A feature-based network analysis and fMRI meta-analysis of the task structure underlying prosocial decision-making reveal three distinct clusters: cooperation, equity, and altruism

Shawn A. Rhoads[1], Jo Cutler[2], Abigail A. Marsh[1]

[1]Department of Psychology, Georgetown University, Washington DC, USA

[2]Centre for Human Brain Health, University of Birmingham, Birmingham, UK

Keywords: prosocial, decision-making, meta-analysis, cooperation, equity, altruism

Correspondence to:

Shawn A. Rhoads (sr1209@georgetown.edu)

37th and O Streets NW White Gravenor Hall 304, Washington DC, 200057

Acknowledgements: We would like to thank Peter E. Turkeltaub, Katherine O'Connell, Kathryn Berluti, Adam E. Green, and Joscelin Rocha-Hidalgo for helpful suggestions, discussions, and feedback regarding this project. We are also grateful to Daniel Campbell-Meiklejohn for support in collecting data, Lin Gan, Alexa King, and Kathleen Neill for their help identifying the task features, and all of the researchers who contributed their time and data for our meta-analysis.

Funding: This work was supported by the National Science Foundation Graduate Research Fellowship Program Award to S.A.R.

Conflicts of interest: Authors have nothing to disclose.

Abstract

Tasks that measure correlates of prosocial decision-making share one common feature: agents can make choices that increase the welfare of a beneficiary. However, prosocial decisions vary widely as a function of other task features. The diverse ways that prosociality is defined and the heterogeneity of prosocial decisions have created challenges for interpreting findings across studies and identifying their neural correlates. To overcome these challenges, the present study aimed to organize the prosocial decision-making task-space of neuroimaging studies. We conducted a systematic search for studies in which participants made decisions to increase the welfare of others during fMRI. We identified shared and distinct features of these tasks and employed an unsupervised graph-based approach to assess how various forms of prosocial decision-making are related in terms of their low-level components (e.g., task features like potential cost to the agent or potential for reciprocity). We uncovered three clusters of prosocial decisions: cooperation, equity, and altruism. This feature-based representation of the task structure was supported by results of a neuroimaging meta-analysis that each category of prosocial decisions recruited diverging neural systems. Results clarify some of the existing heterogeneity in how prosociality is conceptualized and generate insight for future research and task paradigm development.

Introduction

Prosocial decisions—choices that increase the welfare of others—are universal across cultures (Henrich et al., 2005) and are integral for supporting interpersonal relationships at multiple scales, including between dyads (Declerck et al., 2013; Rusbult & Van Lange, 2003), among groups and social networks (de Waal, 2008; Fehr et al., 2002; Fehr & Camerer, 2007; Fehr & Fischbacher, 2003; FeldmanHall, 2017; Fowler & Christakis, 2010), and within societies (Nowak, 2006). And, while largely conserved across species (Burkart et al., 2014; de Waal, 2008; Hare, 2017), the prevalence and variety of prosociality exhibited by humans is unique (Fehr & Schurtenberger, 2018; Zaki & Mitchell, 2013). Although cognitive and neural processes underlying various forms of prosociality have been studied extensively across disciplines spanning psychology, neuroscience, economics, and biology, the heterogeneity of prosocial decisions has led to inconsistencies in how they are operationalized and categorized (Batson & Powell, 2003; de Waal, 2008; Declerck et al., 2013; Fehr et al., 2002; Fehr & Schmidt, 1999; Marsh, 2016; Parnamets et al., 2020; Rand & Nowak, 2013; Rilling et al., 2002; Ruff & Fehr, 2014; Tricomi & Sullivan-Toole, 2015). This can create challenges when interpreting findings across neuroimaging studies or when attempting to understand how different types of prosocial decisions vary in terms of their underlying processes.

Derived from the Latin stem *pro* and root *socius*, signifying "for a companion", prosocial decision-making refers to "decisions made for the benefit of another" (*The American Heritage Dictionary of the English Language*, 2000). Laboratory tasks that measure correlates of prosocial decision-making share one common feature: allowing deciding participants (or agents) to make choices that increase the welfare of a beneficiary. However, prosocial decisions vary widely as a function of other task features. For example, although choosing to forgo resources (usually money) to alleviate the suffering of a stranger (in a charitable donation task) versus choosing to contribute money to maximize equity among known members in a group (in a Public Goods Game) both share a common prosocial core of increasing the welfare of others, these decisions diverge along multiple other characteristics. In the first example, a prosocial agent sacrifices resources in response to another person's distress with the understanding that they will not receive anything in return—suggesting a likely role for empathic concern and planning for prosocial action without any anticipation of reward. In the second example, all prosocial agents are relying on the decisions of others and are hoping to increase the total pool of resources for everyone involved. This suggests a role for monitoring the expected actions of others' decisions and includes the anticipation of self-rewarding outcomes.

In addition to charitable donation tasks and Public Goods Games, other common prosocial paradigms include Dictator Games, Prisoner’s Dilemmas, and Trust Games. Such tasks can be implemented with multiple variations, and the vast number of combinations of task features is a major source of heterogeneity. This heterogeneity raises the question of whether common mechanisms underlie all prosocial choices. One possibility is that prosocial decisions in each distinct task are supported by distinct mechanisms. But it is more likely that taxonomic clusters exist within the task-space of prosocial decision-making that reflect common underlying neural processes (Cutler & Campbell-Meiklejohn, 2019). One way to identify such clusters would be via a bottom-up approach aimed at characterizing the task structure of prosocial decision-making by analyzing the way specific tasks cluster according to their low-level features. In other words, developing one level of a formal representation (or ontology) of cognitive tasks and their inter-relationships (Poldrack & Yarkoni, 2016; Turner & Laird, 2012). The first goal of this paper was to clarify how different prosocial tasks are inter-related and how their low-level features give rise to broad categories of prosocial decisions. Then, using this information across various studies, we employed an unsupervised graph-based approach to generate a preliminary characterization of the neuroimaging task-space comprised of the distinct and shared task features of prosocial decision-making paradigms. Finally, we conducted an fMRI meta-analysis to identify patterns of distinct and overlapping neural activation that correspond to the identified clusters of prosocial processes.

Breaking prosocial decision processes down into their relevant task features may allow a better understanding of how prosocial decisions are inter-related, and how they diverge. In general, the features that distinguish these tasks involve those related to the *beneficiary* (Is the beneficiary a real or imaginary person (or persons) or an organization like a charity? Is their identity apparent to the agent? Is their need or distress known to the agent?), to the *interaction* (Does the beneficiary also make decisions that will affect ultimate outcome? Will the agent and beneficiary interact only once or more than once?), and to the *outcomes* of the agent's decision (What is the magnitude of the benefit to the beneficiary? Will the decision result in rewarding outcomes for the agent? Will it be costly? Will the decision conform to social norms, such as equity? How certain is the outcome?). Multiple combinations of these features likely shape the context, motivations, and outcomes of prosocial decisions, and thus should recruit diverging neural systems.

Features related to the beneficiary. Various features related to the beneficiary of a prosocial decision are known to influence such decisions. Beneficiaries can include specific people, such as close or familiar others (Fareri et al., 2015; Hill et al., 2017; Schreuders et al., 2018; Sharp et al., 2011; Telzer et al., 2011) or in-group members (Balliet et al., 2014; Hackel et al., 2017; Telzer et al., 2015; Wills, Hackel, & Van Bavel, 2018), or can be hypothetical or even non-human (e.g., computers) (Delgado et al., 2005; Fareri et al., 2012). Across contexts, agents are typically more willing to help people than computers (Fareri et al., 2015), and are more willing to help people close to them than strangers (Jones & Rachlin, 2006, 2009; Safin et al., 2013; Strombach et al., 2015). Neural activation during decisions that affect real versus imaginary beneficiaries (e.g., computer) is increased in regions important for theory of mind or inferring the mental states of others such as the temporoparietal junction (TPJ) (FeldmanHall et al., 2012).

In tasks that include real beneficiaries who are previously unknown to the agent, the beneficiary may be another participant in study (Weiland et al., 2012), or an anonymous stranger (Bault et al., 2014; Hutcherson et al., 2015; Strombach et al., 2015) who the agent may have briefly met before the task (Abe et al., 2019; Shaw et al., 2018) or seen in a photograph (Genevsky et al., 2013; Park et al., 2017). Receiving any identifying information about a beneficiary generally increases prosociality, in line with the identifiable victim effect (Jenni & Loewenstein, 1997; Kogut & Ritov, 2005; Lee & Feeley, 2016). This effect also results in greater prosociality toward single individuals versus collectives (Kogut & Ritov, 2005; Lee & Feeley, 2016), including charitable organizations, whether predetermined (Greening et al., 2014), of the agent's choosing (Kuss et al., 2013), or from a list of charities (Hare et al., 2010; Izuma et al., 2010; Tusche et al., 2016). Increases in prosocial decision-making are particularly robust when the need or distress of the beneficiary is salient (FeldmanHall et al., 2015; Genevsky et al., 2013; Kuss et al., 2015; Tusche et al., 2016). Cues that signal need or distress typically elicit empathic concern, which motivates the desire to alleviate it (Batson, 2011; de Waal, 2008; Marsh, 2016; Preston & de Waal, 2002). This form of empathy is supported by activity in neural regions including the anterior insula, ACC, and pre-supplementary motor area (pre-SMA) (Lamm et al.,

2011) and empathic neural responding predicts prosocial decision-making both in and out of the laboratory (Tusche et al., 2016; Vekaria et al., 2020).

Features related to the interaction. Aspects of the interaction between agents and beneficiaries (or other agents) in prosocial tasks also influence agents' decisions, particularly when agents can learn about those with whom they are interacting. In some interactions, only one agent can influence the outcome. For example, in Dictator Games, agents unilaterally allocate resources between themselves and a beneficiary (Engel, 2011). In others, multiple agents can shape the outcome. For example, in social dilemmas or Trust Games, agents can choose to cooperate with others in order to increase the total pool of available resources for everyone involved, or can defect to obtain better outcomes for self (Balliet et al., 2011). Alternatively, in ultimatum games, agents receive feedback about their decisions from beneficiaries, who can accept or reject the offer (Güth et al., 1982).

Some prosocial decisions involve repeated interactions, which, unlike one-shot interactions, provide opportunities to reciprocate or respond to feedback about prior choices (Thielmann et al., 2020). When repeated interactions are expected, it typically motivates cooperation, with agents motivated to pay short-term cooperation "costs" to increase future reciprocity from a partner (Milinski et al., 2001; Rand & Nowak, 2013) and more willing to cooperate with partners who have cooperated previously (Fehr & Schurtenberger, 2018). This may be related to the ability to update expectations of others' likely behavior, a type of social learning is supported by the subgenual anterior cingulate cortex (ACC) (Christopoulos & King-Casas, 2014).

Features related to outcomes. Across interaction types, prosocial decisions are also shaped by their anticipated outcomes. In some cases, prosocial decisions may benefit the agent directly. In Prisoner's Dilemmas or Public Goods Games, for example, decisions to cooperate increase the probability of future reciprocity. In such tasks, agents also take on the role of beneficiaries (Chaudhuri, 2011; Rand & Nowak, 2013), and thus must arbitrate between their own and others' rewards. These tasks often recruit neural systems that support subjective valuation and reward expectancy, such as the ventromedial PFC and ventral striatum (Parnamets et al., 2020; Wills, Hackel, & Van Bavel, 2018; Wills, Hackel, FeldmanHall, et al., 2018). These tasks also carry an element of uncertainty (Bellucci et al., 2017), with the agent's outcome often dependent on a beneficiary or trustee's choices (Mayer et al., 1995). Uncertainty during these decisions may be reflected through activation in the dorsal ACC (Aimone et al., 2014).

Prosocial choices may also yield more abstract rewards, such as conformity to desirable social norms like maximizing equity among multiple parties (via, for example, a 50-50 split of resources) (Fehr & Schurtenberger, 2018; Krupka & Weber, 2013; López-Pérez, 2008). In some cases, agents may choose to act prosocially and forgo resources to avoid deviating from desirable norm, which is known as disadvantageous inequity aversion (Tricomi & Sullivan-Toole, 2015). Equitable interpersonal decisions are thought to engage neural structures involved in computing subjective value such as the medial PFC and ventral striatum, and thus may be motivated through increased intrinsic value placed on the decision (Zaki & Mitchell, 2011), perhaps via their goal of producing increased subjective happiness for agents and beneficiaries (Tabibnia et al., 2007; Tabibnia & Lieberman, 2007). In tasks with repeated interactions, these decisions may also reflect the maintenance of abstract, norm-based rules regarding fairness or reciprocity in dorsolateral PFC (Guroglu et al., 2014; van den Bos et al., 2009).

In other prosocial decision-making tasks (such as Dictator Games or charitable giving tasks) agents can forgo resources (including money, time, effort, or safety) solely to benefit others. In this case, prosocial choices are made despite certain concrete costs to the agent, often to alleviate the beneficiary's distress or need. As described above, such decisions are thought to be driven by activation in regions like anterior insula, which represent negative affective states (e.g., pain or distress) of the beneficiary (FeldmanHall et al., 2015; Tusche et al., 2016). Such choices also may yield indirect gains, including increases in mood or well-being (Aknin et al., 2012; Curry et al., 2018; Dunn et al., 2008), possibly related to the vicarious reward of improving the beneficiary's welfare (Mobbs et al., 2009); such vicarious reward may be supported by activity in ventral striatum and ventromedial PFC.

Given the diversity of extant prosocial decision tasks, two recent meta-analytic studies have been very valuable in describing the neural correlates of prosocial behaviors aggregated across tasks that reflect divergent constellations of the above variables. Belluci and colleagues (2020) aggregated across a wide range of tasks in which participants made decisions about others, rated others' traits, or judged others' behaviors in an effort to find neural activation overlap among prosociality, empathy, and mentalizing. They found four regions to be preferentially engaged across the tasks they incorporated: dorsolateral PFC, ventromedial PFC, dorsal posterior cingulate cortex (PCC), and middle cingulate cortex (MCC). Of these regions, they found a conjunction in dorsal PCC activation during tasks involving prosocial behavior and tasks involving mentalizing (understanding another person's needs and inferring goals across contexts); they also found a conjunction in MCC activation across tasks involving prosocial behavior and tasks involving empathy (resonating with another's needs) (Bellucci et al., 2020). Activation during prosocial behavior in the dorsolateral PFC and ventromedial PFC did not overlap with activation during mentalizing or empathy tasks. This work identified common neural patterns underlying a range of behaviors related to prosociality, but by not considering key differences among types of prosocial decisions, it was not able to identify whether they are supported by distinct processes. Cutler and Campbell-Meiklejohn (2019) provided preliminary evidence that distinct neural regions do indeed support different forms of prosocial decision-making, finding diverging patterns of activation for prosocial behaviors that do not provide an opportunity to gain extrinsic rewards (and thus likely are intrinsically motivated) versus those with the probability of gaining an extrinsic reward. For example, extrinsically motivated decisions recruited greater activity in striatal regions relative to intrinsically motivated decisions. In contrast, intrinsically motivated decisions recruited increased activation in ventromedial PFC relative to extrinsically motivated decisions. Activation in ventromedial PFC also differentiated these types along a posterior (intrinsic) to anterior (extrinsic) axis.

However, the distinction between extrinsic and intrinsic motivation was determined in advance, rather than being driven by objective features of the data. This is also only one of many possible distinctions among forms of prosocial behavior. An alternative means of investigating neural substrates of various prosocial decision tasks could instead take a more bottom-up approach that identifies distinct clusters of tasks that emerge from statistical variation in their objective features or outcomes. For example, a recent behavioral study analyzed the behavioral outcomes of different economic prosocial tasks (such as the percentage of prosocial decisions during each task, the ratio of other-regarding to self-regarding decisions in each, average monetary donations, or summary scores of self-reported measures). Using factor analysis, they determined that the prosocial tasks clustered into four factors that the authors termed: altruistically motivated prosocial behavior, norm-motivated prosocial behavior, strategically motivated prosocial behavior, and self-reported prosocial behavior (Böckler et al., 2016).

We sought to use a similar bottom-up approach to meta-analytically investigate the neural correlates of prosocial decision-making during fMRI. We focused on objective features that distinguish the tasks themselves, which included features related to outcomes of decisions, to the beneficiaries of the decision, and to the interaction between agents and beneficiaries. We first compiled data from 43 unique fMRI studies of prosocial decision-making (including 25 maps and 18 coordinate tables across 1,423 participants). We then dummy-coded task features related to the beneficiary, interaction, and outcome of each decision and employed a data-driven, graph-based approach to identify clusters of studies based on their overlapping versus distinct task features. (As described in detail below, this approached indicated that the prosocial decision-making task-space comprises three clusters: cooperation, equity, and altruism). We next used a meta-analytic approach that combined group-level statistical parametric images with reported peak-coordinates to identify divergent neural activation patterns across these clusters of studies. In so doing, the present study resolves some discrepancies in how prosocial decisions are conceptualized, expands understanding regarding how prosocial decisions are related and distinct, and generates insight for future research.

Method

Literature search and study selection

A literature search using PubMed identified research published prior to June 2019 using keywords either ("fMRI" or "neur*") and one of the following: ("prosocial", "trust game", "fairness", "reciproc*", "cooperat*", "charitable", "public goods", "dictator", "ultimatum", "prisoner*"). The search returned 201 articles. We removed 124 articles that did not meet key criteria, such as non-neuroimaging studies, neuroimaging studies that did not use functional magnetic resonance imaging (fMRI), and literature reviews or meta-analyses. Independently, we identified 123 potential articles from the reference lists of the remaining 77 articles. After removing all duplicate titles from the combined lists, 146 articles remained. We then selected only articles that reported novel whole-brain fMRI data (i.e., data only published once) that were collected while participants made decisions that benefitted another individual (prosocial decisions). We also limited our search to include only data from studies that were able to examine differences in activation during prosocial decisions relative to decisions that benefited the agent alone (selfish decisions). In some cases, this was a contrast between prosocial choices and selfish choices within a task condition or parametric modulation of the amount given. For other studies, the contrast was between decisions during a prosocial condition and a self-only condition. We did not include contrasts involving alternate control conditions (e.g., rest, visuomotor controls), even when these were available, due to significant variation in brain activation (Cutler & Campbell-Meiklejohn, 2019). Upon review of the remaining 146 articles, 69 were identified that met our inclusion criteria.

We sent emails to the corresponding authors of all included studies to request unthresholded, group-level t-statistic map(s) from the study that best fit our criteria. For studies that included pharmacological manipulations or clinical populations, we requested data from only the control group. If maps were not available, we requested coordinates for contrasts of interests or extracted them from manuscripts. If a coordinate table reported Z-scores or Talaraich coordinates, peak values were transformed to *t*-statistics and MNI coordinates, respectively. If the contrast of interest was reported in both directions (e.g. cooperate > defect and defect > cooperate), the selfish contrast peaks were assigned as negative t-values. Ultimately, we obtained the necessary data from 43 unique fMRI studies, including 25 maps and 18 coordinate tables that included data from 1,423 subjects (Table 1).

Table 1. Descriptions of studies included in meta-analysis

Study	N	Proportion female	Map or peak	Task	T	FWHM	Program	Sig	Contrast selected	Cluster Label
(Abe et al., 2019)	19	9 / 19	map	Joint force-production task	3T	8mm^3	SPM12	Pcorr < 0.001	Joint performance vs single performance	C
(Aimone et al., 2014)	28	15 / 30 total	peak	Trust Game, one-shot, binary (investor)	3T	8mm^3	SPM8	Puncorr <.0001	Share vs keep	C
(Bault et al., 2014)	25	12 / 29 total	map	Public Goods Game	3T	5mm^3	FSL	Pcorr < 0.05 (FWE)	PM monetary choice	C
(Chen et al., 2016)	93	50 / 104 total	map	Prisoner's Dilemma, iterative	3T	5mm^3	FSL	Pcorr <0.05 (FWE)	Cooperation vs defect (placebo only)	C
(Decety et al., 2004)	12	6 / 12	peak	Cooperative Pattern Game	3T	4.5 x 4.5 x 6 mm	SPM2	Pcorr < 0.05	Cooperation vs competition	C
(Delgado et al., 2005)	12	6 / 14 total	peak	Trust Game, binary, iterative (investor)	3T	4mm^3	Brain Voyager	Puncorr < 0.001	Share vs keep	C
(Fareri et al., 2015)	26	14 / 26	map	Trust Game, binary, iterative (investor)	3T	4mm^3	Brain Voyager	N/A	Share vs keep (across all partners)	C
(FeldmanHall et al., 2012)	14	8 / 14	peak	Your Pain, My Gain Task	3T	8mm^3	SPM	Pcorr < 0.05	PM monetary choices, no covariates	A
(FeldmanHall et al., 2015)	17	11 / 17	peak	Your Pain, My Gain Task	3T	8mm^3	SPM5	Pcorr < 0.05 (FWE)	PM monetary choice	A
(Fermin et al., 2016)	33	18 / 33	map	Prisoner's Dilemma, one-shot	3T	8mm^3	SPM8	Puncorr <.005	Cooperation vs defect	C
(Fouragnan et al., 2013)	18	0 / 18	peak	Trust Game (with priors manipulation)	4T	8mm^3	SPM8	Pcorr < 0.005	Share vs keep	C
(Garbarini et al., 2014)	16	8 / 16	map	Trust Game, one-shot (responder)	1.5T	4mm^3	Brain Voyager	Pcorr < 0.05	Reciprocate vs defect	E
(Genevsky et al., 2013)	11	6 / 11	peak	Charitable donation task with identifiable victim	3T	4mm^3	AFNI	Puncorr < 0.005	Donation vs no donation	A
(Greening et al., 2014)	18	9 / 18	peak	Charity task	3T	4mm^3	AFNI	Pcorr < 0.05	Charity vs self	A
(Guroglu et al., 2014)	22	17 / 28 total	peak	Modified Dictator Game (self-maximizing inequity (SMI) game)	3T	8mm^3	SPM8	Puncorr < 0.005	Equal split vs unequal	E
(Hare et al., 2010)	22	22 / 22	map	Charitable donation task	3T	8mm^3	SPM5	Puncorr < 0.0005	PM monetary choice	A

(Hutcherson et al., 2015)	51	0 /51	map	Modified Dictator Game, with probable outcomes	3T	8mm^3	SPM8	Puncorr < 0.05	Prosocial choice vs self	A
(Izuma et al., 2010)	23	12 / 23	map	Charitable donation task in presence or absence of observer	3T	6mm^3	SPM5	Puncorr < 0.001	Donation vs no donation (no observers)	A
(Koban et al., 2014)	17	10 / 22 total	map	Modified Dictator Game, Sharing/keeping during conflict	3T	8mm^3	SPM8	Puncorr < 0.001	Equal split vs self during human interpersonal conflict	E
(Kuss et al., 2013)	33	14 / 33	map	Charitable donation task, with probable outcomes	3T	6mm^3	SPM8	Puncorr < 0.001	Costly donation vs self	A
(Lee et al., 2018)	16	0 / 16	map	Tetris-like game, with self, helping, and harming conditions	3T	8mm^3	SPM5	Pcorr < 0.05	Help vs self	A
(Lelieveld et al., 2013)	26	17 / 26	map	Dictator Game with emotional manipulation	3T	6mm^3	SPM5	Puncorr < 0.001	Equal split vs unequal split	E
(Morishima et al., 2012)	27	17 / 30 total	map	Dictator Game	3T	8mm^3	SPM8	Pcorr < 0.05 (FWE)	Donation vs self	A
(Park et al., 2017)	159	100 / 166 total	map	Dictator Game with Faces	3T	4mm^3	AFNI	Puncorr < 0.005	Giving vs not giving (offer only)	A
(Ramsøy et al., 2015)	30	14 / 30	peak	Prisoner's Dilemma, with belief prompts about partner's decision	3T	8mm^3	SPM8	Puncorr < 0.001	Cooperation vs defect	C
(Schneider-Hassloff et al., 2015)	164	78 / 164	map	Prisoner's Dilemma with measures of attachment style	3T	8mm^3	SPM8	Puncorr < 0.001	Cooperation vs defect	C
(Schreuders et al., 2018)	22	12 / 27 total	map	Modified Dictator Game (self-maximizing inequity game)	3T	8mm^3	SPM8	Pcorr < 0.05	Equal split vs self (across interaction partners)	E
(Schreuders et al., 2019)	39	29 / 50 total	map	Modified Dictator Game (self-maximizing inequity game)	3T	8mm^3	SPM8	Pcorr < 0.001	Equal split vs self (across interaction partners)	E
(Sharp et al., 2011)	20	0 / 20	map	Trust Game (investor)	3T	4mm^3	AFNI	Puncorr < 0.001	Share vs keep (controls)	C
(Shaw et al., 2018)	38	0 / 38	peak	Ultimatum Game (proposer)	3T	5mm^3	FSL (preproc) SPM12 (GLM)	Pcorr < 0.001 (FWE)	PM monetary choice (proposers only)	C

(Smith-Collins et al., 2013)	24	24 / 24	peak	Trust Game, iterative (investor)	1.5T	8mm^3	SPM8	Puncorr < .001	Cooperation vs defect	C
(Stanley et al., 2012)	40	22 / 40	map	Trust Game, one-shot (investor)	3T	6mm^3	SPM8	Pcorr < .05 (FWE)	PM monetary choices (only human partners)	C
(Strombach et al., 2015)	27	13 / 27	map	Modified Dictator Game with social distance manipulation	3T	8mm^3	SPM8	Pcorr < 0.005	Equal split vs self	E
(Telzer et al., 2011)	25	13 / 25	map	Family Assistance Task	3T	8mm^3	SPM5	Pcorr < 0.05	Costly donation vs self	A
(Telzer et al., 2015)	29	13 / 29	map	Modified Dictator Game with group membership manipulation	3T	8mm^3	SPM8	Pcorr < 0.05	Donation vs self	A
(Tusche et al., 2016)	23	15 / 33 total	map	Charitable donation task	3T	8mm^3	SPM8	Pcorr < .05 (FWE)	High donation vs low donation	A
(van den Bos et al., 2009)	18	9 / 18	peak	Trust Game, one-shot (responder)	3T	6mm^3	SPM2	Puncorr < 0.001	Reciprocate vs defect	C
(Watanabe et al., 2014)	48	N/A	peak	Pay-it-forward + reputation-based indirect reciprocity game	3T	8mm^3	SPM8	Pcorr < 0.05 (FWE)	Cooperation vs defect (pay-it-forward & reputation-based conditions)	C
(Weiland et al., 2012)	14	8 / 14	peak	Dictator Game	3T	8mm^3	Brain Voyager	Puncorr < 0.005	Fair split (6:6, 7:5) vs unfair (8:4, 9:3, 10:2, 11:1)	E
(Will et al., 2016)	43	17 / 43	map	Dictator Game after Cyberball	3T	8mm^3	SPM8	Puncorr < .001	Equitable choice vs inequitable	E
(Wills, Hackel, & Van Bavel, 2018)	42	32 / 47	peak	Public Goods Game	3T	6mm^3	SPM12	Pcorr < .05	Give vs keep	C
(Wittmann et al., 2016)	24	9 / 24	peak	Cooperation Task	3T	5mm^3	FSL	Pcorr < 0.05 (FWE)	Cooperation versus competition	C
(Zaki & Mitchell, 2011)	15	6 / 15	peak	Modified Dictator Game	3T	6mm^3	SPM	Pcorr < 0.005	Equitable choice vs inequitable	E

Note. Some studies only reported gender as a proportion of the total sample, but not the final analytic sample. C=Cooperative; E=Equitable; A=Altruistic

Identifying task features across studies

We first reviewed the details of the methodologies of all available tasks and identified 13 distinct task features that varied among existing prosocial decision-making tasks across neuroimaging studies (Figure 1). These task features can be broken down into those vary as a function of the beneficiary, the interaction, and the outcome. (Although in theory features related to the agent can also vary, all participants included in the present study were healthy control adults, and we did not identify consistent features related to these participants—for example, consistent individual difference measures—in the available literature). Four independent raters dummy coded ("present" or "absent") the 13 features during the prosocial decision phase for the available contrast in each study with high initial agreement among coders (ICC=.82, $CI_{95\%}$=[.79, .84]). Discrepancies in the initial coding were then resolved through a consensus agreement across each of the four coders.

Identifying clusters within the task structure of prosocial decision-making

We next applied an unsupervised, graph-based approach to assess differential clusters of prosocial decisions based on their task features. We used the identified features to construct a bipartite graph. This graph contained two sets of nodes: nodes representing the 13 different task features and nodes representing the 43 different prosocial decision study contrasts. In this graph, an edge exists between a feature node and a study node if the study contrast contained the task feature (Figure 1, top). The bipartite graph was then projected onto a weighted network of studies, where edge weights between studies represented the Dice similarity coefficient (Dice, 1945) or the degree of overlapping task features relative to the total possible task features. We then ran the Louvain community detection algorithm (Blondel et al., 2008), which assigned study nodes to clusters in two steps. First, the algorithm finds small clusters of studies by optimizing local modularity. Second, it aggregates studies of the same cluster in a hierarchical fashion and builds a new network whose nodes are these clusters. These steps were repeated iteratively until the global modularity was maximized.

Neuroimaging preprocessing and meta-analyses

We next conducted meta-analyses combining reported peak information (coordinates and *t*-statistics) with original statistical parametric maps using the Anisotropic Effect Size Signed Differential Mapping software (AES-SDM, version 5.141; Radua et al., 2014). We selected this analytical technique rather than alternatives, such as coordinate-based activation likelihood (Eickhoff et al., 2009; Turkeltaub et al., 2002), because this approach enabled the utilization of precise, continuous estimates of effect sizes, assessment of between-study heterogeneity, and identification of potential publication bias (Radua et al., 2012). Using AES-SDM, within-study, voxel-level maps of effect sizes (*Hedge's g*) and their variances were re-created for each study. When only reported coordinates and statistics were available for a study, we calculated the effect size at each peak and estimated effect sizes in neighboring voxels based on the Euclidean distance between voxels and the peak using a 20mm FHWM Gaussian function (Radua et al., 2012, 2014). This method of estimation is similar to the estimation of activation likelihood used in peak-probability meta-analytic methods, but the use of effect sizes in the calculation increases the accuracy of estimation of the true signal (Radua et al., 2012). When the *t*-statistics of the peak coordinates were unknown (one study: Delgado et al., 2005), we imputed the effect size with the extent threshold reported in the study.

We conducted three random-effects models to compute a meta-analytic activation for each prosocial category identified using the Louvain community detection algorithm. Each individual study was weighted by the inverse sum of its variance plus the between-study variance as obtained by the DerSimonian-Laird estimator of heterogeneity (DerSimonian & Laird, 1986). Within this random-effects framework, studies with larger sample sizes or lower variability contribute more and effects are assumed to randomly vary between study samples. To assess statistical significance, we implemented a modified permutation test that empirically estimated a null distribution for each meta-analytic brain map. We thus tested the hypothesis that each map's true effect sizes were not the result of a random spatial association among studies within a prosocial category. We applied a threshold of $p<.005$ as recommended by Radua et al. (2012) to optimally balance specificity and sensitivity while yielding results approximately equivalent to $p<.05$ corrected for multiple comparisons. Reported z-scores are specified as SDM-Z, as they do not follow a standard normal distribution. We also conducted three pairwise comparisons of activation maps across each of the three prosocial categories, which followed the same procedures (see Supplemental Table S1).

The effect size maps were imported into AFNI (Cox, 1996) and a conjunction analysis was conducted to examine the overlap of consistently activated regions across altruism, cooperation, and equity. Conjunction was determined using 3dcalc by overlaying the thresholded meta-analytic maps for each category to determine activation overlay.

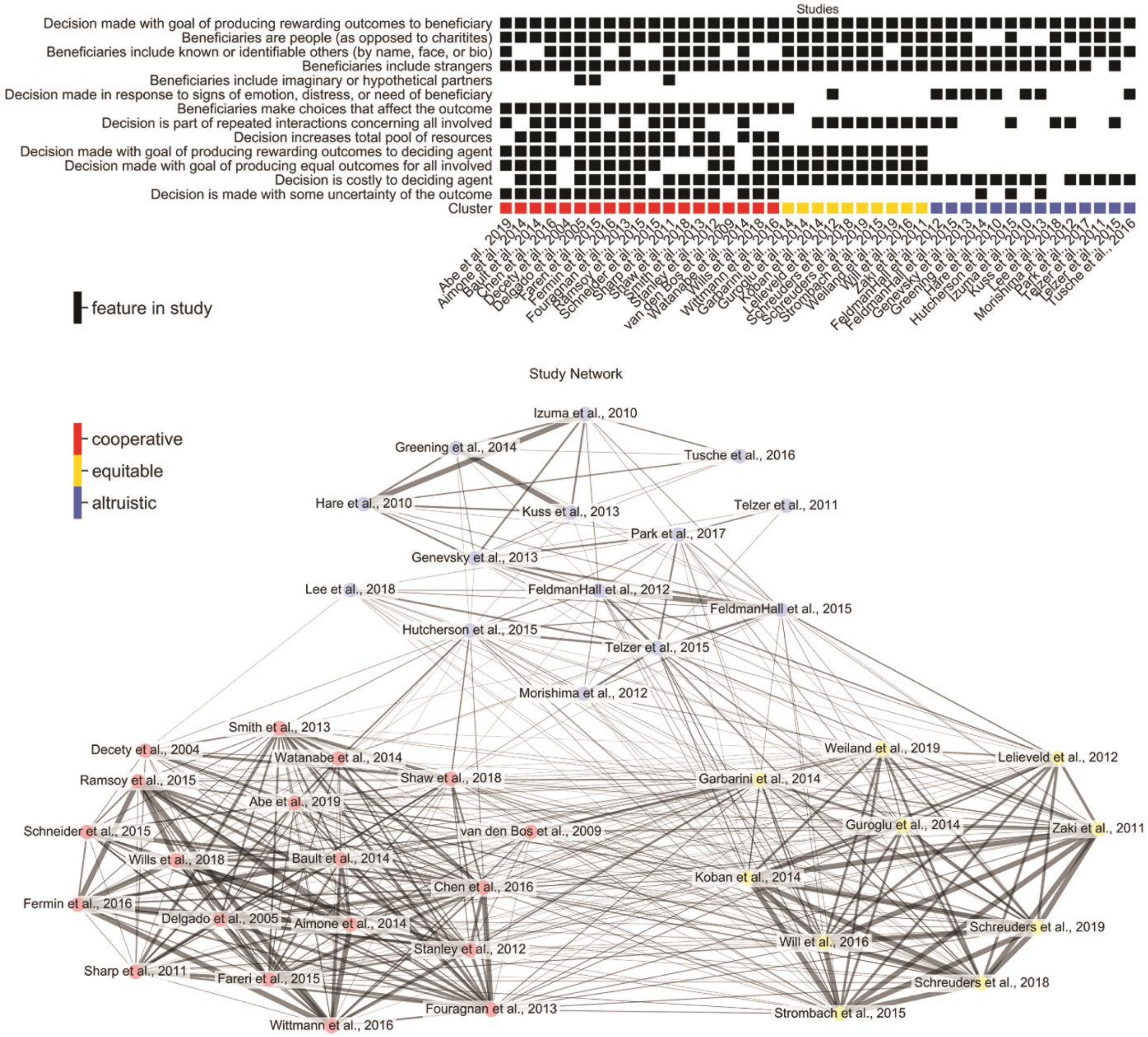
Results

Clustering tasks into categories of prosocial decision-making

The Louvain clustering algorithm revealed three clusters of prosocial decision-making tasks (Figure 1). Upon inspection, we labeled these *cooperative* decisions (blue), *equitable* decisions (yellow), and *altruistic* decisions (red) based on the task features shared within each cluster. Cooperative decisions (N=19; 8 maps, 11 coordinate tables) are those in which as decisions involve multiple agents acting prosocially to maximize resources (tasks included

Prisoner's Dilemmas, Public Goods Games, Ultimatum Games played by proposers, Trust Games, and non-economic cooperative tasks). Equitable decisions (N=10; 7 maps, 3 coordinate tables) are those in which decisions produce equitable outcomes for the agent and beneficiary (tasks included Dictator Games in which 50-50 splits were possible). Finally, altruistic decisions (N=14; 10 maps, 4 coordinate tables) are those in which agents make decisions to forgo resources without receiving anything in return (tasks included charitable donation tasks, Dictator Games in which 50-50 splits were not possible, Your Pain, My Gain tasks, and assistance tasks).

Figure 1. Graph depiction of study network generated from overlapping task features



Note. (top) Studies were labeled with task features to construct a bipartite graph where edges (black boxes in incidence matrix) exist between a feature node and study node if the study contrast contained the task feature. (bottom) The bipartite graph was projected onto a weighted study network, where wider edge weights between studies represented a larger Dice similarity coefficient (or greater similarity according to task features). A community detection algorithm based on modularity maximization revealed three clusters of studies, which were then labeled as cooperative (blue), equitable (yellow), and altruistic (red) decisions based on the task features shared within each cluster. For visualization purposes, the depicted graph was thresholded to only display edges with a Dice similarity coefficient greater than .70 (note that the actual graph was fully-connected).

Meta-analyses

Neural correlates of cooperative decisions. In the cooperative decision cluster, prosocial decisions (in contrast to selfish decisions) were associated with increased activation in right inferior frontal gyrus, bilateral subgenual ACC, left ventral striatum (including caudate nucleus), bilateral insula, bilateral MCC, left supramarginal gyrus extending to superior temporal gyrus, left lateral postcentral gyrus, bilateral ventral tegmental area (VTA), left thalamus, left precuneus, right cerebellum lobule VIII, and bilateral occipital cortex. We did not find significantly increased activation in any region during selfish versus cooperative decision-making (Figure 2, Table 2).

Neural correlates of equitable decisions. In the equitable decision cluster, prosocial decisions (in contrast to selfish decisions) were associated with increased activation in bilateral orbital frontal cortex (OFC), bilateral ventrolateral PFC, bilateral dorsolateral PFC, bilateral medial PFC including rostral ACC, bilateral ventral striatum and caudate, and left occipital cortex. Activation was increased during selfish relative to equitable decisions in left dorsolateral PFC, medial portion of left precentral gyrus, lateral portion of bilateral precentral gyrus, left thalamus, bilateral supramarginal gyrus, left inferior temporal gyrus, bilateral posterior superior temporal sulcus (STS), and bilateral occipital cortex (Figure 2, Table 2).

Neural correlates of altruistic decisions. In the altruistic decision cluster, prosocial decisions (in contrast to selfish decisions) were associated with increased activation in several regions including left ventromedial PFC, bilateral ACC and paracingulate gyrus, bilateral pre-SMA, bilateral anterior insula, right ventrolateral PFC, left bilateral dorsolateral PFC, thalamus, right ventral striatum, right precuneus, and bilateral interior parietal gyrus.

We also found increased activation during selfish decision-making, in contrast to altruistic decision-making, in several regions including right dorsolateral PFC, right ventrolateral PFC right putamen, bilateral posterior insula, bilateral precentral gyrus, right middle temporal gyrus (MTG), superior temporal gyrus (STG), right STS left parahippocampal gyrus, right superior parietal gyrus, bilateral middle occipital gyrus, and left cerebellum crus I and II (Figure 2, Table 2).

Pairwise meta-analytic map comparisons

Pairwise comparisons confirmed that the meta-analytic maps derived from each cluster of studies were distinct from one another. When comparing activation for cooperative relative to equitable decisions, we found increased activation in regions that included left ventrolateral PFC, left SMA, right caudate, left hippocampus, bilateral thalamus, left VTA, left supramarginal gyrus, and left superior parietal gyrus (Supplemental Figure 1a). We did not find increased activity in any region for equitable relative to cooperative decisions. When comparing activation for equitable relative to altruistic decisions, however, we found increased activation for equitable relative to altruistic decisions in regions that included right ventrolateral PFC, bilateral posterior insula, and right STG (Supplemental Figure 1b).

When comparing activation for altruistic relative to equitable decisions, we found increased activation in regions that included dorsolateral PFC, bilateral pre-SMA, left anterior insula, bilateral caudate, left thalamus, and left hippocampus (Supplemental Figure 1c). Finally, when comparing activation for altruistic relative to cooperative decisions, we found increased activation in regions that included left dorsolateral PFC, left SMA, left MTG, and left angular gyrus (Supplemental Figure 1d). We did not find increased activity in any region for cooperative relative to altruistic decisions. See Supplemental Table S1 for results.

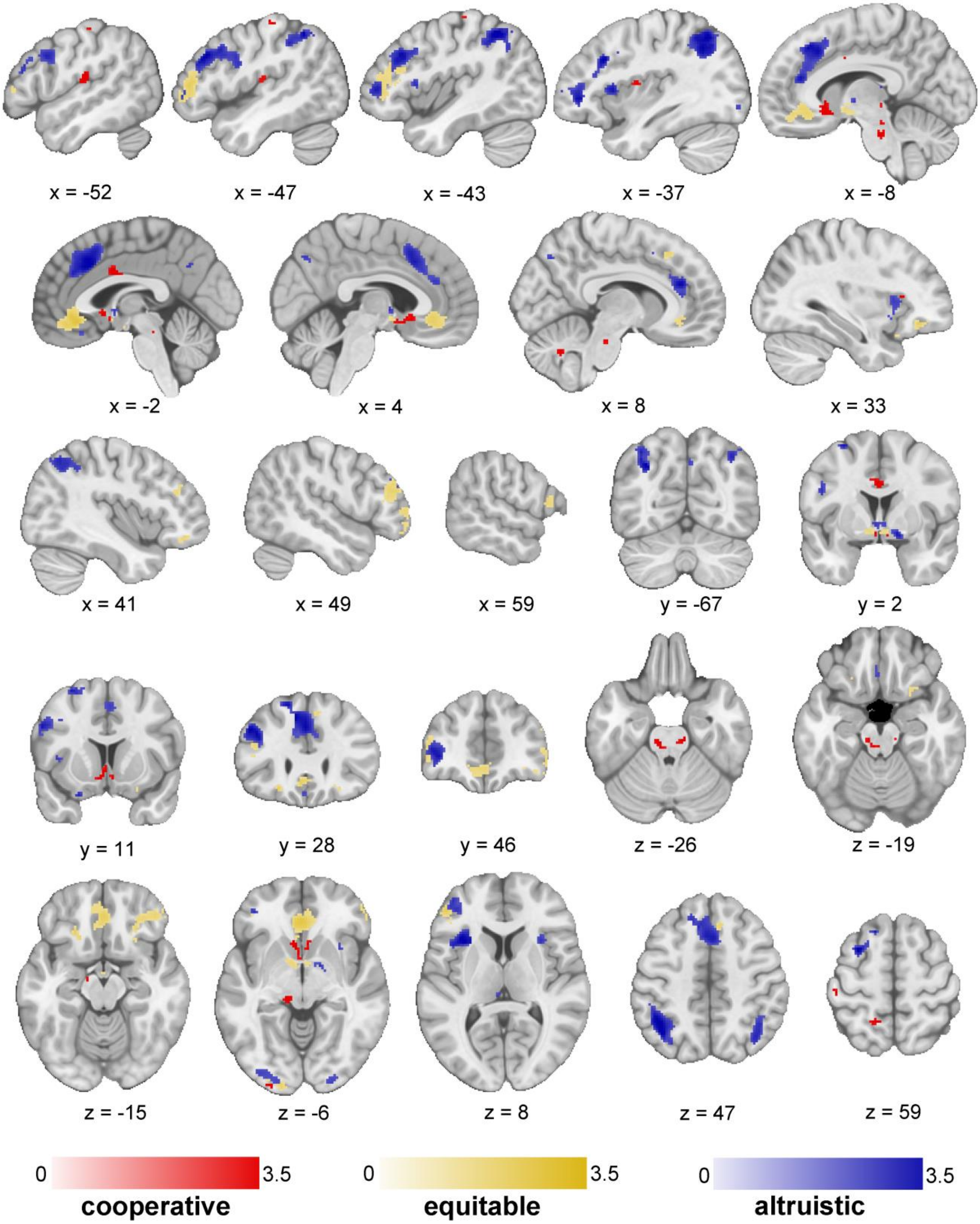
Conjunction across meta-analytic maps

Conjunction analyses identified overlapping regions of activation for altruism ∩ equity and cooperation ∩ equity. Overlapping activation for altruism ∩ equity was observed in bilateral dorsolateral PFC (BA9/46; left: [-44, 34, 20], k=19; right: [46, 40, 24], k=24), left ventrolateral PFC (BA10; [-42, 44, -4], k=32), and left visual cortex (BA18, [-20, -96, -8], k=34). Overlapping activity for equity ∩ cooperation was observed in bilateral ventral striatum (left: [-2, 4, -8]; right: [4, 4, -8]; k=4). We did not find overlapping activation for cooperation ∩ altruism nor across all three clusters of studies.

Data availability

All thresholded and unthresholded meta-analytic activation maps will be openly available on the Open Science Framework at https://osf.io/3vu9w/.

Figure 2. Thresholded results from meta-analyses.



Note. Results from each of three mixed-effects models displaying the main effects of decision-making category (cooperative: red; equitable: yellow; altruistic: blue). SDM-Z maps are corrected using a threshold of $p<.005$ and $k>10$.

Table 2. Results from the meta-analyses.

Cooperative > selfish decisions (increased activity)

Region	**Brodmann Area**	**SDM-Z**	**P**	**Voxels**	**MNI-x**	**MNI-y**	**MNI-z**
Inferior frontal gyrus, triangular part extending to insula (R)	48	3.824	0.000581	11	34	24	10
Ventral striatum extending to caudate and subgenual cingulate (L)		4.682	0.000021	89	-6	16	-2
Subgenual anterior cingulate cortex extending to olfactory cortex (R)	25	3.65	0.001053	28	4	12	-10
Middle cingulate cortex extending to paracingulate gyri (L, R)	24	3.853	0.000523	45	-2	4	32
Insula (L)	48	3.497	0.001740	14	-38	-4	12
Hippocampus (L)		3.408	0.002305	11	-16	-12	-12
Ventral tegmental area (L)		4.102	0.000212	51	-10	-24	-28
Ventral tegmental area (R)	30	3.814	0.000600	30	12	-24	-22
Postcentral gyrus, lateral part (L)	3	3.861	0.000509	14	-50	-24	60
Supramarginal gyrus extending to superior temporal gyrus (L)	48	3.888	0.000464	95	-60	-26	22
Thalamus (L)	50	3.931	0.000400	42	-14	-28	-6
White matter (middle cerebellar peduncles)		3.614	0.001188	13	14	-30	-34
Precuneus (L)		3.744	0.000764	14	-14	-50	58
Cerebellum lobule VIII (R)		3.568	0.001378	15	8	-62	-34
Middle occipital gyrus (L)	19	4.292	0.000103	49	-22	-88	18
Middle occipital gyrus (L, R)	18	3.736	0.000785	11	-28	-98	-6

Equitable > selfish decisions (increased activity)

Region	**Brodmann Area**	**SDM-Z**	**P**	**Voxels**	**MNI-x**	**MNI-y**	**MNI-z**

Inferior frontal gyrus, triangular part (L)	45	2.366	0.000113	305	-48	42	8
Middle frontal gyrus (R)	45	2.093	0.000468	205	48	40	20
Middle frontal gyrus extending to superior orbital gyrus (R)	10, 11	2.391	0.000098	234	30	38	-12
Anterior cingulate cortex extending to paracingulate gyri (L, R)	11	2.731	0.000014	488	-2	32	-10
Inferior frontal gyrus, orbital part (L)	11	2.302	0.000158	24	-24	26	-16
Superior frontal gyrus, medial part (R)	8	1.957	0.000904	18	8	26	48
Inferior frontal gyrus, opercular part (R)	44	1.996	0.000748	36	60	16	14
Striatum (L)		1.994	0.000752	68	-6	-6	-8
Middle occipital gyrus (L)	18	2.39	0.000099	113	-18	-98	-2

Equitable < selfish decisions (decreased activity)

Region	Brodmann Area	SDM-Z	P	Voxels	MNI-x	MNI-y	MNI-z
Middle frontal gyrus (L)	46	-2.534	0.000820	125	-22	42	26
Rolandic operculum (L)		-2.341	0.001943	36	-52	0	16
Inferior temporal gyrus (L)	20	-2.548	0.000771	15	-48	-2	-32
Precentral gyrus (R)	6	-2.533	0.000827	32	48	-4	40
Supplementary motor area (L)	6	-2.831	0.000197	252	-6	-4	68
Precentral gyrus (L)	6	-2.546	0.000778	27	-50	-8	52
Precentral gyrus (L)	6	-2.518	0.000883	17	-38	-12	40
Caudate nucleus (R)		-3.431	0.000007	91	18	-16	24
Precentral gyrus (L)	6	-2.355	0.001822	34	-30	-18	60
Inferior temporal gyrus (L)	20	-2.449	0.001206	12	-44	-18	-24
Thalamus extending to caudate (L)		-3.007	0.000079	288	-14	-20	20
Posterior insula (R)		-3.078	0.000054	27	32	-28	18
Supramarginal gyrus (R)	48	-2.651	0.000474	70	54	-36	28
Posterior cingulate cortex, ventral portion (R)	30	-2.347	0.001894	10	14	-38	6

Posterior superior temporal sulcus (R)	39	-3.021	0.000073	65	48	-40	-2
Supramarginal gyrus (L)	48	-2.581	0.000663	84	-50	-40	28
Inferior temporal gyrus (L)	20	-3.117	0.000043	105	-42	-44	-10
Posterior superior temporal sulcus (R)	39	-2.428	0.001329	11	44	-48	12
Postcentral gyrus (L)	5	-2.937	0.000113	182	-20	-52	64
Posterior superior temporal sulcus (L)	39	-2.592	0.000627	27	-40	-60	12
Extrastriate cortex extending to parahippocampal gyrus (R)	36	-3.239	0.000021	211	30	-62	6
Primary visual cortex (L)	17	-2.488	0.00101	57	-24	-68	8

Altruistic > selfish decisions (increased activity)

Region	Brodmann Area	SDM-Z	P	Voxels	MNI-x	MNI-y	MNI-z
Middle orbital gyrus (L)		3.397	0.000130	200	-38	44	0
Middle frontal gyrus (R)	45	2.827	0.001091	31	46	40	28
Anterior cingulate cortex extending to paracingulate gyri and pre-supplementary motor area (L, R)	32	3.871	0.000021	1278	10	36	22
Gyrus rectus, ventromedial (L, R)	11	2.564	0.002750	10	-2	30	-18
Middle frontal gyrus (L)	44	3.761	0.000031	568	-40	24	32
Anterior insula (L)	48	3.961	0.000014	203	-34	20	8
Anterior insula (R)	48	2.923	0.000773	64	30	18	6
Inferior frontal gyrus, orbital part (L)	38	2.982	0.000624	12	-24	12	-24
Ventral striatum (R)		3.15	0.000331	50	14	2	-8
Thalamus (L)		2.814	0.001146	40	-6	-6	0
Thalamus (L)		2.803	0.001191	23	-14	-30	16
Inferior parietal (excluding supramarginal and angular) gyri (L)	40	3.917	0.000017	705	-40	-54	50
Inferior parietal (excluding supramarginal and angular) gyri (R)	40	3.105	0.000394	219	42	-56	52

Precuneus (L)	7	2.83	0.001082	13	-4	-60	38
Precuneus (R)	7	2.96	0.000673	26	6	-70	44
Middle occipital gyrus (L)	18	2.761	0.001381	194	-22	-96	0
Inferior occipital gyrus (R)	18	2.594	0.002483	65	24	-96	-6

Altruistic < selfish decisions (decreased activity)

Region	Brodmann Area	SDM-Z	P	Voxels	MNI-x	MNI-y	MNI-z
Superior frontal gyrus, dorsolateral (R)	9	-2.602	0.002379	13	18	50	30
Inferior frontal gyrus, pars triangularis part (R)	48	-3.033	0.000539	104	48	34	6
Rolandic operculum (L)	6, 48	-3.372	0.000147	38	-56	0	8
Posterior insula (L)	84	-2.627	0.002198	15	-42	0	8
Anterior superior temporal sulcus (R)		-3.524	0.000080	69	48	-2	-14
Putamen (lenticular nucleus) (R)	48	-2.541	0.002913	10	30	-8	2
Posterior insula (R)	48	-3.902	0.000016	780	42	-10	2
Superior frontal gyrus, dorsolateral (R)	6	-3.119	0.000394	13	18	-10	72
Posterior insula (L)	48	-4.006	0.000010	70	-38	-14	6
Inferior temporal gyrus (L)	20	-3.238	0.000252	12	-42	-14	-28
Middle temporal gyrus (R)	21	-4.377	0.000002	232	58	-24	-8
Inferior temporal gyrus extending to parahippocampal gyrus (L)	20	-3.175	0.000317	32	-42	-26	-18
Postcentral gyrus, medial (R)	3	-2.787	0.001283	15	32	-34	50
Postcentral gyrus, medial (L)	3	-3.018	0.000567	18	-36	-38	56
Fusiform gyrus (L)	37	-3.181	0.000311	88	-22	-44	-18
Middle temporal gyrus (R)		-3.156	0.000342	195	48	-48	0
Superior parietal gyrus (R)	5	-3.479	0.000096	195	14	-50	70
Middle temporal gyrus (L)	37	-4.052	0.000008	764	-44	-68	6
Middle temporal gyrus (R)	37	-3.774	0.000029	136	42	-70	14
Cerebellum, crus I extending to crus II (L)		-2.648	0.002050	11	-22	-82	-30

Discussion

Using data from 43 unique fMRI studies that included 25 statistical maps and 18 coordinate tables across 1,423 subjects, we identified 13 features that distinguish prosocial decisions tasks. We used these features to generate a feature-based representation of prosocial decision tasks that classified prosocial decisions into three sub-clusters that we labeled cooperation, equity, and altruism. That the feature-based structure we generated identifies conceptually and motivationally coherent categories of prosocial decisions was supported by the results of our fMRI meta-analysis, which found evidence suggesting that each category of decision recruits diverging neural systems. Cooperative decisions primarily recruited regions such as dorsal and ventral striatum, VTA, and subgenual ACC. Equitable decisions recruited neural regions such as ventral striatum, dorsolateral PFC, and ventromedial PFC. Altruistic decisions recruited neural regions such as ventral striatum, dorsolateral PFC, ventromedial PFC, pre-SMA, dorsal ACC, and anterior insula.

Our approach demonstrates that the dozens of tasks that have been used to assess the neural correlates of prosocial decisions generally cluster together according to specific shared features. These tasks are adaptations of those used in studies of prosocial decision making outside the scanner and more broadly. We identified key features that distinguish prosocial decisions, including features related to the identity of the beneficiary, the nature of the interaction between agents and beneficiaries, and various outcomes associated with the decision. Then, using an unsupervised graph-based approach, we identified three clusters of tasks that tend to share common core features. For example, most cooperative decisions included outcomes that depended on the decisions of others, repeated interactions between the agent and beneficiaries, and decisions in conditions of uncertainty. Features shared by most equitable decisions included adherence to social norms such as producing equal outcomes, and unilateral decisions made by a single agent and that thus resulted in no uncertainty. Features shared by most altruistic decisions included outcomes that did not produce any benefit to the deciding agent and decisions that were made in response to the need or distress of the beneficiaries. Of note, the prosocial decision-making tasks included in this meta-analysis were described by the original authors using at least 10 different terms that did not consistently correspond to the features of the tasks being used—including cooperation, collaboration, reciprocity, trust, equity, fairness, prosocial behavior, interpersonal behavior, charitable behavior, and altruism—reinforcing the value of a clearer and more consistent prosocial decision-making task-space.

Supporting the identified task structure, the clusters yielded by our approach map closely onto the results of a previous behavioral characterization of prosocial paradigms that applied factor analysis to the behavioral outcomes of these paradigms (Böckler et al., 2016). These outcomes included the percentage of prosocial decisions during tasks, the ratio of other-regarding versus self-regarding decisions, average monetary donations, or summary scores of self-reported measures. Using a similar bottom-up, but otherwise completely distinct approach across 329 participants who completed multiple tasks assessing prosocial behavior, the authors identified clusters of prosocial tasks that correspond to those we identified: altruistically motivated prosocial behavior (corresponding to altruistic decisions), norm-motivated prosocial behavior (corresponding to equitable decisions), and strategically motivated prosocial behavior (corresponding to cooperative decisions). They also identified self-reported prosocial behavior (a category not included in our meta-analysis) as comprising a fourth distinct cluster.

Our findings also extend this work by showing that tasks cluster similarly even when completed by different participants across tasks. This suggests task features are crucial in determining the category of a prosocial decision and has implications for comparing results across different tasks, such as in previous meta-analyses. In addition, our findings suggest that even within a type of task, specific features may determine the category of prosocial decision at hand. For example, Dictator Games that offered an option to split available resources equally (50-50) clustered with equitable decisions, whereas Dictator Games with the option to make other prosocial splits clustered with altruistic decisions. The existence of strong norms related to equity may explain why 50-50 is the most common non-selfish split across Dictator Games when this choice is available (Engel, 2011). Future analyses of Dictator Game tasks, particularly those conducted using fMRI, could benefit from considering that there may be something unique about the decision to split resources equally (50%) rather than it simply existing as an option on a parametric continuum between 49% and 51%.

Our approach also yielded several important observations about the neural substrates of prosocial decisions. Notably, all three clusters of prosocial decisions recruited the striatum, a region consistently found to encode action value during learning and decision-making (Daw & Doya, 2006; Guitart-Masip et al., 2014). However, each category of decisions elicited activation in different regions within the striatum, possibly because actions varied across tasks. In some tasks (primarily cooperative and equitable decisions), prosocial decisions increased the agent's own welfare. In other tasks (primarily altruistic and equitable decisions), prosocial decisions meant forgoing resources. We found that cooperation recruited the left caudate and bilateral ventral striatum, equity recruited right caudate and bilateral ventral striatum, and altruism recruited right ventral striatum. The only overlap of striatal activation we identified occurred in a small volume of four voxels during both cooperative and equitable decisions. This conjunction may reflect the fact that cooperative and equitable decisions benefit both agents and beneficiaries, which may be recapitulated in ventral striatal activity.

This suggests the possibility of an additive effect of striatal activity—an interpretation consistent with observations of a parametric effect of striatal activation and reward magnitude for self (Miller et al., 2014). Ventral striatum is also preferentially engaged in response to rewarding social stimuli relative to rewarding nonsocial stimuli, for instance, when participants cooperate with a human partner relative to a computer partner despite identical monetary gains (Rilling et al., 2002, 2004). This also might explain why we did not observe any regions that were more active during selfish decisions than cooperative decisions. Cooperative decisions, which yield outcomes benefiting both the agent and other beneficiaries, are potentially more rewarding than decisions that only yield self-rewarding outcomes. Cooperative decisions also uniquely recruited activity in bilateral VTA, which projects dopamine to the ventral striatum in response to positive prediction errors and reward cues (D’Ardenne et al., 2008), and activity in which likely reflects the anticipation of both the self- and social-rewards gained from cooperating with others.

Striatal activation to anticipatory reward cues occurs within a larger subjective valuation system, which consistently involves activity in the medial PFC during reward-based decision-making (Bartra et al., 2013) and prosocial decision-making (Bellucci et al., 2020; Cutler & Campbell-Meiklejohn, 2019). As has been previously found (Cutler & Campbell-Meiklejohn, 2019), we observed activation in more anterior portions of the ventromedial PFC (including the rostral ACC) during equitable decisions which produce self-enhancing, norm-based outcomes and activation in more posterior portions during altruistic decisions. These results are consistent with a hypothesized spatial gradient of activation along the medial PFC during prosocial decision-making (Sul et al., 2015), which may integrate information about self and others to encode an overall value during a prosocial decision (Hutcherson et al., 2015).

Activation in both ventral striatum and medial PFC did not overlap between altruistic decisions and cooperative decisions. In contrast to altruistic decisions, which recruited the lingual gyrus portion of left ventromedial PFC, we found activation in bilateral subgenual ACC during tasks that require agents to cooperate with others to achieve a common goal. This suggests that altruistic and cooperative decision represent distinct processes, despite frequent conflation these two categories of decisions in the literature, for example, when altruistic behavior (choosing to benefit others without any self-gain) is labeled "cooperation" (Balliet et al., 2014; Declerck et al., 2013; Gintis, 2014; Peysakhovich et al., 2014; Yang et al., 2019). Because the subgenual ACC supports prosocial learning computations (Christopoulos & King-Casas, 2014; Lockwood et al., 2016) as well as preferences for socially rewarding outcomes (Smith et al., 2010), it may also play a role in updating expectations of others' actions or the value of others' outcomes during iterative cooperative decision-making.

In addition to a subjective-valuation sub-system, altruistic decisions seemed to recruit two other distinct sub-systems underlying goal-directed behavior and empathy, respectively. The goal-directed sub-system included regions typically implicated in controlling action and directing goal-directed behaviors, including the lateral PFC (Hoshi & Tanji, 2004; Kaller et al., 2011; Morris et al., 2014). Importantly, activation for equitable and altruistic decision-making overlapped in dorsolateral PFC, involved in modulating subjective value representations (Carlson & Crockett, 2018; Tusche & Hutcherson, 2018) and in making norm-related decisions (Baumgartner et al., 2011; Knoch et al., 2006). Thus, it may play an important role in guiding prosocial action in accordance with abstract, social rules during social decision-making tasks (Bellucci et al., 2020).

The second sub-system comprised regions commonly involved in representing and empathizing with the distress of others, and included the dorsal ACC, pre-SMA, and anterior insula (Ashar et al., 2017; Decety & Lamm, 2006; Lamm et al., 2011). Activation in these regions emerged only during altruistic decisions, consistent with the theory that affective resonance with others' distress give rise to empathic concern and altruistic motivation, which is a primary motivator of prosocial behavior in the absence of cooperative or equity-maintaining goals (Batson, 2009, 2011; Brethel-Haurwitz et al., 2018; Decety et al., 2016; O'Connell et al., 2019). This finding was likely driven by eight out of the fourteen identified altruistic studies including stimuli that depicted or implied the need or distress of beneficiaries. We also found activation in the precuneus—a key node of the mentalizing system (Koster-Hale et al., 2017)—during altruistic decisions. This region has also been found to be active in response to observing emotional suffering (Immordino-Yang et al., 2009; Masten et al., 2011; Meyer et al., 2013). Although some hypothesize the right TPJ—another core region of the mentalizing system—to be recruited during prosocial decision-making (Chakroff & Young, 2014; Parnamets et al., 2020), we did not find differential activation in this region across prosocial relative to selfish decisions. It is possible that we did not observe differences in mean activation because both selfish and prosocial decisions require the maintenance of others' beliefs and intentions, whereas studies finding TPJ activation usually contrast decisions pertaining to other people with hypothetical decisions pertaining to imaginary people (or computers) (FeldmanHall et al., 2012).

Limitations and Future Directions

These results should be considered in light of some limitations. We could not obtain complete data from a number of potentially relevant studies. At least 69 studies would have been eligible for the analysis if we had been able to retrieve the necessary data. In addition, while the 13 features we identified captured the distinctions across the tasks included in our analysis, they may not be representative of all features that could potentially overlap across tasks. With access to more data, future work could test how clustering algorithms such as the one we employed can generalize to unseen tasks with different combinations of features, or even generate new tasks with unique sets of task features.

We identified how tasks were inter-related using a bottom-up feature-based approach (assuming equally important features), and how these inter-relationships give rise to overarching prosocial categories. This is in contrast to the more top-down approaches based on expert-models that have been used to map cognitive constructs like creativity and control onto tasks (Kenett et al., 2020; Lenartowicz et al., 2010). Future work could combine these approaches to generate a finer-grained task-space for prosocial decision-making including all possible levels of its cognitive ontology: the categories identified in the present study, finer-detailed sub-categories, task paradigms and their features weighted according to their relative importance, and contrast estimates. In so doing, we could go further in solving discrepancies within the prosocial decision-making literature, such as delineating more specific categories of prosocial decision-making within the identified task-space, which may only reflect the top level of a prosocial decision-making hierarchy. For example, active decisions to forgive (Fourie et al., 2020), norm-enforcing decisions (i.e., social influence on agreements or valuation) (Chang & Sanfey, 2013; Wu et al., 2016; Yang et al., 2019; Zinchenko & Arsalidou, 2018), and third-party altruistic punishment decisions for norm violations (Buckholtz et al., 2008; David et al., 2017; Fehr et al., 2004; Jordan et al., 2016) were not considered in this study because we sought to only examine decisions that directly benefited another person, but may reflect more specific prosocial decisions under the umbrella of the identified categories.

As with many neuroimaging tasks, prosocial decision-making tasks adapted for neuroimaging are tightly controlled, and often designed to minimize variability and maximize the statistical power of detecting effects. However, they are not designed with high ecological validity and may not map onto the contexts of real-world prosocial decisions such as holding a door open, splitting a meal, volunteering, or donating blood or an organ. Instead, they are primarily monetary in nature, repetitive, and may increase behaviors related to social desirability in the laboratory (Richman et al., 1999). Recent work has focused on making neuroimaging paradigms more "naturalistic," such as viewing or listening to narratives or interacting in real-time with another person in the laboratory (Hasson & Frith, 2016; Redcay & Moraczewski, 2019; Redcay & Schilbach, 2019; Wheatley et al., 2019) or in the real-world (Dikker et al., 2019). Other work has concentrated on characterizing the behavioral and neural features of individuals who engage in extreme forms of real-world prosociality (Brethel-Haurwitz et al., 2018; Marsh et al., 2014; O’Connell et al., 2019; Vekaria et al., 2020). Understanding the neurobiology underlying more ecologically-valid altruistic decisions will be crucial for understanding the broader picture of prosocial decision-making.

Related to this, we were not able to consider how individual differences in phenotypic traits may contribute to the neural activity patterns observed across prosocial decision-making tasks due to the limited number of studies that collect or report consistent data on participant characteristics, yet this remains an open question. Finally, we only considered univariate maps that contrasted prosocial versus selfish decisions (or assessed a parametric increase). Due to the high variability of contrasts across studies, this allowed us to generate consistent neuroimaging contrasts that only indexed activation during the decision phase across studies. This approach is distinct from other meta-analytic work compiling coordinate-based maps across any contrasts in prosocial tasks (Bellucci et al., 2020; Yang et al., 2019), which run the risk of creating dependence across experiment maps that negatively impacts the validity of meta-analytic results (Müller et al., 2018). Ideally, future work would incorporate neural activation maps derived from computational modeling of behavior (Charpentier & O'Doherty, 2018; Lockwood et al., 2020; Lockwood & Klein-Flügge, 2020; Suzuki & O'Doherty, 2020; Tognoli et al., 2017), which holds promise for understanding individual differences in social learning and decision-making (Patzelt et al., 2018). For example, computational modeling of decisions can identify latent subjective states (e.g., mood, anxiety), beliefs about others (e.g., trust, morality), or other subjective biases about agents that are not directly observable from behavior. Thus, mapping these latent parameters onto the task features that give rise to them and their neural representations will be a necessary next step in characterizing a cognitive ontology of prosociality.

Conclusion

Despite limitations, the present study provides a framework for understanding how prosocial decisions are inter-related and distinct, and can be applied to a variety of experimental task paradigms. Using a bottom-up approach, we identified a feature-based representation of the task-space underlying prosocial decisions. Results revealed that three clusters of prosocial decisions identified this way—cooperative, equitable, and altruistic decisions—recruit neural systems that diverge in ways that shed light on the key motivations and mechanisms that support each category of prosocial decision compared to selfish decisions. These findings clarify some of the existing heterogeneity in how prosociality is conceptualized and generate insight for future research in task paradigm development and the improvement of formal cognitive ontologies.

References

Abe, M. O., Koike, T., Okazaki, S., Sugawara, S. K., Takahashi, K., Watanabe, K., & Sadato, N. (2019). Neural correlates of online cooperation during joint force production. *NeuroImage*, *191*, 150–161. https://doi.org/10.1016/j.neuroimage.2019.02.003

Aimone, J. A., Houser, D., & Weber, B. (2014). Neural signatures of betrayal aversion: An fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1782). https://doi.org/10.1098/rspb.2013.2127

Aknin, L. B., Dunn, E. W., & Norton, M. I. (2012). Happiness Runs in a Circular Motion: Evidence for a Positive Feedback Loop between Prosocial Spending and Happiness. *Journal of Happiness Studies*, *13*(2), 347–355. https://doi.org/10.1007/s10902-011-9267-5

Ashar, Y. K., Andrews-Hanna, J. R., Dimidjian, S., & Wager, T. D. (2017). Empathic Care and Distress: Predictive Brain Markers and Dissociable Brain Systems. *Neuron*, *94*(6), 1263-1273.e4. https://doi.org/10.1016/j.neuron.2017.05.014

Balliet, D., Mulder, L. B., & Van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, *37*(4), 594–615. https://doi.org/10.1037/a0023489

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: a meta-analysis. *Psychological Bulletin*, *140*(6), 1556–1581. https://doi.org/10.1037/a0037737

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Batson, C. D. (2009). These Things Called Empathy: Eight Related but Distinct Phenomena. In *The Social Neuroscience of Empathy* (pp. 4–14). https://doi.org/10.7551/mitpress/9780262012973.003.0002

Batson, C. D. (2011). Altruism in Humans. In *Altruism in Humans*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195341065.001.0001

Batson, C. D., & Powell, A. A. (2003). Altruism and Prosocial Behavior. In *Handbook of Psychology*. https://doi.org/10.1002/0471264385.wei0519

Bault, N., Pelloux, B., Fahrenfort, J. J., Ridderinkhof, K. R., & Van Winden, F. (2014). Neural dynamics of social tie formation in economic decision-making. *Social Cognitive and Affective Neuroscience*, *10*(6), 877–884. https://doi.org/10.1093/scan/nsu138

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., & Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, *14*(11), 1468–1474. https://doi.org/10.1038/nn.2933

Bellucci, G., Camilleri, J. A., Eickhoff, S. B., & Krueger, F. (2020). Neural signatures of prosocial behaviors. *Neuroscience and Biobehavioral Reviews*, *118*, 186–195. https://doi.org/10.1016/j.neubiorev.2020.07.006

Bellucci, G., Chernyak, S. V., Goodyear, K., Eickhoff, S. B., & Krueger, F. (2017). Neural signatures of trust in reciprocity: A coordinate-based meta-analysis. *Human Brain Mapping*, *38*(3), 1233–1248. https://doi.org/10.1002/hbm.23451

Blondel, V., Guillaume, J., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of community hierarchies in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), 1–12. https://doi.org/10.1088/1742-5468/2008/10/P10008

Böckler, A., Tusche, A., & Singer, T. (2016). The Structure of Human Prosociality: Differentiating Altruistically Motivated, Norm Motivated, Strategically Motivated, and Self-Reported Prosocial Behavior. *Social Psychological and Personality Science*, *7*(6), 530–541. https://doi.org/10.1177/1948550616639650

Brethel-Haurwitz, K. M., Cardinale, E. M., Vekaria, K. M., Robertson, E. L., Walitt, B., VanMeter, J. W., & Marsh, A. A. (2018). Extraordinary Altruists Exhibit Enhanced Self–Other Overlap in Neural Responses to Distress. *Psychological Science*, *29*(10), 1–11. https://doi.org/10.1177/0956797618779590

Buckholtz, J. W., Asplund, C. L., Dux, P. E., Zald, D. H., Gore, J. C., Jones, O. D., & Marois, R. (2008). The Neural Correlates of Third-Party Punishment. *Neuron*, *60*(5), 930–940. https://doi.org/10.1016/j.neuron.2008.10.016

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., Huber, J., Isler, K., Kosonen, Z. K., Martins, E., Meulman, E. J., Richiger, R., Rueth, K., Spillmann, B., Wiesendanger, S., & Van Schaik, C. P. (2014). The evolutionary origin of human hyper-cooperation. *Nature Communications*, *5*(1), 1–9. https://doi.org/10.1038/ncomms5747

Carlson, R. W., & Crockett, M. J. (2018). The lateral prefrontal cortex and moral goal pursuit. *Current Opinion in Psychology*, *24*, 77–82. https://doi.org/10.1016/j.copsyc.2018.09.007

Chakroff, A., & Young, L. (2014). The prosocial brain: Perceiving others in need and acting on it. In L. M. Padilla-Walker & G. Carlo (Eds.), *Prosocial development: A multidimensional approach* (pp. 90–111). Oxford University Press.

Chang, L. J., & Sanfey, A. G. (2013). Great expectations: Neural computations underlying the use of social norms in decision-making. *Social Cognitive and Affective Neuroscience*, *8*(3), 277–284. https://doi.org/10.1093/scan/nsr094

Charpentier, C. J., & O'Doherty, J. P. (2018). The application of computational models to social neuroscience: promises and pitfalls. *Social Neuroscience*, *13*(6), 637–647. https://doi.org/10.1080/17470919.2018.1518834

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, *14*(1), 47–83. https://doi.org/10.1007/s10683-010-9257-1

Chen, X., Hackett, P. D., DeMarco, A. C., Feng, C., Stair, S., Haroon, E., Ditzen, B., Pagnoni, G., & Rilling, J. K. (2016). Effects of oxytocin and vasopressin on the neural response to unreciprocated cooperation within brain regions involved in stress and anxiety in men and women. *Brain Imaging and Behavior*, *10*(2), 581–593. https://doi.org/10.1007/s11682-015-9411-7

Christopoulos, G. I., & King-Casas, B. (2014). With you or against you: Social orientation dependent learning signals guide actions made for others. *NeuroImage*, *104*, 326–335. https://doi.org/10.1016/j.neuroimage.2014.09.011

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Curry, O. S., Rowland, L. A., Van Lissa, C. J., Zlotowitz, S., McAlaney, J., & Whitehouse, H. (2018). Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor. *Journal of Experimental Social Psychology*, *76*, 320–329. https://doi.org/10.1016/j.jesp.2018.02.014

Cutler, J., & Campbell-Meiklejohn, D. (2019). A comparative fMRI meta-analysis of altruistic and strategic decisions to give. *NeuroImage*, *184*, 227–241. https://doi.org/10.1016/j.neuroimage.2018.09.009

D’Ardenne, K., McClure, S. M., Nystrom, L. E., Cohen, J. D., D’Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267.

https://doi.org/10.1126/science.1150605

David, B., Hu, Y., Krüger, F., & Weber, B. (2017). Other-regarding attention focus modulates third-party altruistic choice: An fMRI study. *Scientific Reports*, *7*, 1–12. https://doi.org/10.1038/srep43024

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204. https://doi.org/10.1016/j.conb.2006.03.006

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, *59*, 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Decety, J., Bartal, I. B. A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: Highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150077. https://doi.org/10.1098/rstb.2015.0077

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Decety, J., & Lamm, C. (2006). Human Empathy Through the Lens of Social Neuroscience. *The Scientific World Journal*, *6*, 1146–1163. https://doi.org/10.1100/tsw.2006.221

Declerck, C. H., Boone, C., & Emonds, G. (2013). When do people cooperate? The neuroeconomics of prosocial decision making. *Brain and Cognition*, *81*(1), 95–117. https://doi.org/10.1016/j.bandc.2012.09.009

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, *8*(11), 1611–1618. https://doi.org/10.1038/nn1575

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology*, *26*(3), 297–302. https://doi.org/10.2307/1932409

Dikker, S., Michalareas, G., Oostrik, M., Serafimaki, A., Kahraman, M., Struiksma, M. E., & Poeppel, D. (2019). Crowdsourcing neuroscience: inter-brain coupling during face-to-face interactions outside the laboratory. *BioRxiv*, 1–49.

Dunn, E. W., Aknin, L. B., & Norton, M. I. (2008). Spending money on others promotes happiness. *Science*, *319*(5870), 1687–1688. https://doi.org/10.1126/science.1150952

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, *30*(9), 2907–2926. https://doi.org/10.1002/hbm.20718

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7

Fareri, D. S., Chang, L. J., & Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in Neuroscience*, *6*, 1–17. https://doi.org/10.3389/fnins.2012.00148

Fareri, D. S., Chang, L. J., & Delgado, M. R. (2015). Computational Substrates of Social Value in Interpersonal Collaboration. *Journal of Neuroscience*, *35*(21), 8170–8180. https://doi.org/10.1523/jneurosci.4775-14.2015

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences*, *11*(10), 419–427. https://doi.org/10.1016/j.tics.2007.09.002

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785–791. https://doi.org/10.1038/nature02043

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, *13*(1), 1–25. https://doi.org/10.1007/s12110-002-1012-7

Fehr, E., Rockenbach, B., Borst, A., & Schultz, W. (2004). Human altruism: economic, neural, and evolutionary perspectives. *Current Opinion in Neurobiology*, *14*(6), 784–790. https://doi.org/10.1016/j.conb.2004.10.007

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, *2*(7), 458–468. https://doi.org/10.1038/s41562-018-0385-5

FeldmanHall, O. (2017). How Does Social Network Position Influence Prosocial Behavior? *Trends in Cognitive Sciences*, *21*(8), 572–573. https://doi.org/10.1016/j.tics.2017.05.005

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage*, *105*, 347–356. https://doi.org/10.1016/j.neuroimage.2014.10.043

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., & Mobbs, D. (2012). Differential neural circuitry and self-interest in real vs hypothetical moral decisions. *Social Cognitive and Affective Neuroscience*, *7*(7), 743–751. https://doi.org/10.1093/scan/nss069

Fermin, A. S. R., Sakagami, M., Kiyonari, T., Li, Y., Matsumoto, Y., & Yamagishi, T. (2016).

Representation of economic preferences in the structure and function of the amygdala and prefrontal cortex. *Scientific Reports*, *6*, 1–11. https://doi.org/10.1038/srep20982

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., & Coricelli, G. (2013). Reputational Priors Magnify Striatal Responses to Violations of Trust. *Journal of Neuroscience*, *33*(8), 3602–3611. https://doi.org/10.1523/jneurosci.3086-12.2013

Fourie, M. M., Hortensius, R., & Decety, J. (2020). Parsing the components of forgiveness: Psychological and neural mechanisms. *Neuroscience and Biobehavioral Reviews*, *112*, 437–451. https://doi.org/10.1016/j.neubiorev.2020.02.020

Fowler, J. H., & Christakis, N. A. (2010). Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences*, *107*(12), 5334–5338. https://doi.org/10.1073/pnas.0913149107

Garbarini, F., Boero, R., D’Agata, F., Bravo, G., Mosso, C., Cauda, F., Duca, S., Geminiani, G., & Sacco, K. (2014). Neural correlates of gender differences in reputation building. *PLoS ONE*, *9*(9), e106285. https://doi.org/10.1371/journal.pone.0106285

Genevsky, A., Vastfjall, D., Slovic, P., & Knutson, B. (2013). Neural Underpinnings of the Identifiable Victim Effect: Affect Shifts Preferences for Giving. *Journal of Neuroscience*, *33*(43), 17188–17196. https://doi.org/10.1523/JNEUROSCI.2348-13.2013

Gintis, H. (2014). The bounds of reason: Game theory and the unification of the behavioral sciences. In *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press. https://doi.org/10.1080/10848770.2013.838043

Greening, S., Norton, L., Virani, K., Ty, A., Mitchell, D., & Finger, E. (2014). Individual differences in the anterior insula are associated with the likelihood of financially helping versus harming others. *Cognitive, Affective and Behavioral Neuroscience*, *14*(1), 266–277.

https://doi.org/10.3758/s13415-013-0213-3

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003

Guroglu, B., Will, G. J., & Crone, E. A. (2014). Neural correlates of advantageous and disadvantageous inequity in sharing decisions. *PLoS ONE*, *9*(9), 12–14. https://doi.org/10.1371/journal.pone.0107996

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, *3*(4), 367–388. https://doi.org/10.1016/0167-2681(82)90011-7

Hackel, L. M., Zaki, J., & Van Bavel, J. J. (2017). Social identity shapes social valuation: evidence from prosocial behavior and vicarious reward. *Social Cognitive and Affective Neuroscience*, *12*(8), 1219–1228. https://doi.org/10.1093/scan/nsx045

Hare, B. (2017). Survival of the Friendliest: Homo Sapiens Evolved via Selection for Prosociality. *Annual Review of Psychology*, *68*(1), 155–186. https://doi.org/10.1146/annurev-psych-010416-044201

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., & Rangel, A. (2010). Value Computations in Ventral Medial Prefrontal Cortex during Charitable Decision Making Incorporate Input from Regions Involved in Social Cognition. *Journal of Neuroscience*, *30*(2), 583–590. https://doi.org/10.1523/JNEUROSCI.4089-09.2010

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), 20150366.

https://doi.org/10.1098/rstb.2015.0366

Henrich, J., Boyd, R., Bowles, S., Gintis, H., Fehr, E., Camerer, C., McElreath, R., Gurven, M., Hill, K., Barr, A., Ensminger, J., Tracer, D., Marlow, F., Patton, J., Alvard, M., Gil-White, F. J., & Henrich, N. (2005). “Economic man” in cross-cultural perspective: Ethnography and experiments from 15 small-scale societies. *Behavioral and Brain Sciences*, *28*(6), 795815.

Hill, P. F., Yi, R., Spreng, R. N., & Diana, R. A. (2017). Neural congruence between intertemporal and interpersonal self-control: Evidence from delay and social discounting. *NeuroImage*, *162*, 186–198. https://doi.org/10.1016/j.neuroimage.2017.08.071

Hoshi, E., & Tanji, J. (2004). Area-selective neuronal activity in the dorsolateral prefrontal cortex for information retrieval and action planning. *Journal of Neurophysiology*, *91*(6), 2707–2722. https://doi.org/10.1152/jn.00904.2003

Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A Neurocomputational Model of Altruistic Choice and Its Implications. *Neuron*, *87*(2), 451–463. https://doi.org/10.1016/j.neuron.2015.06.031

Immordino-Yang, M. H., McColl, A., Damasio, H., & Damasio, A. (2009). Neural correlates of admiration and compassion. *Proceedings of the National Academy of Sciences*, *106*(19), 8021–8026. https://doi.org/10.1073/pnas.0810363106

Izuma, K., Saito, D. N., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience*, *22*(4), 621–631. https://doi.org/10.1162/jocn.2009.21228

Jenni, K., & Loewenstein, G. (1997). Explaining the “Identifiable Victim Effect.” *Journal of Risk and Uncertainty*, *14*(3), 235–257. https://doi.org/10.1023/A:1007740225484

Jones, B. A., & Rachlin, H. (2006). Social discounting. *Psychological Science*, *17*(4), 283–286.

https://doi.org/10.1111/j.1467-9280.2006.01699.x

Jones, B. A., & Rachlin, H. (2009). Delay, Probability, and Social Discounting in A Public Goods Game. *Journal of the Experimental Analysis of Behavior*, *91*(1), 61–73. https://doi.org/10.1901/jeab.2009.91-61

Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, *530*(7591), 473–476. https://doi.org/10.1038/nature16981

Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., & Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cerebral Cortex*, *21*(2), 307–317. https://doi.org/10.1093/cercor/bhq096

Kenett, Y. N., Kraemer, D. J. M., Alfred, K. L., Colaizzi, G. A., Cortes, R. A., & Green, A. E. (2020). Developing a neurally informed ontology of creativity measurement. *NeuroImage*, *221*, 117166. https://doi.org/10.1016/j.neuroimage.2020.117166

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, *314*(5800), 829–832. https://doi.org/10.1126/science.1129156

Koban, L., Pichon, S., & Vuilleumier, P. (2014). Responses of medial and ventrolateral prefrontal cortex to interpersonal conflict for resources. *Social Cognitive and Affective Neuroscience*, *9*(5), 561–569. https://doi.org/10.1093/scan/nst020

Kogut, T., & Ritov, I. (2005). The “identified victim” effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18*(3), 157–167. https://doi.org/10.1002/bdm.492

Koster-Hale, J., Richardson, H., Velez, N., Asaba, M., Young, L., & Saxe, R. (2017). Mentalizing regions represent distributed, continuous, and abstract dimensions of others’ beliefs.

NeuroImage, *161*, 9–18. https://doi.org/10.1016/j.neuroimage.2017.08.026

Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, *11*(3), 495–524. https://doi.org/10.1111/jeea.12006

Kuss, K., Falk, A., Trautner, P., Elger, C. E., Weber, B., & Fliessbach, K. (2013). A reward prediction error for charitable donations reveals outcome orientation of donators. *Social Cognitive and Affective Neuroscience*, *8*(2), 216–223. https://doi.org/10.1093/scan/nsr088

Kuss, K., Falk, A., Trautner, P., Montag, C., Weber, B., & Fliessbach, K. (2015). Neuronal correlates of social decision making are influenced by social value orientation-an fMRI study. *Frontiers in Behavioral Neuroscience*, *9*(40), 1–8. https://doi.org/10.3389/fnbeh.2015.00040

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Lee, M., Ahn, H. S., Kwon, S. K., & Kim, S. (2018). Cooperative and Competitive Contextual Effects on Social Cognitive and Empathic Neural Responses. *Frontiers in Human Neuroscience*, *12*, 1–17. https://doi.org/10.3389/fnhum.2018.00218

Lee, S., & Feeley, T. H. (2016). The identifiable victim effect: a meta-analytic review. *Social Influence*, *11*(3), 199–215. https://doi.org/10.1080/15534510.2016.1216891

Lelieveld, G. J., van Dijk, E., Güroğlu, B., van Beest, I., van Kleef, G. A., Rombouts, S. A. R. B., & Crone, E. A. (2013). Behavioral and neural reactions to emotions of others in the distribution of resources. *Social Neuroscience*, *8*(1), 52–62. https://doi.org/10.1080/17470919.2012.735621

Lenartowicz, A., Kalar, D. J., Congdon, E., & Poldrack, R. A. (2010). Towards an Ontology of Cognitive Control. *Topics in Cognitive Science*, *2*(4), 678–692. https://doi.org/10.1111/j.1756-8765.2010.01100.x

Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is There a “Social” Brain? Implementations and Algorithms. *Trends in Cognitive Sciences*, *24*(10), 802–813. https://doi.org/10.1016/j.tics.2020.06.011

Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences*, *113*(35), 9763–9768. https://doi.org/10.1073/pnas.1603198113

Lockwood, P. L., & Klein-Flügge, M. (2020). Computational modelling of social cognition and behaviour—a reinforcement learning primer. *Social Cognitive and Affective Neuroscience*, 1–11. https://doi.org/10.1093/scan/nsaa040

López-Pérez, R. (2008). Aversion to norm-breaking: A model. *Games and Economic Behavior*, *64*(1), 237–267. https://doi.org/10.1016/j.geb.2007.10.009

Marsh, A. A. (2016). Neural, cognitive, and evolutionary foundations of human altruism. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(1), 59–71. https://doi.org/10.1002/wcs.1377

Marsh, A. A., Stoycos, S. A., Brethel-Haurwitz, K. M., Robinson, P., VanMeter, J. W., & Cardinale, E. M. (2014). Neural and cognitive characteristics of extraordinary altruists. *Proceedings of the National Academy of Sciences*, *111*(42), 15036–15041. https://doi.org/10.1073/pnas.1408440111

Masten, C. L., Morelli, S. A., & Eisenberger, N. I. (2011). An fMRI investigation of empathy for “social pain” and subsequent prosocial behavior. *NeuroImage*, *55*(1), 381–388.

https://doi.org/10.1016/j.neuroimage.2010.11.060

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.2307/258792

Meyer, M. L., Masten, C. L., Ma, Y., Wang, C., Shi, Z., Eisenberger, N. I., & Han, S. (2013). Empathy for the social suffering of friends and strangers recruits distinct patterns of brain activation. *Social Cognitive and Affective Neuroscience*, *8*(4), 446–454. https://doi.org/10.1093/scan/nss019

Milinski, M., Semmann, D., Bakker, T. C. M., & Krambeck, H. J. (2001). Cooperation through indirect reciprocity: Image scoring or standing strategy? *Proceedings of the Royal Society B: Biological Sciences*, *268*(1484), 2495–2501. https://doi.org/10.1098/rspb.2001.1809

Miller, E. M., Shankar, M. U., Knutson, B., & McClure, S. M. (2014). Dissociating Motivation from Reward in Human Striatal Activity. *Journal of Cognitive Neuroscience*, *26*(5), 1075–1084. https://doi.org/10.1162/jocn

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., Schweizer, S., Frith, C. D., & Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science*, *324*(5929), 900. https://doi.org/10.1126/science.1170539

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., & Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, *75*(1), 73–79. https://doi.org/10.1016/j.neuron.2012.05.021

Morris, R. W., Dezfouli, A., Griffiths, K. R., & Balleine, B. W. (2014). Action-value comparisons in the dorsolateral prefrontal cortex control choice between goal-directed actions. *Nature Communications*, *5*(1), 1–10. https://doi.org/10.1038/ncomms5390

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., Tench, C. R., Yarkoni, T., Nichols, T. E., Turkeltaub, P. E., Wager, T. D., & Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, *84*, 151–161. https://doi.org/10.1016/j.neubiorev.2017.11.012

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

O'Connell, K., Brethel-Haurwitz, K. M., Rhoads, S. A., Cardinale, E. M., Vekaria, K. M., Robertson, E. L., Walitt, B., Vanmeter, J. W., & Marsh, A. A. (2019). Increased similarity of neural responses to experienced and empathic distress in costly altruism. *Scientific Reports*, *9*(10774), 1–11. https://doi.org/10.1038/s41598-019-47196-3

Park, B. K., Blevins, E., Knutson, B., & Tsai, J. L. (2017). Neurocultural evidence that ideal affect match promotes giving. *Social Cognitive and Affective Neuroscience*, *12*(7), 1083–1096. https://doi.org/10.1093/scan/nsx047

Parnamets, P., Shuster, A., Reinero, D. A., & Van Bavel, J. J. (2020). A value-based framework for understanding cooperation. *Psychological Science*, *29*(3), 227–234. https://doi.org/10.1177/0963721420906200

Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational Phenotyping: Using Models to Understand Individual Differences in Personality, Development, and Mental Illness. *Personality Neuroscience*, *1*, 1–9. https://doi.org/10.1017/pen.2018.14

Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a "cooperative phenotype" that is domain general and temporally stable. *Nature Communications*, *5*(1), 1–8. https://doi.org/10.1038/ncomms5939

Poldrack, R. A., & Yarkoni, T. (2016). From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure. *Annual Review of Psychology*, *67*(1), 587–612. https://doi.org/10.1146/annurev-psych-122414-033729

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/S0140525X02000018

Radua, J., Mataix-Cols, D., Phillips, M. L., El-Hage, W., Kronhaus, D. M., Cardoner, N., & Surguladze, S. (2012). A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. *European Psychiatry*, *27*(8), 605–611. https://doi.org/10.1016/j.eurpsy.2011.04.001

Radua, J., Rubia, K., Canales-Rodríguez, E. J., Pomarol-Clotet, E., Fusar-Poli, P., & Mataix-Cols, D. (2014). Anisotropic kernels for coordinate-based meta-analyses of neuroimaging studies. *Frontiers in Psychiatry*, *5*(13), 1–8. https://doi.org/10.3389/fpsyt.2014.00013

Ramsøy, T. Z., Skov, M., Macoveanu, J., Siebner, H. R., & Fosgaard, T. R. (2015). Empathy as a neuropsychological heuristic in social decision-making. *Social Neuroscience*, *10*(2), 179–191. https://doi.org/10.1080/17470919.2014.965341

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Redcay, E., & Moraczewski, D. (2019). Social cognition in context: A naturalistic imaging approach. *NeuroImage*, *216*, 116392. https://doi.org/10.1016/j.neuroimage.2019.116392

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*, *20*(8), 496–505. https://doi.org/10.1038/s41583-019-0179-4

Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, *84*(5), 754. https://doi.org/10.1037//0021-9010.84.5.754

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A Neural Basis for Social Cooperation. *Neuron*, *35*(2), 395–405. https://ac.els-cdn.com/S0896627302007559/1-s2.0-S0896627302007559-main.pdf?_tid=b30ec185-cd6d-4ad2-acd6-02d8be7580af&acdnat=1551065251_9b20b10ab6977b1373b83aaf75eaa9ae

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*(4), 1694–1703. https://doi.org/10.1016/j.neuroimage.2004.04.015

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562. https://doi.org/10.1038/nrn3776

Rusbult, C. E., & Van Lange, P. A. M. (2003). Interdependence, Interaction, and Relationships. *Annual Review of Psychology*, *54*(1), 351–375. https://doi.org/10.1146/annurev.psych.54.101601.145059

Safin, V., Locey, M. L., & Rachlin, H. (2013). Valuing rewards to others in a prisoner's dilemma game. *Behavioural Processes*, *99*, 145–149. https://doi.org/10.1016/j.beproc.2013.07.008

Schneider-Hassloff, H., Straube, B., Nuscheler, B., Wemken, G., & Kircher, T. (2015). Adult attachment style modulates neural responses in a mentalizing task. *Neuroscience*, *303*, 462–473. https://doi.org/10.1016/j.neuroscience.2015.06.062

Schreuders, E., Klapwijk, E. T., Will, G. J., & Güroğlu, B. (2018). Friend versus foe: Neural correlates of prosocial decisions for liked and disliked peers. *Cognitive, Affective and Behavioral Neuroscience*, *18*(1), 127–142. https://doi.org/10.3758/s13415-017-0557-1

Schreuders, E., Smeekens, S., Cillessen, A. H. N., & Güroğlu, B. (2019). Friends and foes:

Neural correlates of prosocial decisions with peers in adolescence. *Neuropsychologia*, *129*, 153–163. https://doi.org/10.1016/j.neuropsychologia.2019.03.004

Sharp, C., Burton, P. C., & Ha, C. (2011). “Better the devil you know”: A preliminary study of the differential modulating effects of reputation on reward processing for boys with and without externalizing behavior problems. *European Child and Adolescent Psychiatry*, *20*(11–12), 581–592. https://doi.org/10.1007/s00787-011-0225-x

Shaw, D. J., Czekóová, K., Staněk, R., Mareček, R., Urbánek, T., Špalek, J., Kopečková, L., Řezáč, J., & Brázdil, M. (2018). A dual-fMRI investigation of the iterated Ultimatum Game reveals that reciprocal behaviour is associated with neural alignment. *Scientific Reports*, *8*(1), 1–13. https://doi.org/10.1038/s41598-018-29233-9

Smith-Collins, A. P. R., Fiorentini, C., Kessler, E., Boyd, H., Roberts, F., & Skuse, D. H. (2013). Specific neural correlates of successful learning and adaptation during social exchanges. *Social Cognitive and Affective Neuroscience*, *8*(8), 887–897. https://doi.org/10.1093/scan/nss079

Smith, D. V., Hayden, B. Y., Truong, T. K., Song, A. W., Platt, M. L., & Huettel, S. A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience*, *30*(7), 2490–2495. https://doi.org/10.1523/JNEUROSCI.3319-09.2010

Stanley, D. A., Sokol-Hessner, P., Fareri, D. S., Perino, M. T., Delgado, M. R., Banaji, M. R., & Phelps, E. A. (2012). Race and reputation: Perceived racial group trustworthiness influences the neural correlates of trust decisions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1589), 744–753. https://doi.org/10.1098/rstb.2011.0300

Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I. I., Tobler, P. N., & Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences*, *112*(5), 1619–1624. https://doi.org/10.1073/pnas.1414715112

Sul, S., Tobler, P. N., Hein, G., Leiberg, S., Jung, D., Fehr, E., & Kim, H. (2015). Spatial gradient in value representation along the medial prefrontal cortex reflects individual differences in prosociality. *Proceedings of the National Academy of Sciences*, *112*(25), 7851–7856. https://doi.org/10.1073/pnas.1423895112

Suzuki, S., & O'Doherty, J. P. (2020). Breaking human social decision making into multiple components and then putting them together again. *Cortex*, *127*, 221–230.

Tabibnia, G., & Lieberman, M. D. (2007). Fairness and cooperation are rewarding: Evidence from social cognitive neuroscience. *Annals of the New York Academy of Sciences*, *1118*, 90–101. https://doi.org/10.1196/annals.1412.001

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2007). The Sunny Side of Fairness: Preference for Fairness Activates Reward Circuitry (and Disregarding Unfairness Activates Self-Control Circuitry). *Psychological Science*, *19*(4), 339–347. https://doi.org/10.1111/j.1467-9280.2008.02091.x

Telzer, E. H., Ichien, N., & Qu, Y. (2015). The ties that bind: Group membership shapes the neural correlates of in-group favoritism. *NeuroImage*, *115*, 42–51. https://doi.org/10.1016/j.neuroimage.2015.04.035

Telzer, E. H., Masten, C. L., Berkman, E. T., Lieberman, M. D., & Fuligni, A. J. (2011). Neural regions associated with self control and mentalizing are recruited during prosocial behaviors towards the family. *NeuroImage*, *58*(1), 242–249. https://doi.org/10.1016/j.neuroimage.2011.06.013

The American Heritage Dictionary of the English Language (4th ed.). (2000). Houghton Mifflin

Harcourt.

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and Prosocial Behavior: A Theoretical Framework and Meta-Analysis. *Psychological Bulletin*, *146*(1), 30–90. https://doi.org/10.1037/bul0000217.supp

Tognoli, E., Dumas, G., & Kelso, J. A. S. (2017). A roadmap to computational social neuroscience. *Cognitive Neurodynamics*, *12*(1), 135–140. https://doi.org/10.1007/s11571-017-9462-0

Tricomi, E., & Sullivan-Toole, H. (2015). Fairness and Inequity Aversion. In *Brain Mapping: An Encyclopedic Reference* (Vol. 3). Elsevier Inc. https://doi.org/10.1016/B978-0-12-397025-1.00142-1

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, *16*(3), 765–780. https://doi.org/10.1006/nimg.2002.1131

Turner, J. A., & Laird, A. R. (2012). The cognitive paradigm ontology: Design and application. *Neuroinformatics*, *10*(1), 57–66. https://doi.org/10.1007/s12021-011-9126-x

Tusche, A., Bockler, A., Kanske, P., Trautwein, F.-M., & Singer, T. (2016). Decoding the Charitable Brain: Empathy, Perspective Taking, and Attention Shifts Differentially Predict Altruistic Giving. *Journal of Neuroscience*, *36*(17), 4719–4732. https://doi.org/10.1523/JNEUROSCI.3392-15.2016

Tusche, A., & Hutcherson, C. A. (2018). Cognitive regulation alters social and dietary choice by changing attribute representations in domain-general and domain-specific brain circuits. *ELife*, *7*, e31185. https://doi.org/10.7554/eLife.31185

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A. R. B., & Crone, E. A. (2009).

What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Social Cognitive and Affective Neuroscience*, *4*(3), 294–304. https://doi.org/10.1093/scan/nsp009

Vekaria, K. M., O'Connell, K., Rhoads, S. A., Brethel-Haurwitz, K. M., Cardinale, E. M., Robertson, E. L., Walitt, B., VanMeter, J. W., & Marsh, A. A. (2020). Activation in bed nucleus of the stria terminalis (BNST) corresponds to everyday helping. *Cortex*, *127*, 67–77. https://doi.org/10.1016/j.cortex.2020.02.001

Watanabe, T., Takezawa, M., Nakawake, Y., Kunimatsu, A., Yamasue, H., Nakamura, M., Miyashita, Y., & Masuda, N. (2014). Two distinct neural mechanisms underlying indirect reciprocity. *Proceedings of the National Academy of Sciences*, *111*(11), 3990–3995. https://doi.org/10.1073/pnas.1318570111

Weiland, S., Hewig, J., Hecht, H., Mussel, P., & Miltner, W. H. R. (2012). Neural correlates of fair behavior in interpersonal bargaining. *Social Neuroscience*, *7*(5), 537–551. https://doi.org/10.1080/17470919.2012.674056

Wheatley, T., Boncz, A., Toni, I., & Stolk, A. (2019). Beyond the Isolated Brain: The Promise and Challenge of Interacting Minds. *Neuron*, *103*(2), 186–188. https://doi.org/10.1016/j.neuron.2019.05.009

Will, G. J., Crone, E. A., Van Lier, P. A. C., & Güroğlu, B. (2016). Neural correlates of retaliatory and prosocial reactions to social exclusion: Associations with chronic peer rejection. *Developmental Cognitive Neuroscience*, *19*, 288–297. https://doi.org/10.1016/j.dcn.2016.05.004

Wills, J. A., Hackel, L., FeldmanHall, O., Parnamets, P., & Van Bavel, J. J. (2018). The Social Neuroscience of Cooperation. *PsyArXiv*, 1–31. https://doi.org/10.31234/osf.io/8kcd4

Wills, J. A., Hackel, L., & Van Bavel, J. (2018). Shifting prosocial intuitions: Neurocognitive evidence for a value based account of group-based cooperation. *PsyArXiv*.

Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. S. (2016). Self-Other Mergence in the Frontal Cortex during Cooperation and Competition. *Neuron*, *91*(2), 482–493. https://doi.org/10.1016/j.neuron.2016.06.022

Wu, H., Luo, Y., & Feng, C. (2016). Neural signatures of social conformity: A coordinate-based activation likelihood estimation meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, *71*, 101–111. https://doi.org/10.1016/j.neubiorev.2016.08.038

Yang, Z., Zheng, Y., Yang, G., Li, Q., & Liu, X. (2019). Neural signatures of cooperation enforcement and violation: a coordinate-based meta-analysis. *Social Cognitive and Affective Neuroscience*, *14*(9), 919–931. https://doi.org/10.1093/scan/nsz073

Zaki, J., & Mitchell, J. P. (2011). Equitable decision making is associated with neural markers of intrinsic value. *Proceedings of the National Academy of Sciences*, *108*(49), 19761–19766. https://doi.org/10.1073/pnas.1112324108

Zaki, J., & Mitchell, J. P. (2013). Intuitive Prosociality. *Current Directions in Psychological Science*, *22*(6), 466–470. https://doi.org/10.1177/0963721413492764

Zinchenko, O., & Arsalidou, M. (2018). Brain responses to social norms: Meta-analyses of fMRI studies. *Human Brain Mapping*, *39*(2), 955–970. https://doi.org/10.1002/hbm.23895