in insula, temporal gyrus, or TPJ.

To test whether our neural findings could reflect group differences in the processes engaged in discounting decisions more broadly (as opposed to specifically social discounting), we conducted a conjunction analysis between our contrast map assessing regions which in altruists differed from controls in subjective valuation encoding of social distance (Fig. 5) to three Z-statistic maps assessing subjective valuation encoding of physical effort, probability, and time delays (45). Results revealed that the altruist-control difference in BOLD activation tracking subjective valuation during social discounting only partially overlapped with BOLD activation tracking subjective valuation during nonsocial discounting. Importantly, this conjunction of voxels only partially overlapped with the rostral ACC/medial PFC cluster from the present study and did not overlap with any of the other clusters (Supplementary Fig. S8), suggesting that the processes altruists engage during social discounting are unlikely to stem from differences in discounting more broadly.

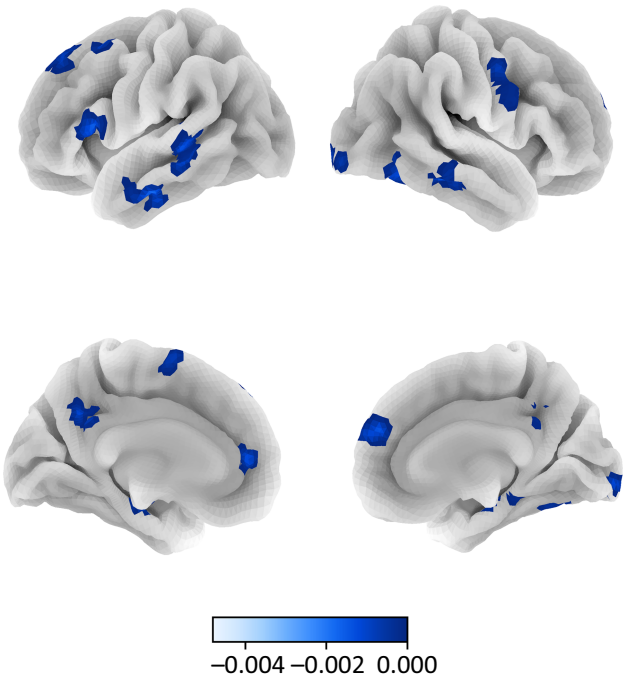

Fig. 5. Regions in which altruists exhibited a different relationship between neural activation and social valuation in contrast to all controls. *Note.* Colormap corresponds to the difference in subjective social value-related activity (FDR-corrected $q < 0.05$) between altruists and controls. Surface corresponds to the following key: left lateral = top left; right lateral = top right; left medial = bottom left; right medial = bottom right.

Discussion

Combining both ecological and experimental approaches, we conducted the first pre-registered assessment of the neural underpinnings of generosity for socially distant others in real-world extraordinary altruists and demographically matched controls, approximately half of whom completed an LKM intervention. We tested two alternative hypotheses regarding the origins of these patterns: (i) that altruists' generosity for distant others reflects their more successfully overcoming an egoism bias (7, 8) or (ii) that their generosity reflects how they represent the subjective value of distant others' outcomes. Our findings provided no support for greater effortful suppression of selfish preferences in altruists. Instead, we found that group differences in social discounting corresponded to activity in a constellation of regions reflecting the subjective value placed on others' welfare (v), including rostral ACC and amygdala. Relative to typical adults, value-related activation patterns were more uniform across social distances among altruists—a pattern consistent with altruists' more uniform generosity for close and distant others. We thus conclude that altruists are more generous to strangers because of how they encode the subjective value of their welfare.

Recent research in humans and nonhuman animals has implicated medial PFC, specifically rostral ACC, and amygdala, in prosocial decision-making. In nonhuman primates, ACC responding is selective for others' rewards (46) and is also found to be critical for value-based social choices (29, 47). Excitotoxic lesions to this region also impair prosociality (29). In rodents, populations of cells in these regions have been found to compute subjective valuation of others' outcomes, with enhanced coordination in firing patterns between neurons in the ACC and amygdala facilitating prosocial decisions (31). Although functional neuroimaging cannot yield similarly specific results, we found for the first time that in typical adults, mean activity in left rostral ACC increased, and mean activity in amygdala decreased as the subjective valuation of others' welfare (v) increased during decision-making. By contrast, altruists in our study exhibited more uniform mean activity in both of these regions than controls, paralleling their more

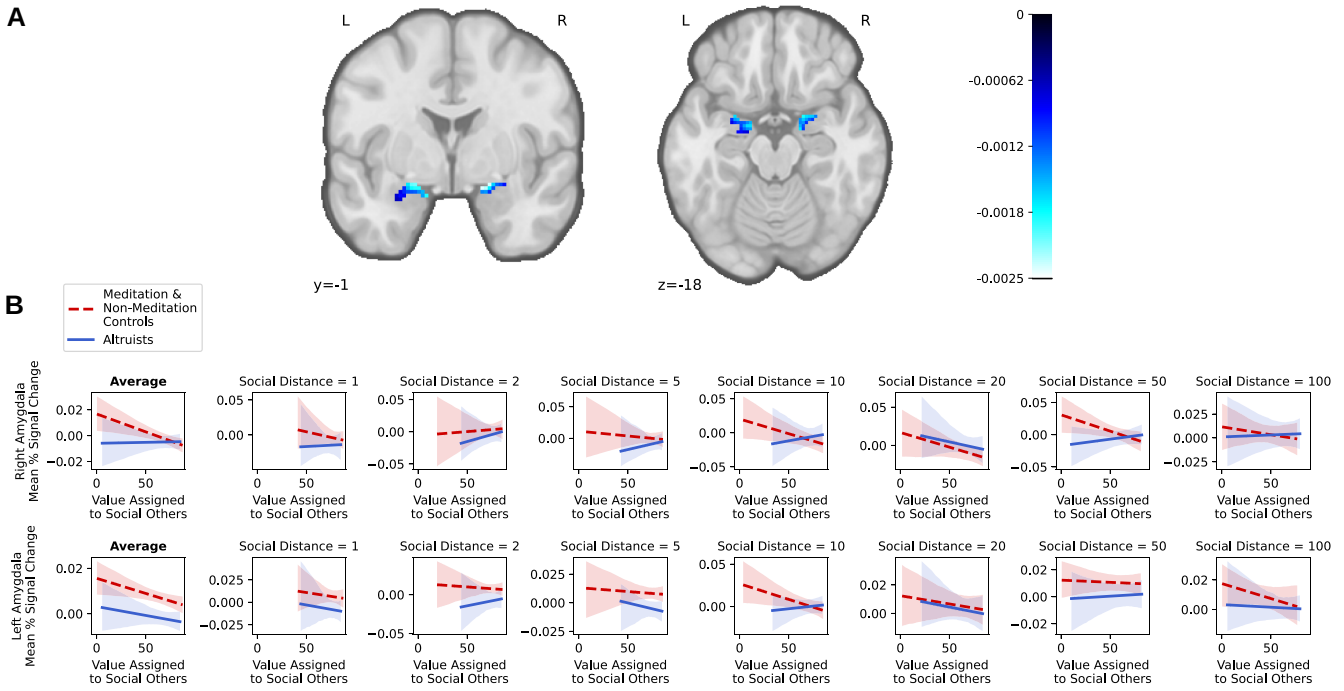

Fig. 6. Amygdala voxels in which altruists exhibited a different relationship between neural activation and social valuation in contrast to all controls. *Note.* A) Colormap corresponds to the difference in subjective social value-related activity between altruists and controls in a pre-registered a priori region of interest (BA; FDR-corrected $q < 0.05$). A smaller cluster of voxels in this region also survived multiple-comparison correction across the whole brain (FDR-corrected $q < 0.05$). B) Plots of mean activation magnitude in these regions as a function of subjective social valuation, group, and social distance. The average plot corresponds to the signal change by value aggregated across all social distances.

consistent generosity across social distances (i.e. shallower discounting slopes) (11, 12).

There are a few reasons this might be the case. This includes our current interpretation that it reflects more uniform subjective valuation of others' welfare in altruists. It may also reflect another form of implicit or affective responding that is more uniform for various others during decision-making. However, evidence in nonhuman species linking activity in amygdala and ACC to the genesis of prosocial preferences supports the former possibility. These patterns are generally consistent with evidence from prior neuroimaging research in humans as well. For example, ACC activation has been found to track the value of others' rewarding outcomes, the probability that others will receive rewards, and prediction errors after observing the outcome of others' choices (48–51). And previous research in extraordinary altruists finds both increased amygdala volume in altruists and increased activation in this region corresponds to altruists' empathic responding to strangers' distress (19, 52). These findings have been interpreted to suggest that variation in the structure and function of amygdala and ACC may underpin a "caring continuum" (3), ranging from highly uncaring populations (such as those with psychopathy) (53–56) to highly caring populations (17–19).

Our findings do not support the notion that individual variation in costly real-world generosity reflects variability in overcoming an egoism bias, as previously suggested by models of prosocial restraint (9, 10). We found no group differences for generous versus selfish decisions. We instead found similar activation patterns across the full sample reliably linked to prosocial versus selfish choices (36). In addition, altruists exhibited activity in right TPJ for generous versus selfish decisions that *decreased* with a greater temptation to be selfish. Given that we also did not observe within-subject differences among controls, our finding could be related to methodological differences between our task and the task used previously (7, 8). Notably, we opted not to include a visuo-spatial depiction of social distance (a number line) in our paradigm; we instead used a simple numeric indicator of social distance. We made this choice because tasks using the number line have reported results in the right angular gyrus (BA 39), a region known to encode information about spatial, temporal, and social distance (57). It is possible that increased activation in this region during generous versus selfish decisions in previous tasks reflected the spatial location of avatars on the number line, because selfish choices would be correlated with avatars appearing toward the right of the screen, whereas more generous choices would be correlated with avatars appearing towards the left of the screen (given the hyperbolicity of participants' choices). In discounting tasks that use a number line, it is thus difficult to discern whether activation in this region reflects changes in spatial and social distance versus overcoming egoism bias during generous choices, or both. It should also be noted that recent neurostimulation research on prosocial decision-making corroborates our findings, showing that inhibitory transcranial magnetic stimulation to right TPJ does not alter selfish choices (58). Taken together, our results are instead more consistent with neuroimaging evidence that activity in TPJ reflects a conflict between personal motives rather than a conflict between selfish and altruistic motives (59, 60). Based on the evidence reported previously (12) and in the present study, it is also unlikely that our observed differences in behavior or neural activation tracking social value are related to extraordinary altruists' ability (or lack thereof) to differentiate between elements on a rank-ordered list. In the present study, altruists did not differ from controls in their ratings of nominated social others in terms of emotional closeness nor in terms of relationship types across social distances. Despite similar

judgments about these people, they nonetheless exhibited divergent patterns of generous behavior as a function of social distance.

Another potential explanation for why our results differ from those reported previously may be related to the possibility that their sample was more selfish as a function of a social distance relative to than our control samples (7). The estimated discounting function of the current control sample is similar to previous studies (11, 12, 61). But, it is also possible that our control sample showed slightly greater prosocial preferences (despite careful recruitment and matching to the altruist sample on various demographics) than in some prior studies. This could be for several reasons. First, the study was conducted during the global pandemic, which introduced several challenges related to participant recruitment. Thus, the controls that participated in our study may be more generous than typical controls, as prosocial behavior increased during this time period (62, 63). It could also be due to meditation study recruitment drawing participants from less selfish populations. Lastly, it is possible that this difference is due to potential geographic or cultural differences in social discounting; noting that work has shown consistent social discounting between rural and urban backgrounds in China (24) and between Germany and China (25), but none between the United States and Europe.

Our results also found that altruists also exhibited value-related activity that differed from controls in a network of other regions that varied nonuniformly. These included the right dorsomedial PFC, in which altruists' activation generally decreased while controls' activation generally increased as a function of social value. This finding in controls is consistent with work showing that dorsomedial PFC activity during social judgment tasks increases with prosociality (i.e. increases as the value of others' outcomes increases), suggesting that helping others may stem from the spontaneous ability to understand others (64). And in dorsal striatum (right putamen), altruists exhibited increasing activity in contrast to controls, who exhibited decreasing activity, as the subjective value placed on others increased. Because activity in this region increases during reward-based decision-making as a function of resulting personal reward magnitudes (65, 66), this result may reflect how rewarding individuals find prosocial actions on average. The observed pattern in this region is consistent with altruists finding prosocial decisions more rewarding than controls, a possibility that future research could test.

We found that generous decisions are reported to be subjectively easier on average across all groups, which is generally more consistent with the conclusion that generosity is driven by prosocial preferences than by inhibition of selfishness. In addition, we found that the strength of this relationship was stronger for both altruists and LKM controls relative to wait-list controls. One interpretation of this finding is that people who have greater prosocial preferences (e.g. altruists) find it easier to make prosocial decisions, in line with existing theories (16). However, it is important to note that subjective difficulty ratings cannot be assumed to index selfish impulses because a decision can be difficult for many reasons other than an implicit bias to be selfish. Recent work has demonstrated that computational models of altruistic behavior (e.g. sequential sampling models) can account for behavioral and neural effects attributed to self-control and the value of generosity, with generous choices slower than selfish choices when the relative weight placed on one's own welfare is higher and faster when the weight placed on others' welfare is higher (60). Although our task was not optimized for this model, our data are consistent with this work. Qualitatively, the difference in response times between generous and selfish choices for controls decreases as social distance increases for controls but is fairly consistent across social distances for altruists. Taken together, altruists may be behaviorally responding similarity for all social others because the weight they place on all others' welfare is similar to the weight placed on their own. However, future work should test this directly.

Finally, our study also indicated that undergoing an 8-week LKM training protocol aimed at increasing care and kindness towards others at varying social distances did not reduce social discounting, despite our predictions and despite some evidence linking similar training to laboratory-based prosociality (13, 14). Instead, LKM training made decisions for distant others more difficult for participants. We do not believe these null findings reflected the quality of the training modules, which were explicitly designed to elicit care for others at varying social distances, and which were created by one of the world's most renowned LKM training experts, Sharon Salzberg (39). Instead, these findings suggest that although LKM training may promote some forms of prosociality, it may not change how people fundamentally value distant others' welfare, a conclusion consistent with findings from a recent review and meta-analysis (44) and noteworthy given increased popularization of meditation techniques in both scientific and nonscientific contexts. While LKM participants in this study completed nearly all of the training sessions and consistently reported enjoying them, we did not find either behavioral or neural evidence that these sessions changed how they encoded the subjective value of others' welfare. It is possible, however, that LKM might increase other types of prosocial outcomes such as responding empathically others' pain or distress. Future work could assess the effects of the present LKM training on social discounting in a loss context (67). In the present LKM training, participants were also free to direct loving-kindness towards anyone who met the criteria for a given week. It is possible that identifying targets that overlapped with those nominated for the decision-making task could have reduced social discounting. Future work should examine this possibility.

Our findings should also be considered in light of certain limitations. For one, we did not employ an active meditation control group. This was necessary to allow direct comparison between altruists (who did not undergo meditation training) and wait-list controls. It should be noted, however, that the use of a wait-list control group generally *increases* meditation intervention effect sizes relative to active control conditions (44), so this feature of our paradigm is unlikely to be confounding. Consistent with this conclusion, a similar LKM training protocol has been used (the same recordings, but over 4 weeks instead 8) with an active control condition (progressive muscle relaxation meditation) and produced similar null effects related to perceived social connectedness and real-world prosocial behavior (68). In addition, we did not collect baseline discounting or neuroimaging data prior to LKM training in order to reduce bias in our control groups. This makes it possible that our LKM group could have exhibited increased social discounting preferences at baseline such that within-group changes after the 8-week program were missed. However, our control groups were carefully recruited concurrently, randomly assigned to each condition, and matched on multiple psychosocial and demographic features to each other and to altruists, so we do not believe this is likely.

It is important to note that matching controls to extraordinary altruists on demographics led to a fairly demographically homogenous sample of participants—likely due to the fact that kidney

donors are often selected based on factors related to age, sex, and health, among other factors that can represent barriers to healthcare. However, other populations of extraordinary altruists are more demographically diverse than the present sample (e.g. bone marrow donors, heroic rescuers) (11, 69, 70) and are also characterized by decreased social discounting (11). It should also be noted that because all altruists were recruited after donating, our results are not aimed at prospectively predicting extraordinary altruism. Our results instead characterize divergent behavioral and neural patterns among individuals who have already engaged in extraordinary acts of altruism. It is possible that altruists' divergent behavioral and neural patterns reflect their awareness that they were recruited for a study on altruism (an issue relevant to all research on special groups, including all clinical samples) and thus are responding to experimental demand. However, several lines of evidence argue against this possibility. Altruists do not score highly on self-report altruism scales that are more straightforwardly related to altruism and thus arguably less susceptible to demand effects (11, 71). In addition, we do not find that the characteristics actually differentiating altruists from controls are those that the average person predicts will set them apart (11), which suggests altruists are not simply responding in altruistic stereotype-consistent ways.

Prior work has established that decreased social discounting with monetary costs captures an important latent variable relevant to valuation of others' welfare and to various ecologically valid forms of altruism (11). However, it should be noted that monetary costs are only one form of costs that can be incurred from altruistic behavior. Willingness to undergo physical effort costs (72, 73) could also be closely linked to the costly helping related to organ donation; future work could test the extent to which altruists will exert more effort in the lab to benefit social others' welfare.

Despite these limitations, these results resolve key questions about the origins of real-world costly altruism. We find the first evidence that real-world altruists' generosity for distant others reflect their true valuation of distant others' welfare, not overcoming selfish biases. More, this work adds to the emerging evidence linking value-related calculations in rostral ACC and amygdala to prosocial decision-making across species and shows that activation patterns in these regions diverge in real-world extraordinary altruists. Altruists may engage in costly generosity for socially distant others due to their relatively unbiased value-related neural responses across both socially close and distant others in these and other regions, such that they respond generously to both distant and close others.

The finding that altruists' generosity reflects their true valuation of distant others' welfare may explain why even a custom-designed 8-week LKM intervention was unable to reproduce their generous choices or neural activation profiles in a sample of typical adults. Genuinely valuing the welfare of strangers may reflect powerful life experiences and fundamental beliefs about other human beings that cannot be quickly or easily reproduced (74, 75). Interventions that can alter how the subjective value of distant others' welfare is encoded may be more successful.

Materials and methods

All recruitment procedures described below (e.g. inclusion/exclusion criteria, hypotheses) were outlined in our pre-registration (https://osf.io/u8adg) and approved by the Georgetown University Institutional Review Board.

Participants

All participants were provided informed consent to participate in a study on the behavioral and neural correlates of altruism. Eighty participants completed the social discounting portion of our study, which surpassed our pre-registered target sample size. Our pre-registered sample size ($N = 78$) was calculated with G*Power 3.1 (76) using previously published social discounting data in altruistic kidney donors and matched controls (12) to detect group differences in behavior with large effect sizes (≥ 0.80) while maintaining an alpha of 0.05. As a post hoc calculation of power, we used the whole-brain, unthresholded t-statistic map contrasting generous and selfish decisions from a prior neuroimaging study (7) to estimate the sample size we would need to generate a similar map. Using neuropowertools (http://neuropowertools.org/) (77), we determined that sample sizes of $N = 62$ and $N = 31$ would be required to achieve ≥ 0.80 power according to Random Field Theory correction and no correction, respectively. Our final sample sizes meeting inclusion criteria ($N = 62$, $N = 77$) exceeded these estimates and would have ≥ 0.95–0.99 power according to these simulations. Nondirected kidney donors from across North America and typical adults from the Washington D.C. metropolitan area were recruited simultaneously. All altruistic kidney donors were verified as having volunteered to donate a kidney to a stranger. Control participants were verified to have never donated an organ to another person and were matched to altruists on age, gender, and IQ. Among control participants, we employed a randomized wait-list control using blocked randomization with stratification by age (above/below 40 years old) and gender to select the LKM training group.

All participants did not have any current psychological illness (e.g. depression, bipolar disorder, psychosis) as assessed by scores above clinical thresholds on the SCL-90 (78) or use of psychotropic medication, did not have any history of head injury followed by 10 + minutes loss of consciousness, or neurological disorder (epilepsy, Tourette's disorder, etc.), did not have previous formal meditation training and/or current meditation practice, did not have any contraindication to safe MRI scanning (e.g. ferrous metal implants, pregnancy), and did not have an IQ score less than 80 as assessed by the Kaufman Brief Intelligence Test—Second Edition (79).

Eighty participants meeting the criteria above completed three runs of a social discounting task during fMRI. Of these participants, one participant was excluded due to excessive motion during scanning (more than 15% of volumes greater than 0.5 mm framewise displacement), one participant was excluded because they asked to change over 20% of their choices after completing the scan (indicating that they did not understand the task), and one participant was excluded because they missed more than half of all choices for a social other. Our final sample included 77 participants (26 altruists, 25 controls randomized to an 8-week meditation training, 26 wait-list controls; 62.34% female; mean age = 41.49 years, SD = 7.65 years).

Selection of social others

Prior to testing, subjects were asked to provide names and photographic images of individuals from their own social network who correspond to each of the social distances, $N = 1$, 2, 5, 10, 20, and 50. They were provided the following instructions: "Think about a list of 100 people in your social environment, with person #1 being the closest relationship to you and person #100 being someone who you do not know. Please provide us the full name and photos for the person you would consider #1, #2, #5, #10, #20,

#50." They were also asked to supply a nonoccluded image of the person's face looking into the camera and wearing a neutral or mildly positive expression (not analyzed in the present study) (80). Although within social networks, it is possible for more than one person to represent a given social distance, subjects only selected one person for each distance (7). The experiment also included social distance level 100, but this level represented strangers whose name may not be known (21), so subjects were not required to provide a name or photo for this level. Thus, seven social distances were included.

To validate that groups ranked people similarly across social distances, we administered a pen/paper questionnaire that evaluated their emotional closeness towards each nominated social other (#1, #2, #5, #10, #20, #50) and for two strangers of the opposite sex (#100) using a Likert scale ("How emotionally close do you feel with this person?" 1 = Not at all; 9 = Very much). These measures were completed by 79 out of 80 participants. Ratings for the strangers were averaged prior to analyses. Results from a linear mixed effects model found no significant differences between meditation controls ($\beta = 0.32$, SE = 0.25, 95% CI = [−0.17, 0.81], $P = 0.20$) or altruists ($\beta = -0.22$, SE = 0.54, 95% CI = [−0.71, 0.28], $P = 0.39$) relative to wait-list controls (Supplementary Table S10; Supplementary Fig. S9). Means and standard errors are also reported across groups for ratings on emotional closeness, trust, positive affect, and negative affect, as well as relationship duration and interaction frequency (Supplementary Table S11). Participants also described their relationship with each person nominated via a free response prompt, which was coded according to thematic analysis (81). Six relationship categories emerged in addition to the stranger category: friend, family member, coworker, spouse, dating partner, and acquaintance. A series of chi-square tests of independence was performed at each social distance to examine the relationship between the group and the relationship type of the nominated targets. The proportion of nominations did not differ among groups for any of the social distances (Supplementary Fig. S10): At Social Distance #1, $\chi^2(8, N = 79) = 10.22$, $P = 0.25$; Social Distance #2, $\chi^2(10, N = 79) = 8.58$, $P = 0.57$; Social Distance #5, $\chi^2(8, N = 79) = 6.19$, $P = 0.63$; Social Distance #10, $\chi^2(8, N = 79) = 6.25$, $P = 0.62$; Social Distance #20, $\chi^2(6, N = 79) = 5.03$, $P = 0.54$; Social Distance #50, $\chi^2(6, N = 79) = 11.38$, $P = 0.08$.

LKM training

The 8-week LKM training program was custom-designed for the purposes of this study by renowned meditation expert, Sharon Salzberg. In a randomized control trial, this training followed a structured approach during which loving kindness is directed toward a different target each week: (i) the self, (ii) a benefactor, (iii) a friend who is doing well, (iv) a neutral person, (v) a friend who is struggling, (vi) a difficult person, (vii) groups, and finally (viii) to all beings without distinction. Participants completed five training sessions per week. The first session on a given week lasted 27 min and the remaining four sessions lasted 15 min each (87 min per week; 11.6 h in total over 8 weeks).

All control participants were recruited into the study at least 8 weeks prior to neuroimaging and were randomized into groups upon enrollment. Meditation participants were tasked to complete five trainings per week for 8 weeks prior to their neuroimaging scan date (40 total). Web-based confirmation of participation among LKM controls included in analyses indicated that they completed at least 77.5% of trainings (range = 77.5–100%, mean = 89.6%, SD = 0.07%).

Social discounting paradigm

During scanning, participants made 189 decisions across three independent runs involving the seven social distances described above. The task consisted of 21 blocks with nine trials each. In keeping with established procedures (7), participants were prompted to imagine the seven possible individuals on the list ($N = 1, 2, 5, 10, 20, 50$, or 100) and make nine dichotomous choices about keeping and/or sharing amounts of money with each. For each trial, participants were asked to indicate if they would prefer to keep an amount of money for themselves alone (selfish option) or to keep an amount of money for themselves and also share an amount of money with the Nth person on their list (generous option). The selfish options ranged from keeping $185 to $65 in the order of decreasing $10 increments, and the generous option always kept a set amount of money (i.e. keeping $85, $75, or $65) and sharing that same amount with the Nth person (i.e. sharing $85, $75, or $65). Participants made their selections by pressing buttons held in the left and right thumbs. Each trial lasted 5 s, and the interval between trials was randomly jittered with an average of 4 s (jittered ±1 s). The ordering of the blocks was pseudo-randomized. Stimuli consisted of text only in which participants chose between two options: "$XX for YOU" or "$XX for YOU and $XX for N=##", where ## indicated the number of the social distance. Participants were trained on the task before entering the scanner using instructions sheets and verbal testing. They were asked to confirm that they knew the identity of each N before scanning (see Supplementary Table S12 for instructions).

Participants were informed prior to the task that payments were yoked to task responses so that the experiment avoided deception and met the full standards of behavioral economics. Based on their actual choices, participants were paid 10% from a randomly chosen trial, and, if the participant chose the generous option, the other person in question also received 10% of the amount selected for them. If relevant, participants were asked after the task to provide an email address or phone number for contacting that person so that payment could be made via PayPal or Amazon gift card. If the randomly chosen trial featured a person at social distance 100, a random phone number generator was used to select a person to receive the payment using Venmo.

Post-scan questionnaire

After the scanning session, participants' answers to the $75 magnitude questions (63 trials) during the social discounting task were presented back to them. Participants then rated how difficult each decision was for each trial on a scale from 1 (very easy) to 7 (very difficult). Furthermore, each participant completed a questionnaire regarding each of the targets selected during the tasks.

Behavioral modeling

Following standard practice (82, 83), we estimated an "indifference point" when each respondent switched from selfish choices to sharing with a given person number on the list for each of the different generous options (sharing $85, $75, or $65; referred to as the "generous option blocks" below). The indifference point was calculated for each social other (N) by solving for the value at which the probability (P) between choosing a selfish versus generous option was 50% via binomial logistic regression:

$$P(\text{share})_{\text{trials}} = \frac{\exp(b_0 + b_1 \times \text{Amount}_{\text{trials}})}{1 + \exp(b_0 + b_1 \times \text{Amount}_{\text{trials}})},$$

$$\frac{P}{1-P} = \exp(b_0 + b_1 \times \text{Amount}_{\text{trials}}).$$

Setting $P = 0.5$ yields:

$$\ln\left(\frac{P}{1-P}\right) = \ln\left(\frac{0.5}{1-0.5}\right) = b_0 + b_1 \times \text{Amount}_{\text{indifference trial}},$$

$$\ln(1) = 0 = b_0 + b_1 \times \text{Amount}_{\text{indifference trial}},$$

$$-\frac{b_0}{b_1} = \text{Amount}_{\text{indifference trial}}.$$

If the selfish option was chosen for all trials in a block, the indifference points were assumed at \$80, \$70, or \$60 for the \$85, \$75, or \$65 generous options, respectively. If the generous option was chosen for all trials in a block, the indifference points were assumed at \$170, \$160, or \$150 for the \$85, \$75, or \$65 generous options, respectively. Amounts willing to forgo (v) were calculated by subtracting the generous option magnitude (\$85, \$75, or \$65) from each indifference point, and then averaging across the generous option blocks. Thus, we had seven "amount willing to forgo" (v) observations corresponding to one of seven social others (N) for every participant (i).

Social distance (N) was centered at $N = 1$ (decision towards closest other); groups were coded as indicator variables with wait-list controls entered as the baseline group. We used the nlme package v3.1.160 (https://cran.r-project.org/package=nlme) (84) in R version 4.2.2 to fit data to a hyperbolic mixed-effects model using maximum likelihood estimation, which allowed both intercepts (v_{0i}), and discounting rates ($\log k_i$) to vary across participants with group entered at level 2. Because discounting rates are non-parametrically distributed, we employed a variation in the hyperbolic discounting function to improve model convergence, estimating $\log k$ rather than k (85):

$$v = \frac{v_{0i}}{1 + k_i \times N_i} = \frac{v_{0i}}{1 + \exp(\log k_i) \times N_i}.$$

The model assessed how discounting rates ($\log k_i$) differed across altruists and LKM controls relative to controls:

Level 1 (hyperbolic function):

$$v_{Ni} = \frac{v_{0i}}{1 + \exp(\log k_i) \times N_i} + e_{Ni}.$$

Level 2:

$$v_{0i} = b_{00} + r_{0i},$$

$$\log k_i = \beta_{10} + \beta_{11} \times \text{Altruists}_i + \beta_{12} \times \text{Meditation controls}_i + r_{1i}.$$

Bayesian model comparison using the Bayesian Information Criterion confirmed that the hyperbolic model outperformed a null (intercept-only) model, linear model, and exponential model (Supplementary Figs S11 and S12), as previously demonstrated (11, 12). To validate the estimated $\log k$ parameters across participants, we calculated the correlation with a model-agnostic measure of social discounting: area under the curve (AUC), which does not assume hyperbolicity in responding (86). AUC was calculated for each participant by normalizing amount willing to forgo (v) as a percentage of maximum v, normalizing social distance (N) as a percentage of maximum N, connecting the crossover points by straight lines, then summing the areas of the trapezoids formed. As generosity increases, $\log k$ decreases and AUC increases (both representing shallower slopes). We observed a statistically significant negative correlation (Spearman $\rho = -0.89$), with $\log k$ capturing more variation across participants (see Supplementary Fig. S13 for relationship and mean AUC across groups), validating our use of $\log k$ to index discounting rates.

Response time modeling

We tested how log-transformed response times varied as a function of group and choice using a linear mixed-effects model using the glmer function within the lme4 v1.1.31 package in R version 4.2.2 that allowed intercepts to vary across participants.

Difficultly rating modeling

We tested how choice difficulty varied as a function of social distance using a linear mixed-effects model (again using the nlme package in R version 4.2.2) that nested ratings within social distance blocks and allowed slopes and intercepts to vary across participants. We also tested how choices varied as function of difficulty and decision type (generous versus selfish) using a logistic mixed-effects model using the glmer function within the lme4 v1.1.31 package in R version 4.2.2 that allowed intercepts to vary across participants. See Supplementary Figs S14 to S17 for plots of difficulty as a function of trials within block.

MRI data acquisition

All scanning took place at the Center for Functional and Magnetic Resonance Imaging at Georgetown University, which supplies a 3 T Siemens Magnetom Prisma Fit MRI scanner. A 64-channel head coil was used for data collection. Briefly, following a localizer, we collected one double-echo field map (phase encoding direction = A >> P). Functional data were first collected four runs of a task not discussed in this manuscript, which measured neural responses to faces of each of the nominated others at varying social distances in the social discounting task. These runs were followed by a magnetization-prepared rapid gradient-echo (MPRAGE) high-resolution T1-weighted structural scan: 1 mm isotropic voxels, repetition time (TR) = 2,300 ms, time to echo (TE) = 2.99 ms, phase encoding direction = A >> P, matrix: 270 × 288 × 176 mm. Finally, three runs of functional data during the social discounting task were collected: 3 mm isotropic voxels, 204 mm FoV, TR = 2,000 ms, TE = 30 ms, phase encoding direction = A >> P, slices = 37, accel mode = GRAPPA factor 2, bandwidth 2,450 Hz/Px. All runs were counterbalanced across participants.

Preprocessing

All preprocessing was implemented using fMRIPrep 21.2.1 (87), which includes anatomical T1-weighted brain extraction, anatomical surface extraction, head-motion estimation and correction, susceptibility-derived distortion estimation and unwarping, slice-timing correction, intrasubject registration, and spatial normalization (intersubject registration). All functional-to-anatomical registration outputs were quality-checked prior to analyses. We deviated slightly from our initial preprocessing plan in the pre-registration in favor of using ICA-AROMA in the long-term release of fMRIPrep 21.2.1 to remove artifacts (e.g. motion, physiological noise) and smoothing the data using a 6 mm^3 full width at half maximum (FWHM) Gaussian kernel. All data were also scaled to have a mean of 100 prior to first-level statistical modeling.

All first-level and second-level neuroimaging analyses were performed using AFNI version 22.0.02 (88). We used the 3dDeconvolve function for all first-level analyses and used the 3dMVM function (89) for all second-level group analyses. Family-wise error rate (FWER)-correction was applied by first using AFNI's 3dFWHMx function to estimate the spatial

smoothness of the population-level residuals obtained from 3dMVM and then using a permutation approach to determine cluster-size thresholding via AFNI's 3dClustSim function, which randomizes and permutes input datasets using 10,000 Monte Carlo simulations. This approach was developed to reduce the false positive rate in response to Eklund et al. (2016). Cluster significance was determined using an underlying voxel height threshold of $P < 0.001$ and a cluster forming threshold to control the false discovery rate (FDR) at $P < 0.05$.

Neural activation for generous versus selfish decisions

To assess whether group differed in terms of their univariate activation during generous relative to selfish decisions, we fit a first-level GLM to BOLD data from the 62 participants who exhibited variation in generosity (i.e. included those who did not choose the generous option for >99% of trials, as did 4 wait-list controls, 1 LKM control, and 10 altruists; Supplementary Table S4) that included event-related regressors for selfish choices, generous choices, six motion parameters, white matter signal, and cerebrospinal fluid (CSF) signal. We also included an additional regressor for missed choices among 10 participants. When applicable, we also censored volumes with motion greater than our pre-registered threshold (framewise displacement >.5 mm). This temporal data reduction approach produced two whole brain maps of interest reflecting a participant's percent signal change for generous decisions and selfish decisions. We then submitted these to a group-level analysis to assess whether their contrast (generous minus selfish) varied across groups. Whole brain group-level maps were thresholded using an FDR-corrected $q < 0.05$ and cluster size of 10 voxels and a FWER corrected $P < 0.001$. We also tested whether groups differed in neural activation for generous choices versus baseline (fixation) among all participants with complete social discounting behavior. To do this, we fit a first-level GLM to BOLD data from the 10 participants excluded in the generous versus selfish analysis that included event-related regressors for generous choices, which produced whole-brain maps reflecting percent signal change for generous decisions that were then submitted to a group-level analysis to assess whether activation during generous choices (against implicit baseline) varied across groups.

Neural activity that scales with the temptation to be selfish

Following Strombach et al. (7), we next tested whether activity in right TPJ increases as the temptation to be selfish increases. To do this, for each participant and for each trial, we first computed a "selfish temptation value" by subtracting the subjective valuation placed on others' welfare from the value of a potential selfish reward for a given choice. We then fit another GLM that included two additional regressors that assessed the amplitude modulation of responses as a function of this temptation bias for generous and selfish choices in addition to the event-related regressors for selfish choices, generous choices, six motion parameters, white matter signal, and CSF signal (and a regressor for missed choice among 10 participants). When applicable, we also censored volumes with motion greater than our pre-registered threshold (framewise displacement >0.5 mm). This temporal data reduction approach produced two whole brain maps of interest reflecting the degree to which a participant's brain activity during generous decisions increased as the selfish temptation increases and the degree to which a participant's brain activity during selfish decisions increased as the selfish temptation increases. We again submitted these to a group-level analysis to assess whether their contrast (generous minus selfish) varied across groups. Whole brain group-level maps were thresholded using an FDR-corrected $q < 0.05$ and cluster size of 10 voxels. We also carried out this analysis using small-volume FDR-correction $q < 0.05$ in a pre-registered a priori region-of-interest: a 6-mm sphere around the right TPJ [MNI: x = 51, y = −49, z = 34].

Psychophysiological interaction

We also followed up with a PPI analysis using the same right TPJ sphere as a seed region. To do this, we computed each participant's average time series within this seed and created two interaction regressors between the normalized time series and each type of decisions (i.e. generous, selfish). Our GLM consisted of the following regressors: a physiological regressor (i.e. the entire time series of the right TPJ), a psychological regressor for the onset of the generous choices, a PPI regressor for the generous choices, a psychological regressor for the onset of the selfish choices, a PPI regressor for the selfish choices. The onset and PPI regressors were convolved with a canonical form of the hemodynamic response. The model also included six motion parameters, white matter signal, and CSF signal (and a regressor for missed choice among 10 participants). When applicable, we also censored volumes with motion greater than our pre-registered threshold (framewise displacement >0.5 mm). This temporal data reduction approach produced two whole brain maps of interest reflecting the degree of temporal correspondence (i.e. functional connectivity) to the right TPJ during generous and selfish choices.

To identify regions whose connectivity was higher during generous than during selfish choices, we submitted these to a group-level analysis to assess whether their varied across groups. Whole brain group-level maps were thresholded using an FDR-corrected $q < 0.05$ and cluster size of 10 voxels. We again carried out this analysis using small-volume FDR-correction $q < 0.05$ in a pre-registered a priori 6-mm sphere around the right TPJ [MNI: x = 51, y = −49, z = 34].

Neural activity that scales with the value assigned to others' welfare

Finally, we tested whether neural activity varied as a function of the value assigned to others' welfare (as estimated from the hyperbolic discounting model) differently across groups. To do this, we estimated a new GLM that included the following regressors: a block regressor for decisions about each social other, a regressor that modulated the amplitude of responses as a function of the subjective value participants assigned to the welfare of each social other, six motion parameters, white matter signal, and CSF signal. Because we used the estimated values from the hyperbolic model, we were able to include the full sample ($N = 77$) for this analysis and thus included an additional regressor for missed choices among 13 participants (three more participants than in the previous neuroimaging analyses). When applicable, we also censored volumes with motion greater than our pre-registered threshold (framewise displacement >0.5 mm). This temporal data reduction approach produced one whole brain map of interest reflecting the degree to which a participant's brain activity increased as the value assigned to a social other's welfare increased. We then submitted these to a group-level analysis to assess whether this amplitude-modulated signal varied across groups. Whole brain group-level maps were thresholded using an FDR-corrected $q < 0.05$ and cluster size of 10 voxels and an

FWER-corrected $P < 0.001$. We also carried out this analysis using small-volume FDR-correction $q < 0.05$ in a pre-registered a priori region-of-interest (BA from the Harvard–Oxford atlas (90–93)). To visualize how the signal varied as a function of the subjective value assigned to social others and as a function of social distance (Supplementary Fig. S6), we estimated another GLM that included seven block regressors for each social other ($N = 1$, 2, 5, 10, 20, 50, and 100), six motion parameters, white matter signal, and CSF signal (and a regressor for missed choice among 13 participants). This produced seven whole brain maps of interest reflecting participants' percent signal change for each of the seven social others included in the task. Supplementary Fig. S6 depicts the mean percent signal change (y-axes) as a function of social value (x-axes), group (blue = altruists, red = controls), and social distance (columns) in the each of the clusters (rows) revealed from the amplitude modulation results from Figs 5 and 6.

Bootstrap resampling

To test in which regions neural activation encoded subjective valuation of others' welfare differently in altruists versus controls, we also ran a bootstrap resampling test. To do this, we randomly generated new samples of controls and altruists (each resample contained data from $N = 26$ participants) using a bootstrapping with replacement procedure and 5,000 permutations. On each permutation, we calculated the mean difference between the altruists and controls for all voxels within a gray matter mask. Thus, at each voxel, we generated a distribution of 5,000 mean difference values. We then applied a voxel height threshold determined using the 99% bootstrap CIs (i.e. thresholded voxels in which the bootstrap distribution overlapped with zero) and FWER cluster correction.

Acknowledgments

We would like to thank Kinney Van Hecke and the Center for Functional and Molecular Imaging staff for help with data collection. We also thank Lin Gan for helping organize neuroimaging coordinate tables, and Thalia P. Wheatley, Adam E. Green, and Kostadin Kushlev for helpful discussions and feedback related to this study. We also wish to express our gratitude to the participants who contributed their time and energy to this work. This manuscript was posted as a preprint: https://doi.org/10.31234/osf.io/8erfa.

Supplementary material

Supplementary material is available at *PNAS Nexus* online.

Funding

This project was supported by National Science Foundation (NSF) grant (#1729406) to A.A.M. and M.A.D. and NSF Graduate Research Fellowship to S.A.R.

Author contributions

A.A.M. and M.A.D. conceptualized the study; S.A.R. and K.O. designed and pre-registered the study procedure; A.S.V. assisted with designing the MRI acquisition protocol; M.L.P. and H.S.E. coordinated participant recruitment and data collection; S.A.R., K.O., K.B., M.L.P. and H.S.E. collected data; P.A. and J.L.L. aggregated and analyzed data on social relationships; S.A.R. analyzed all other data and wrote the manuscript with A.A.M; and all co-authors revised and approved the manuscript.

Data availability

Unthresholded univariate brain maps are available on the Open Science Framework (https://osf.io/mk9cu).

Code availability

All code for preprocessing and analysis are available on the Open Science Framework (https://osf.io/mk9cu).

References

1 de Waal FBM. 2008. Putting the altruism back into altruism: the evolution of empathy. *Annu Rev Psychol*. 59:279–300.
2 Batson CD. 2011. *Altruism in humans*. Oxford (UK): Oxford University Press.
3 Marsh AA. 2019. The caring continuum: evolved hormonal and proximal mechanisms explain prosocial and antisocial extremes. *Annu Rev Psychol*. 70:20–25.
4 Marsh AA. 2016. Neural, cognitive, and evolutionary foundations of human altruism. *Wiley Interdiscip Rev Cogn Sci*. 7:59–71.
5 Miller DT. 1999. The norm of self-interest. *Am Psychol*. 81:5–16.
6 Pew Research Center. 2019. Trust and distrust in America. Retrieved from https://www.pewresearch.org/politics/2019/07/22/trust-and-distrust-in-america/.
7 Strombach T, *et al*. 2015. Social discounting involves modulation of neural value signals by temporoparietal junction. *Proc Natl Acad Sci*. 112:1619–1624.
8 Soutschek A, Ruff CC, Strombach T, Kalenscher T, Tobler PN. 2016. Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Sci Adv*. 2:e1600992.
9 Stevens JR, Hauser MD. 2004. Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn Sci*. 8:60–65.
10 Kocher MG, Martinsson P, Myrseth KOR, Wollbrant CE. 2017. Strong, bold, and kind: self-control and cooperation in social dilemmas. *Exp Econ*. 20:44–69.
11 Rhoads SA, *et al*. 2023. Unselfish traits and social decision-making patterns characterize six populations of real-world extraordinary altruists. *Nat Commun*. 14:1807.
12 Vekaria KM, Brethel-Haurwitz KM, Cardinale EM, Stoycos SA, Marsh AA. 2017. Social discounting and distance perceptions in costly altruism. *Nat Hum Behav*. 1:1–7.
13 Galante J, Bekkers MJ, Mitchell C, Gallacher J. 2016. Loving-kindness meditation effects on well-being and altruism: a mixed-methods online RCT. *Appl Psychol Heal Well Being*. 8:322–350.
14 Condon P, Desbordes G, Miller WB, DeSteno D. 2013. Meditation increases compassionate responses to suffering. *Psychol Sci*. 24:2125–2127.
15 Rand DG, Epstein ZG. 2014. Risking your life without a second thought: intuitive decision-making and extreme altruism. *PLoS One*. 9:e109687.
16 Rand DG. 2016. Cooperation, fast and slow: meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychol Sci*. 27:1192–1206.
17 O'Connell K, *et al*. 2019. Increased similarity of neural responses to experienced and empathic distress in costly altruism. *Sci Rep*. 9:1–11.

18 Brethel-Haurwitz KM, *et al.* 2018. Extraordinary altruists exhibit enhanced self–other overlap in neural responses to distress. *Psychol Sci*. 29:1–11.
19 Marsh AA, *et al.* 2014. Neural and cognitive characteristics of extraordinary altruists. *Proc Natl Acad Sci*. 111:15036–15041.
20 Eisenberg N, Fabes RA. 1990. Empathy: conceptualization, measurement, and relation to prosocial behavior. *Motiv Emot*. 14: 131–149.
21 Jones BA, Rachlin H. 2006. Social discounting. *Psychol Sci*. 17: 283–286.
22 Jones BA, Rachlin H. 2009. Delay, probability, and social discounting in a public goods game. *J Exp Anal Behav*. 91:61–73.
23 Böckler A, Tusche A, Singer T. 2016. The structure of human prosociality: differentiating altruistically motivated, norm motivated, strategically motivated, and self-reported prosocial behavior. *Soc Psychol Personal Sci*. 7:530–541.
24 Ma Q, Pei G, Jin J. 2015. What makes you generous? The influence of rural and urban rearing on social discounting in China. *PLoS One*. 10:e0133078.
25 Strombach T, *et al.* 2014. Charity begins at home: cultural differences in social discounting and generosity. *J Behav Decis Mak*. 27: 235–245.
26 Locey ML, Jones BA, Rachlin H. 2011. Real and hypothetical rewards in self-control and social discounting. *Judgm Decis Mak*. 6: 552–564.
27 Hill PF, Yi R, Spreng RN, Diana RA. 2017. Neural congruence between intertemporal and interpersonal self-control: evidence from delay and social discounting. *Neuroimage*. 162:186–198.
28 Apps MAJ, Rushworth MFS, Chang SWC. 2016. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron*. 90:692–707.
29 Basile BM, Schafroth JL, Karaskiewicz CL, Chang SWC, Murray EA. 2020. The anterior cingulate cortex is necessary for forming prosocial preferences from vicarious reinforcement in monkeys. *PLoS Biol*. 18:e3000677.
30 Chang SWC, *et al.* 2015. Neural mechanisms of social decision-making in the primate amygdala. *Proc Natl Acad Sci*. 112: 16012–16017.
31 Dal Monte O, Chu CCJ, Fagan NA, Chang SWC. 2020. Specialized medial prefrontal–amygdala coordination in other-regarding decision preference. *Nat Neurosci*. 23:565–574.
32 Hackel LM, Wills JA, Van Bavel JJ. 2020. Shifting prosocial intuitions: neurocognitive evidence for a value-based account of group-based cooperation. *Soc Cogn Affect Neurosci*. 15:371–381.
33 Hackel LM, Zaki J, Van Bavel JJ. 2017. Social identity shapes social valuation: evidence from prosocial behavior and vicarious reward. *Soc Cogn Affect Neurosci*. 12:1219–1228.
34 Zaki J, Mitchell JP. 2011. Equitable decision making is associated with neural markers of intrinsic value. *Proc Natl Acad Sci*. 108: 19761–19766.
35 Zoh Y, Chang SWC, Crockett MJ. 2021. The prefrontal cortex and (uniquely) human cooperation: a comparative perspective. *Neuropsychopharmacology*. 47:119–133.
36 Rhoads SA, Cutler J, Marsh AA. 2021. A feature-based network analysis and fMRI meta-analysis reveal three distinct types of prosocial decisions. *Soc Cogn Affect Neurosci*. 16:1214–1233.
37 Scheggia D, *et al.* 2022. Reciprocal cortico-amygdala connections regulate prosocial and selfish choices in mice. *Nat Neurosci*. 25: 1505–1518.
38 Parnamets P, Shuster A, Reinero DA, Van Bavel JJ. 2020. A value-based framework for understanding cooperation. *Psychol Sci*. 29: 227–234.
39 Salzberg S. 1996. *Lovingkindness meditation: learning to love through insight meditation*. Louisville (CO): Sounds True.
40 Hutcherson CA, Seppala EM, Gross JJ. 2008. Loving-kindness meditation increases social connectedness. *Emotion*. 8:720–724.
41 Fredrickson BL, Cohn MA, Coffey KA, Pek J, Finkel SM. 2008. Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources. *J Pers Soc Psychol*. 95:1045–1062.
42 Kang Y, Gray JR, Dovidio JF. 2014. The nondiscriminating heart: lovingkindness meditation training decreases implicit intergroup bias. *J Exp Psychol Gen*. 143:1306–1313.
43 Reb J, Junjie S, Narayanan J. 2010. Compassionate dictators? The effects of loving-kindness meditation on offers in a dictator game. IACM 23rd Annual Conference Paper.
44 Kreplin U, Farias M, Brazil IA. 2018. The limited prosocial effects of meditation: a systematic review and meta-analysis. *Sci Rep*. 8: 1-10.
45 Seaman KL, *et al.* 2018. Subjective value representations during effort, probability and time discounting across adulthood. *Soc Cogn Affect Neurosci*. 13:449–459.
46 Chang SWC, Gariépy JF, Platt ML. 2013. Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci*. 16:243–250.
47 Rudebeck PH, Buckley MJ, Walton ME, Rushworth MFS. 2006. A role for the macaque anterior cingulate gyrus in social valuation. *Science*. 313:1310–1312.
48 Hill MR, Boorman ED, Fried I. 2016. Observational learning computations in neurons of the human anterior cingulate cortex. *Nat Commun*. 7:12722.
49 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*. 35:13720–13727.
50 Apps MAJ, Ramnani N. 2014. The anterior cingulate gyrus signals the net value of others' rewards. *J Neurosci*. 34:6190–6200.
51 Rhoads SA, *et al.* 2022. Neurocomputational basis of learning when choices simultaneously affect both oneself and others. *PsyArXiv* rf4x9. https://doi.org/10.31234/osf.io/rf4x9, preprint: not peer reviewed.
52 Brethel-Haurwitz KM, *et al.* 2017. Amygdala—midbrain connectivity indicates a role for the mammalian parental care system in human altruism. *Proc R Soc B Biol Sci*. 284:20171731.
53 Koenigs M. 2012. The role of prefrontal cortex in psychopathy. *Rev Neurosci*. 23:253–262.
54 Abe N, Greene JD, Kiehl KA. 2018. Reduced engagement of the anterior cingulate cortex in the dishonest decision-making of incarcerated psychopaths. *Soc Cogn Affect Neurosci*. 13:797–807.
55 Marsh AA, *et al.* 2013. Empathic responsiveness in amygdala and anterior cingulate cortex in youths with psychopathic traits. *J Child Psychol Psychiatry Allied Discip*. 54:900–910.
56 Vieira JB, *et al.* 2014. Psychopathic traits are associated with cortical and subcortical volume alterations in healthy individuals. *Soc Cogn Affect Neurosci*. 10:1693–1704.
57 Parkinson C, Liu S, Wheatley T. 2014. A common cortical metric for spatial, temporal, and social distance. *J Neurosci*. 34: 1979–1987.
58 Brethel-Haurwitz KM, Oathes DJ, Kable JW. 2021. Causal role of the right temporoparietal junction in selfishness depends on the social partner. *Soc Cogn Affect Neurosci*. 17:541–548.
59 Morishima Y, Schunk D, Bruhin A, Ruff CC, Fehr E. 2012. Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*. 75:73–79.

60 Hutcherson CA, Bushong B, Rangel A. 2015. A neurocomputational model of altruistic choice and its implications. *Neuron*. 87:451–463.
61 Sellitto M, Neufang S, Schweda A, Weber B, Kalenscher T. 2021. Arbitration between insula and temporoparietal junction subserves framing-induced boosts in generosity during social discounting. *Neuroimage*. 238:118211.
62 Rhoads SA, Marsh AA, *et al*. 2023. Doing good and feeling good: relationships between altruism and well-being for altruists, beneficiaries, and observers. In: Helliwell JF, editor. *World happiness report*. 11th ed. New York (NY): United Nations Sustainable Development Solutions Network. p. 103–130.
63 Helliwell JF, Huang H, Wang S, Norton M. 2022. Happiness, benevolence, and trust during COVID-19 and beyond. In: Helliwell JF, Layard R, Sachs JD, editors. *World happiness report 2022*. New York (NY): Sustainable Development Solutions Network. p. 14–52.
64 Waytz A, Zaki J, Mitchell JP. 2012. Response of dorsomedial prefrontal cortex predicts altruistic behavior. *J Neurosci*. 32: 7646–7650.
65 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.
66 FitzGerald THB, Friston KJ, Dolan RJ. 2012. Action-specific value signals in reward-related regions of the human brain. *J Neurosci*. 32:16417–16423.
67 Story GW, *et al*. 2020. Social discounting of pain. *J Exp Anal Behav*. 114:308–325.
68 O'Connell K. 2021. *From perception to social connection: assessing deficits and modulation of empathic processes to study human social behavior*. Washington (DC): Georgetown University.
69 Kraft-Todd GT, Rand DG. 2019. Rare and costly prosocial behaviors are perceived as heroic. *Front Psychol*. 10:1–7.
70 Vekaria KM, *et al*. 2019. The role of prospection in altruistic bone marrow donation decisions. *Heal Psychol*. 39:316–324.
71 Brethel-Haurwitz KM, Stoycos SA, Cardinale EM, Huebner B, Marsh AA. 2016. Is costly punishment altruistic? Exploring rejection of unfair offers in the ultimatum game in real-world altruists. *Sci Rep*. 6:18974.
72 Hartmann H, Forbes PAG, Rütgen M, Lamm C. 2022. Placebo analgesia reduces costly prosocial helping to lower another person's pain. *Psychol Sci*. 33:1867–1881.
73 Lockwood PL, *et al*. 2017. Prosocial apathy for helping others when effort is required. *Nat Hum Behav*. 1:1–10.
74 Amormino P, *et al*. 2022. Beliefs about humanity, not higher power, predict extraordinary altruism. *J Res Pers*. 101:104313.
75 Lim D, DeSteno D. 2016. Suffering and compassion: the links among adverse life experiences, empathy, compassion, and prosocial behavior. *Emotion*. 16:175–182.
76 Erdfelder E, Faul F, Buchner A, Lang AG. 2009. Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 41:1149–1160.
77 Durnez J, *et al*. 2016. Power and sample size calculations for fMRI studies based on the prevalence of active peaks. *bioRxiv* 049429. https://doi.org/10.1101/049429, preprint: not peer reviewed.
78 Derogatis LR, Unger R. 2010 Symptom checklist-90-revised. In: Weiner IB, Craighead WE, editors. *The corsini encyclopedia of psychology*. Hoboken (NJ): Wiley. https://doi.org/10.1002/9780470479216.corpsy0970.
79 Kaufman AS, Kaufman NL. 2004. *Kaufman brief intelligence test*. 2nd ed. Bloomington (MN): Pearson, Inc.
80 Taylor MJ, *et al*. 2009. Neural correlates of personally familiar faces: parents, partner and own faces. *Hum Brain Mapp*. 30: 2008–2020.
81 Braun V, Clarke V. 2012. Thematic analysis. In: *APA handbook of research methods in psychology, vol 2: research designs: quantitative, qualitative, neuropsychological, and biological*. Washington (DC): American Psychological Association. p. 57–71.
82 Soutschek A, *et al*. 2017. The dopaminergic reward system underpins gender differences in social preferences. *Nat Hum Behav*. 1: 819–827.
83 Odum AL. 2012. Delay discounting: I'm a k, you're a k. *J Exp Anal Behav*. 96:427–439.
84 Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2020. nlme: linear and nonlinear mixed effects models.
85 Young ME. 2017. Discounting: a practical guide to multilevel analysis of indifference data. *J Exp Anal Behav*. 108:97–112.
86 Myerson J, Green L, Warusawitharana M. 2001. Area under the curve as a measure of discounting. *J Exp Anal Behav*. 76:235–243.
87 Esteban O, *et al*. 2018. FMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods*. 16:111-116.
88 Cox RW. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 29:162–173.
89 Chen G, Adleman NE, Saad ZS, Leibenluft E, Cox RW. 2014. Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. *Neuroimage*. 99:571–588.
90 Makris N, *et al*. 2006. Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophr Res*. 83:155–171.
91 Frazier JA, *et al*. 2005. Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. *Am J Psychiatry*. 162:1256–1265.
92 Desikan RS, *et al*. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 31:968–980.
93 Goldstein JM, *et al*. 2007. Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability. *Biol Psychiatry*. 61: 935–945.