

**OVERVIEW**

# Neurocomputational models of altruistic decision-making and social motives: Advances, pitfalls, and future directions

Anita Tusche<sup>1,2,3</sup>  | Lisa M. Bas<sup>1</sup> <sup>1</sup>Department of Psychology, Queen's University, Ontario, Kingston, Canada<sup>2</sup>Department of Economics, Queen's University, Ontario, Kingston, Canada<sup>3</sup>Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California, USA**Correspondence**

Anita Tusche, Department of Psychology, Queen's University, 62 Arch Street, Humphrey Hall, ON K7L 3N6, Canada.  
Email: anita.tusche@gmail.com

**Funding information**

National Institute of Mental Health Conte Center, Grant/Award Number: 2P50 MH094258; Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2019-04329; Canada Foundation for Innovation, John R. Evans Leaders Fund: CFI JELF, Grant/Award Number: 38812

**Edited by:** Joe Kable, Editor**Abstract**

This article discusses insights from computational models and social neuroscience into motivations, precursors, and mechanisms of altruistic decision-making and other-regard. We introduce theoretical and methodological tools for researchers who wish to adopt a multilevel, computational approach to study behaviors that promote others' welfare. Using examples from recent studies, we outline multiple mental and neural processes relevant to altruism. To this end, we integrate evidence from neuroimaging, psychology, economics, and formalized mathematical models. We introduce basic mechanisms—pertinent to a broad range of value-based decisions—and social emotions and cognitions commonly recruited when our decisions involve other people. Regarding the latter, we discuss how decomposing distinct facets of social processes can advance altruistic models and the development of novel, targeted interventions. We propose that an accelerated synthesis of computational approaches and social neuroscience represents a critical step towards a more comprehensive understanding of altruistic decision-making. We discuss the utility of this approach to study lifespan differences in social preference in late adulthood, a crucial future direction in aging global populations. Finally, we review potential pitfalls and recommendations for researchers interested in applying a computational approach to their research.

This article is categorized under:

Economics > Interactive Decision-Making

Psychology > Emotion and Motivation

Neuroscience > Cognition

Economics > Individual Decision-Making

**KEYWORDS**

decision neuroscience, drift diffusion models, prosociality, social affect and cognition (theory of mind), social choice tasks

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *WIREs Cognitive Science* published by Wiley Periodicals LLC.

## 1 | INTRODUCTION

Human beings are inherently social, and a considerable part of our thoughts, decisions, and behaviors concerns the people around us. One social phenomenon that has received significant attention is human altruism (Declerck & Boone, 2015; Dovidio et al., 2017; Fehr & Rockenbach, 2004; Filkowski et al., 2016; Luo, 2018). Altruism describes a motivational state to promote someone else's welfare, even at a risk or cost to ourselves (Batson & Powell, 2003). It can be distinguished from prosocial behaviors driven by other motivations (e.g., strategic considerations or social norm compliance) (Böckler et al., 2016). Contrary to assumptions of classic economic theory (Neumann & Morgenstern, 1947), people are frequently willing to forgo personal gains to benefit others. Interestingly, this willingness to care about others' welfare and to act altruistically differs across people. Recent studies show that individual differences in altruistically motivated behaviors are stable over time, generalize across specific measurement tools (Böckler et al., 2016; Böckler, Tusche, Schmidt, & Singer, 2018; Peysakhovich et al., 2014; Yamagishi et al., 2013), and are linked to well-being (Hui et al., 2020; Le et al., 2018; Nelson et al., 2016). The wealth of research on this topic is hardly surprising. Societies depend on the altruistic behaviors of their members. There is a general consensus that caring about others' welfare—and acting on this concern—is central to human decision-making, successful social interactions, and society's functioning at large. Yet, to this day, altruism presents a fascinating puzzle as it challenges assumptions of narrow self-interest that are central to rational choice theory. What mechanisms drive decisions to bear substantial costs to benefit another? How do characteristics of the individual and choice setting shape other-regard and the altruistic decision process?

Decades of empirical research have advanced our ability to answer these questions. Studies on social decision-making have uncovered basic mechanisms of social preferences. Prior work has shed light on the role of personality (Thielmann et al., 2020; Zhao & Smillie, 2015), social context (Bruch & Feinberg, 2017; Rand et al., 2015), social emotion and cognition on the decision process (Barasch et al., 2014; Batson, 2011<sup>1</sup>; Lerner et al., 2015). This progress in our understanding is largely due to the interdisciplinary and computational nature of contemporary research on this topic. The mutual reinforcement of fast-growing fields such as social neuroscience (Adolphs, 2010; Lieberman, 2007) and neuroeconomics (Konovalov & Krajbich, 2019; Reuter & Montag, 2016) has provided novel insights into motivations, mental and neural computations, and contextual factors that guide altruistic behaviors. This article will review theoretical and empirical models of other-regard in altruistic decision-making through the lens of a cross-disciplinary, computational framework. Our goal is to promote the following notion: an interdisciplinary, computational framework provides a more foundational, more general explanation than economic, neurobiological, or psychological measures alone. It allows making broader and more accurate predictions, and ultimately, to formulate more precise interventions through which social decision-making can be altered. We advocate the benefits of such a framework to researchers from diverse fields that study social decision-making and social motives and do not yet apply computational modeling to their research questions. While certain research lines within neuroeconomics and social neuroscience have started to embrace computational approaches, there is still tremendous potential for growth (e.g., advancing computational models of affective responses in social decision-making). Our goal is to outline the benefits, pitfalls, and future directions of interdisciplinary, computational approaches to study altruism and social motives. Recent reviews address related points in the context of computational strategies involved in social learning and social behavior, focusing on reinforcement learning (RL) (Charpentier & O'Doherty, 2018; Lockwood & Klein-Flügge, 2020). Unlike prior work, we will focus on the computational perspective on the narrow concept of altruism, specifically the processing of and caring about another's welfare. However, the basic principles of the benefits of such an interdisciplinary, computational framework extend to related concepts such as cooperation, reciprocity, and altruistic punishment (for definition and review of neural substrates, see Filkowski et al., 2016; Reuter & Montag, 2016; Rilling & Sanfey, 2011).

The structure of the article is as follows. In Section 2, we introduce the interdisciplinary, computational framework of contemporary research on altruism. To this end, we will briefly review popular experimental measures that permit the use of formal mathematical models of social preferences. It provides background information that we will refer to in subsequent parts of the article. To illustrate the utility of a computational framework, Section 3 highlights some specific insights from a recent neurocomputational model of altruism (focusing on popular drift diffusion models [DDMs]). Section 4 introduces alternative computational models of altruism and social motives (focusing on RL). Section 5 discusses processes that guide altruistic decision-making and their substrates in the brain. We will review these processes in light of whether altruistic decisions are different from other kinds of value-based decisions. The focus will be on affective and cognitive social processes relevant to understanding others (empathy and mentalizing). Section 6 outlines promising future directions of the neurocomputational framework of altruistic decision-making. We highlight open

questions regarding lifespan changes in altruism and the integration of specific components of the decision process. Finally, Section 7 discusses some methodological and practical challenges of an interdisciplinary, computational framework and offers recommendations.

## 2 | PROMISES OF AN INTERDISCIPLINARY, COMPUTATIONAL FRAMEWORK

### 2.1 | Synthesizing theories, measures, and analysis tools from different research disciplines

Studies on human altruism have benefitted from the successful cross-fertilization of social psychology, neuroscience, and behavioral economics. Traditionally, these research disciplines focus on different facets of social behaviors and levels of description. Simply put, these research disciplines differ regarding *how* and *what* aspects of social decisions they examine. For instance, neuroimaging studies have concentrated on identifying brain networks reliably recruited during altruistic and strategic social decision-making (Cutler & Campbell-Meiklejohn, 2019). Seminal psychological research has extensively studied emotional, cognitive, and contextual factors that can motivate altruistic behaviors, such as empathy or perceiving others in need or distress (Batson, 2011). Finally, economics traditionally focuses on observed behavior to reveal other-regard (“revealed preferences”), with less emphasis on the underlying mental processes. Economic game theory has proven a powerful theoretical framework to describe and predict these observed social behaviors (Camerer, 2011; van Dijk & De Dreu, 2020).<sup>2</sup> It provides a common and parsimonious model applicable to a broad range of social behaviors and research questions (for an introduction to game theory and areas of applications, see Samuelson, 2016). Contemporary research on human altruism takes advantage of the multilevel perspectives from these different disciplines. Bridging the gap between distinct research traditions has enabled the field to utilize their respective strengths, insights, and explanatory power.

In many ways, these advances were facilitated by the widespread use of measures and analysis tools from neighboring disciplines. Research on human altruism has been the focus of many research fields that traditionally employ different assessment tools. For example, psychology regularly uses questionnaires that measure people's general prosocial tendencies (“traits”) (Rushton et al., 1981) or assess real-world prosocial behaviors like spontaneous helping, volunteering, or donations (Bethlehem et al., 2017; Gaesser et al., 2020; Smith, 1981). Donations have also been popular in early neuroimaging studies (Genevsky et al., 2013; Hare et al., 2010; Ma et al., 2011; Tusche et al., 2016). More recently, crowdfunding has been an exciting way to study ecologically valid altruism (André et al., 2017; Bretschneider & Leimeister, 2017; Genevsky et al., 2017; Giudici et al., 2018; Ryu et al., 2020). Other measures that require hypothetical distributions of resources have their roots in social psychology and economics. Choices between a preselected set of fixed distributions allow characterizing individuals' social value orientation (McClintock & Van Avermaet, 1982; Murphy et al., 2011; Van Lange, 1999). For instance, choices may reveal an individual's preference for maximization of her own payoffs, relative gains over that of another (winning), or joint payoffs (collective efficiency or “social welfare”). The examples illustrate the long tradition and the broad range of measures used to study altruistic motivations and underlying mental and neural processes.

Applications of each assessment tool have yielded important theoretical and empirical insights. However, the acceleration of cross-disciplinary, computational research on altruism over the last decades is in large part due to the widespread use of tasks from (behavioral) game theory (henceforth “games”). Games are traditionally used in behavioral and experimental economics but have also been successfully utilized in other fields like psychology, sociology, and biology (McNamara & Leimar, 2010; Murnighan & Wang, 2016; Pruitt & Kimmel, 1977). Games model social behavior in interdependent social settings. Numerous games have been developed to model behavior in different classes of social situations that people encounter in the real world (Kelley et al., 2003). There are many reasons for the popularity of this measurement tool. Games are standardized experimental tasks, usually posed in abstract terms (stripped of the contextual richness of real-world decision problems), require participants to make simple choices concerning resource distribution, specify exactly how choices relate to payoffs and outcomes, and resulting payoffs are usually realized right away (“incentivized” choices) (Camerer, 2011; Pruitt & Kimmel, 1977). These paradigms are easy to implement, easy to understand, allow for large numbers of repetitions, and specify precise variables. For these reasons, games have been frequently used in neuroimaging studies on altruism (Harbaugh et al., 2007; Moll et al., 2006; Morishima et al., 2012). Experimental games have yielded crucial insights into the neural and psychological processes underlying social decision-making and other-regard (for a recent overview, see van Dijk & De Dreu, 2020).

## 2.2 | Experimental games and formal models of altruistic decision-making

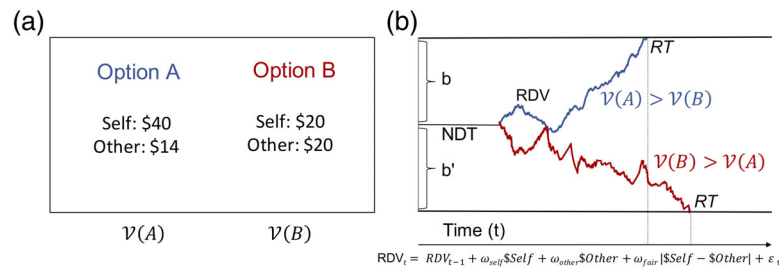
One significant advantage of game-theoretical paradigms is that they permit the use of formal models to study social decision-making. Formal mathematical models can describe *how* specific considerations (variables) interact and are integrated to yield observed behaviors (Crockett, 2016). The widespread use of mathematical models and computational analysis tools is likely the biggest innovation in the field of social decision-making—including research on altruism—over the last two decades. There is some debate that spans across disciplines about what constitutes (and distinguishes) formal, mathematical, and computational models (Diederich & Busemeyer, 2012). We define computational models as a mathematical model that relies on computer simulation to study the workings of a complex system (Roberts & Hutcherson, 2019). For reasons of simplification, we refer to formal models as an umbrella term. Formal models have been used to study numerous psychological and neural processes that can guide prosocial behaviors. For example, researchers have gained important insights into the processes underlying impression formation (e.g., “deserving of help?”), social learning (e.g., “will they reciprocate?”), or biases due to group memberships (e.g., “us” vs. “them”) (for a recent synopsis, see Hackel & Amodio, 2018).

To illustrate how the combination of formal models and standardized experimental games have invigorated studies on altruistic decision-making, we next turn to the dictator game. The dictator game is a canonic game-theoretical paradigm frequently used to study altruistic decision-making in the laboratory (Engel, 2011; Pisor et al., 2020) (for an overview of the origins of this game, see Guala & Mittone, 2010). Its classic form requires participants to unilaterally divide a fixed amount of money between themselves and another person. Payoffs are determined solely by the decision-maker (allocator or “dictator”), and the recipient must accept the proposed split. Choices are usually incentivized, and allocators and recipients receive the payoffs immediately after the experiment. In anonymous settings with no chance of future interactions (one-shot game), strategic concerns do not have to enter the decision process. Allocating nothing to the recipient maximizes the decision-maker’s income. Nevertheless, a considerable fraction of individuals violates the income maximization hypothesis of rational choice theory (Scott, 2000) and allocate something to the other person. On average, allocators are willing to give up 28% of the “pie” (Engel, 2011). Various factors related to the individual and the choice setting have been identified that systematically moderate generosity levels in this game (Engel, 2011; Henrich et al., 2005; Larney et al., 2019). Behavior in this task is widely assumed to reflect the value that an individual places on others’ well-being (altruism) and the willingness to conform to social fairness norms (Guala & Mittone, 2010). Recent findings also suggest that behavior in this task can indicate people’s more general propensity to engage in altruistically motivated behaviors (Böckler et al., 2016; Böckler, Tusche, Schmidt, et al., 2018; Böckler, Tusche, & Singer, 2018).

Since its conception (Forsythe et al., 1994; Kahneman et al., 1986), various modified versions of the dictator game have been put forward. Neuroimaging studies on social preferences require many repetitions (trials) to reliably assess noisy measures of brain function (Huettel et al., 2004). Applications of computational models also require numerous observations and trials (Section 7 discusses potential challenges related to this issue). Adaptations of the dictator game are well suited to address this need. Participants may be asked to repeatedly accept (or reject) a proposed monetary allocation that affects payoffs for themselves and another person (group/charity). Alternatively, participants may choose between two offers that either benefit the other at the expense of oneself (altruistic) or oneself at the cost of another (selfish) (Figure 1a) (Hein et al., 2016; Hutcherson et al., 2015; Shuster & Levy, 2020; Tusche & Hutcherson, 2018). Across versions, the dictator game captures critical features of economic reasoning that guide social behaviors in real life: in altruistic choice settings, we often have to trade off others’ welfare against personal benefits. Presented with the option to help someone or not, we might consider factors such as our gain, the benefit for another individual (group), or the fairness of the outcome (e.g., “How much better/worse am I off than the other(s)?”). These choice-relevant considerations are captured in the dictator game in the form of specific variables, each expressed in numeric values ( $\$Self$ ,  $\$Other$ , and  $|\$Self-\$Other|$ , see Figure 1a). This feature of the dictator game enables the use of formal mathematical models.

## 2.3 | Utility of formal models of altruism

A formal mathematical framework provides several important advantages to study social decision-making and other-regard (for excellent overviews, see Crockett, 2016; Hackel & Amodio, 2018; Roberts & Hutcherson, 2019). First, formal models enable making precise quantitative predictions of decisions across people, contexts, and time. In other words, a



**FIGURE 1** Game-theoretical paradigm and computational model of altruistic choice. (a) Altruistic choice in a modified dictator game (“choose the option that you prefer”). (b) Computational model of altruism (multi-attribute extension of a drift diffusion model). The model characterizes the decision process as the accumulation of a noisy value signal that evolves dynamically over time ( $t$ ). The value signal represents the relative desirability of available choice options (relative decision value, RDV). Blue and red lines represent trajectories of the value signal in favor of Option A or B, respectively. The value signal results from a linear combination of weighted features of the decision problem (e.g., gains for self, others, and fairness; represented as numerical values in each trial:  $\$Self$ ,  $\$Other$ ,  $|\$Self - \$Other|$ ). A decision is reached when the value signal crosses a critical threshold (upper or lower barrier,  $b$  or  $b'$ ). Fitting the model to the observed behavior (choices and reaction times) yields parameters that characterize the decision process. Potential parameters include the weights that individuals place on each choice-relevant feature ( $w_{Self}$ ,  $w_{Other}$ ,  $w_{Fair}$ ); non-decision time (NDT; accounts for sensory and motor-related processes unrelated to the value comparison process itself); the decision threshold; or the start point of the value signal (accounts for initial biases in favor of one available alternative). Drift diffusion models of altruism have been shown to capture generous choices, reaction times (Krajbich et al., 2015), and even neural responses in altruism tasks (Hutcherson et al., 2015; Tusche & Hutcherson, 2018)

well-trained model can produce specific and accurate predictions about social behaviors for individuals (groups) and settings other than those used to estimate the model in the first place (out-of-sample predictions). This feature has obvious value for numerous applied settings. Second, formal computational models allow comparing observed behavior—choices and reaction times—to optimality (Bogacz et al., 2006). We use optimality here to refer to decisions that would maximize certain social preferences (e.g., others welfare,  $\max(\$Other)$ ; own gains,  $\max(\$Self)$ ; or fairness,  $\min(|\$Self - \$Other|)$ ). This application has yielded important insights into altruistic decision-making. For instance, neurocomputational models of altruism suggested that a significant portion of generous choices may represent mistakes rather than individuals' actual altruistic preferences (Hutcherson et al., 2015). Third, by fitting a formal model to observed behavior, researchers can estimate “hidden” components of the altruistic decision process that might not be easily inferred from behavioral observation alone. In other words, models allow decomposing observed behavioral data into several latent mental processes (Forstmann et al., 2011). Here, in addition to actual latent variable inferences, estimates of the free parameters of a model play a crucial role. In a nutshell, these estimates represent a set of parameter values that best account for the real data (e.g., choice, gaze patterns, or brain data) for a given model (Wilson & Collins, 2019).<sup>3</sup> For instance, the estimates of a parameter can quantify the relative weight that individuals place on distinct considerations that guide decisions (e.g., “other welfare” or “fairness”) (Figure 1b). Models thus enable researchers to capture the precise utilities that people ascribe to others' welfare or equity concerns. These parameter estimates can provide important insights. Imagine observing a generous choice in a standardized experimental task (e.g., a modified dictator game). The observed behavior (generous) may have resulted from different motives: the allocator might care *little* about their own payoffs in this low stake choice setting, care *heavily* about others welfare, care *heavily* about maximizing the fairness of the outcome (which by chance is realized by the generous choice option), or a combination of these components. People's insights into their underlying motives and their relative input into the decision process may be limited. Estimated parameters of a model can discriminate between these interpretations (e.g., by examining the weights of self-regarding motives [ $\$Self$ ], others gain [ $\$Other$ ], or fairness [ $|\$Self - \$Other|$ ] on choices that best describe the observed behavior). They thus provide an alternative way to study social motives and complement classic approaches of simply asking people for their preferences and motivations (Insko & Schopler, 2013). Fourth, these estimated parameters of computational models can be linked to functional and structural brain measures. This enables researchers to reveal the neural underpinnings of “hidden” (latent) factors of altruistic decision-making. A neurally informed model of altruistic choice, in turn, enables researchers to predict how changes in the brain will affect observable behaviors. Fifth, neurocomputational models allow testing for common mechanisms of the decision process across choice domains (both on the behavioral and neural levels). Recent studies have started to identify unifying principles that generalize across different decision problems (Krajbich et al., 2015; Soutschek et al., 2016; Tusche &

Hutcherson, 2018). For example, Krajbich et al. (2015) showed that a formal model of dietary decision-making accurately predicted choices and reaction times in a social choice task (dictator game). This evidence points towards a common mechanism that underlies different types of decisions. Designing novel interventions that target common principles may hold solutions to effectively improve behavior across a range of domains in which people frequently struggle to align decisions with “virtuous” goals. For instance, shared mechanisms underlying food and social decisions may be targeted to support goals related to eating healthier and being more generous. In that regard, computational models can have important implications for policy and practitioners. We will further discuss the notion of common decision mechanisms across social and dietary choices—and their neural substrates—in Section 3.2.

## 2.4 | Modeling social preferences: DDMs of altruistic choice

In essence, models force researchers to formalize their attempts to describe the altruistic decision process. Over the last decades, several mathematical models have been developed to characterize how social motives drive behaviors. Popular models of social preferences capture people’s utility for fairness (also referred to as inequality aversion) (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999; Morishima et al., 2012), collective gains (“social welfare”) (Charness & Rabin, 2002), spitefulness and altruism (Levine, 1998), the consistency of altruistic preferences (Andreoni & Miller, 2002), and the impact of social identity on social motives (Akerlof & Kranton, 2005; Bénabou & Tirole, 2011; Chen & Li, 2009).

Here, we will discuss one popular class of models that inspired recent neurocomputational models of altruism: drift diffusion models (DDMs) (Bogacz et al., 2006; Ratcliff & McKoon, 2008; Ratcliff et al., 2016). Diffusion models represent one example from the broader family of sequential-sampling models (Forstmann et al., 2016; Ratcliff et al., 2016). They have a long history of predicting simple perceptual decisions and account for reaction times and brain activation (Gold & Shadlen, 2003; Heekeren et al., 2008). Only recently have these models found their way into the domain of social decision-making (Roberts & Hutcherson, 2019). Diffusion models provide a window into the underlying decision process. In a nutshell, the model assumes that decisions result from the noisy moment-by-moment accumulation of a value signal (Figure 1b). You may think of the value signal as a reflection of the relative “attractiveness” of the available choice options. In our example (Figure 1b), it is based on a weighted sum of three choice-relevant considerations (variables): payoffs for oneself, others, and the outcome’s fairness (Hutcherson et al., 2015; Tusche & Hutcherson, 2018). The model assumes that the decision-maker (and her brain) continuously accumulates noisy information about the choice options until an internal decision boundary is reached. In other words, when enough evidence has accumulated in favor of one option (and the value signal reaches the respective threshold), a decision in favor of this option is made. Thresholds in this model are subjective, capturing the notion that people can adjust their decision criteria (higher decision thresholds resulting in slower but more accurate choices; lower thresholds can be reached faster by the accumulating value signal but might be more error-prone). This computational model has been shown to capture observed patterns of choices, reaction times, and brain activation during altruistic decisions (Hutcherson et al., 2015), all within one powerful analytical framework.

## 3 | INSIGHTS FROM A NEUROCOMPUTATIONAL MODEL OF ALTRUISTIC CHOICE

So far, we have discussed that game-theoretical paradigms permit the use of formal models of social preferences, advantages of a computational framework, and a recent (neuro)computational model of altruism (DDM). Let us bring these components together by way of example. For illustration, we turn to a recent neuroimaging study on altruism (Tusche & Hutcherson, 2018). In this study, participants performed an adapted version of a dictator game while their brain responses were measured using functional magnetic resonance imaging (fMRI). In each trial, participants choose between a proposal that affected their own monetary outcome and that of an anonymous partner, and a default of \$20 for both. Compared to the default allocation, accepting or rejecting the on-screen proposal could benefit the participant or the partner, yielding selfish or generous choices. The researchers assumed that participants mostly considered three factors (variables) to guide their decisions: payoffs for oneself (\$Self), payoffs for the other (\$Other), and the fairness of the outcome ( $|\$Self - \$Other|$ ). Each of these variables was expressed in terms of a specific numeric value in each trial. As discussed earlier, formal models provide a precise, quantitative description of *how* these considerations are

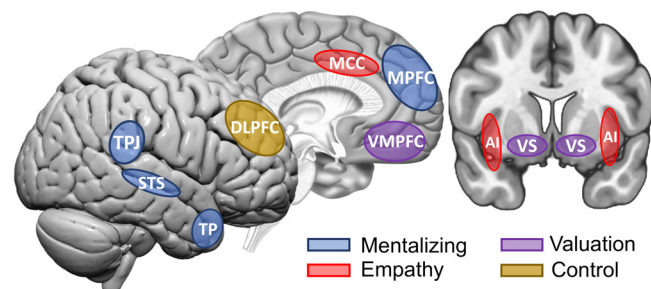
transformed into observed behaviors (generous or selfish choices). To this end, the researchers fitted a multi-attribute extension of the DDM to observed behavior—choices and reaction times (Ratcliff & McKoon, 2008; Ratcliff et al., 2016; Ratcliff & Smith, 2004). The model captured choices, reaction times, and neural data during altruistic choice with high accuracy, mirroring previous evidence (Hutcherson et al., 2015). More importantly, this neurocomputational model of altruism provided novel insights, some of which we will highlight in the following.

### 3.1 | Goal-dependent changes in altruistic choice

People might behave generously in one situation but not in others. Computational modeling approaches offer insights into the mechanisms underlying the variance in altruistic behaviors across contexts. Let us consider evidence from the neuroimaging study introduced above (Tusche & Hutcherson, 2018). Participants in this study made altruistic decisions under three conditions. They were asked to think about the impact of their choice on their partner (directing attention to other's welfare), or the ethics of their choice (directing attention to the fairness of the outcome), or to choose as they naturally would (baseline condition). Not surprisingly, observed generosity levels differed across choice settings. People behaved more altruistically when prompted to deliberate consequences for another individual or social norms. The computational model sheds light on the precise *mechanisms* underlying altered generosity. It yielded estimates representing the weight that individuals placed on each variable (e.g., others gain) in a particular choice setting. The researchers found that the instructed goals systematically altered these weights: contexts that directed attention to a certain choice feature, say others gain, yielded *increased* weight for this *goal-consistent* consideration, capturing its increased input on the altruistic decision process. On the contrary, the weight of *goal-inconsistent* features on choices *decreased* (e.g., one's own gain). Neuroimaging data provide further support for this computational model of altruistic choice. The study found that the brain encodes relevant choice features (gains for self, others, and fairness) and integrated decision values. Notably, neural information on these choice features varied as a function of contextual goals. These changes in neural information closely matched the predictions of the DDM. Taken together, the neurocomputational model explains how people flexibly align social behaviors with current goals and how it is implemented in the brain. It provides us with a formal algorithmic model (Love, 2015) of goal-dependent choice in altruism. More generally, it explains why people can be extremely generous and cooperative in some contexts but not in others.

### 3.2 | Unifying mechanisms across decision problems

Formal models also allow comparing the decision process across different domains. Seminal empirical evidence suggests that there are unifying principles that generalize across various decision problems (Krajbich et al., 2015). For illustration, let us go back to the neuroimaging study by Tusche and Hutcherson (2018). The study probed for commonalities (and differences) in two domains: altruistic and dietary choice. Both choice tasks involved the experimental manipulation of contextual goals. Similar to the altruism task (see above), participants completed a food choice task under different attentional goals (“focus on taste,” “focus on health”). A multi-attribute extension of a DDM was again fitted to observed behavior, separately for each choice context. The results demonstrate a unifying principle underlying goal-consistent behavior in both domains. Successful regulators *increased* the input of goal-consistent considerations on choice (e.g., a food's healthiness when pursuing health goals) and *decreased* the contributions of goal-inconsistent features (e.g., taste). Interestingly, individuals differed in the ability to recruit this mechanism. Individuals who effectively aligned the inputs of altruistic variables with current goals were also better at recruiting this mechanism during dietary choices. The study also linked this unifying mechanism to activation patterns in the dorsolateral prefrontal cortex (DLPFC) (Figure 2). These findings reveal basic mechanisms of regulatory success that can impact altruistic decision-making, as well as other choice domains. It provides us with a neurally informed, mechanistic understanding of why some people seem to have an easier time aligning behaviors with their goals (e.g., “be more altruistic,” “eat healthier”). Notably, the model also identified components of the decision process that were specific to altruistic choice. For example, goal-consistent changes in concern for others gains recruited the temporoparietal junction (TPJ) (Figure 2). This brain area is widely believed to play a role in understanding others. Taken together, these findings point to generic mechanisms in altruism that generalize to other domains, as well as processes specifically relevant to altruism and social choice settings (but not dietary choice). Neural evidence and formal models can inform each other to tease these components apart and inform us about their interactions.



**FIGURE 2** Brain regions involved in (pro)social decision-making. Process-specific networks are color-coded for illustrative purposes. The schematic assignment of brain areas to mental functions represents a simplified account of popular assumptions about their functional role in (pro)social decision-making. We acknowledge that these brain areas have also been linked to other mental processes. AI, anterior insula; DLPFC, dorsolateral prefrontal cortex; MCC, mid cingulate cortex; MPFC, medial prefrontal cortex; STS, superior temporal sulcus; TP, temporal pole; TPJ, temporo-parietal junction; VS, ventral striatum; VMPFC, ventromedial prefrontal cortex

#### 4 | OTHER COMPUTATIONAL MODELS OF SOCIAL DECISIONS: RL

We outlined one popular computational model (DDM) and its utility to study altruistic behavior observed in the dictator game. The choice of a formal model—and the behavioral measures to which the model is fitted—will heavily depend on the specific research question. Summarizing the variety of computational approaches and paradigms to study social decision-making exceeds the scope of this paper (for a review relevant to social cognition, see Rusch et al., 2020). To exemplify the range of applications of formal models of social choice, we briefly highlight one topic that has received a lot of attention: social learning (for reviews, see Charpentier & O’Doherty, 2018; Lockwood & Klein-Flügge, 2020; Olsson et al., 2020); Bolenz et al. (2017) review lifespan differences in social learning and decision-making).

Many social choice tasks—including games—are designed to minimize social learning. Let us consider the example of the dictator game. Participants are fully aware of the rewards for themselves and others associated with their decisions. The number of choices (trials) is often limited, and people usually interact with a “new” partner in every round. Moreover, choices are often anonymous, with little or no information about the other person. These features distinguish behavior observed in these games from real-world decisions in interactive, repeated, and dynamic settings. Outside of the laboratory, people often make an initial assessment based on the limited information available and then adjust it based on subsequent interactions or observations. In other words, people *learn*. Through trial-and-error, they update their beliefs and expectations about others (e.g., their preferences, motives, or moral character; Hackel & Amodio, 2018) or social rewards linked with choices. Studying the impact of these processes on social preferences requires selecting appropriate computational models.

RL represents the most popular family of models to capture learning and adaptive decision-making. The RL framework can be used in non-social contexts and social settings that involve interactions with other individuals. In a nutshell, RL models provide a window into *how* people learn from feedback in repeated interactions or observations to make decisions. Paradigms used to study this process often introduce a mismatch of expected and actual outcomes. Peoples’ attempts to minimize this mismatch in following interactions and decisions represent people’s learning. We do not intend to provide a thorough overview of RL or formal mathematical notations. A comprehensive introduction to the theoretical framework, models, and neural underpinning can be found in Joiner et al. (2017), Lee and Seo (2016), and Olsson et al. (2020). An in-depth discussion is provided by Sutton and Barto (2018). Neural underpinnings of social learning related to social cognition (see Section 5) and decision-making are outlined in Olsson et al. (2020) (also see Cheong et al., 2017). Researchers interested in actually using RL models find guidance and best practices in Lockwood and Klein-Flügge (2020) and Zhang et al. (2020). In the following, we will highlight applications of RL in research on altruism and social motives.

The RL framework has been used in studies on altruism, social motives, social rewards, and their neural correlates (e.g., Fareri et al., 2015; Kuss et al., 2013; Kwak & Huettel, 2016; Rosenthal et al., 2019; Vanyukov et al., 2019). The relevance of social learning in altruism is obvious. To support others, we need to understand how our decisions affect the people around us. Interestingly, people vary in the degree to which they learn about the benefit (or harm) of their choices for themselves and others (Kwak et al., 2014; Lockwood et al., 2016; Sul et al., 2015). Individuals with better



learning about rewards for others report more engagement in real-world altruistic behaviors (Kwak et al., 2014). This result links biases in social learning with self-reported social preferences outside of the laboratory. RL can also help us understand how people update their expectations about generosity and effects on giving in the dictator game (Pereda et al., 2017). Moreover, RL models have provided insights into neurocomputational underpinnings of other-regard and peoples' propensity for altruistic behavior (e.g. Kuss et al., 2013; Sul et al., 2015). In sum, modeling traditions like RL shed light on how social learning shapes behavior towards the people around us.

Interestingly, RL can be combined with sequential sampling models like DDMs (for a review of theoretical grounding and mutual benefits, see Miletic et al., 2020; for a tutorial and a recent software solution, see Pedersen & Frank, 2020). Integrating traditionally separate modeling traditions into a unified framework provides an exciting future direction for computational approaches to study altruism and social motives. For instance, it might provide a mechanistic account of how the processes driving decision making (e.g., processing speed or response caution captured in DDMs) are adjusted during social learning (as captured in RL models).

## 5 | SOCIAL PROCESSES IN A VALUE-BASED FRAMEWORK OF ALTRUISTIC CHOICE

### 5.1 | Multiple components of altruistic decision-making

In Section 3, we examined recent evidence on generic and domain-specific components of altruistic decision-making through the lens of neurocomputational models. These results shine a light on a fundamental question: how “special” is altruistic decision-making? The extent to which dedicated mental processes and brain systems guide social decisions is heavily debated (Lockwood et al., 2020). The alternative notion of common processes subserving value-based decision-making more generally is intuitive. In complex and constantly changing environments, we are faced with countless decision settings. Specialized mental and neural computations for all possible types of scenarios seem inefficient. Consistent with this notion, there is a wide consensus that basic mechanisms of value-based decision-making are integral across various settings, regardless of whether they involve other people or not (for an excellent introduction to the value-based framework of decisions, see Pärnamets et al., 2020). Yet, lesion studies and evidence from clinical populations hint at some level of specificity for computations for social behavior (Corradi-Dell'Acqua et al., 2020; Overgaauw et al., 2020; Rosenthal et al., 2019). Here, we adopt the following position: altruistic decisions arise through the cooperation of multiple distinct but interrelated mechanisms (Cutler & Campbell-Meiklejohn, 2019; Suzuki & O'Doherty, 2020). Some of these mechanisms generalize across a wide variety of settings and decision problems. Prominent examples of generic processes include valuation and cognitive control. Other components of the decision process and their neural substrates become more central when the topic is other people. A fast-growing number of studies point to the role of social cognition, affect, and social context on altruistic decisions in the brain. For further illustration of this notion, we turn to evidence from neuroimaging studies.

Over the last two decades, neuroscience research has examined how social motives and decisions are processed in the brain. The fast-growing number of empirical neuroimaging studies has given rise to several recent systematic reviews (Filkowski et al., 2016; Luo, 2018) and meta-analyses (Bellucci et al., 2017; Cutler & Campbell-Meiklejohn, 2019; Gabay et al., 2014; Zinchenko, 2019). There is broad agreement that prosocial decision-making requires a number of different mental and neural computations. This applies to altruistically and strategically motivated prosocial behaviors (Cutler & Campbell-Meiklejohn, 2019). For example, there is a widespread consensus that social decision-making evokes the processing of desired outcomes. Simply put, decisions result from assigning and comparing values for all choice options. Computing subjective values of choice alternatives recruits brain areas such as the ventromedial prefrontal cortex (VMPFC) and the ventral striatum (VS) (Figure 2). This value network's engagement in social and altruistic decisions is well documented (for recent overviews and meta-analyses, see Bellucci et al., 2020; Cutler & Campbell-Meiklejohn, 2019; Ruff & Fehr, 2014; Suzuki & O'Doherty, 2020). In fact, empirical evidence suggests that this brain network's functional role generalizes across various rewards (Sescousse et al., 2013) and decision problems (Bartra et al., 2013; Clithero & Rangel, 2014; Tusche & Hutcherson, 2018). Damage to the core areas of the brain's valuation network (VMPFC) diminishes altruistic giving (Krajbich et al., 2009).

Notably, these value signals are subject to modifications from other brain networks and mental processes. Inputs into value computations can come from generic cognitive mechanisms or processes relevant to social cognition. One prominent example of the former is executive control and its role in accommodating, for example, contextual goals or salient social norms. Here, the prefrontal cortex, particularly the DLPFC (Figure 2), has been frequently implicated in

altruistic choice (Bellucci et al., 2020; Carlson & Crockett, 2018). Neurocomputational models have started to reveal the precise mechanism of goal-dependent changes in value signals in the DLPFC (Tusche & Hutcherson, 2018) (e.g., see Section 3.2). Neuroimaging studies have also identified inputs from “social” brain networks on value computations during altruistic choice (Hare et al., 2010; Park et al., 2017). Some of these brain regions, such as the medial prefrontal cortex (MPFC) and TPJ, are part of the mentalizing network (discussed below) (Figure 2). For instance, seminal studies have linked brain structure and activation in the TPJ to individuals’ decisions to forgo their own gains in favor of others’ benefits (Morishima et al., 2012; Strombach et al., 2015; Tankersley et al., 2007). Other brain areas, such as the anterior insula (AI) or the mid cingulate cortex (MCC), are involved in the processing of social emotions such as empathy (Figure 2). In the following, we will review the impact of these social processes on altruistic choice.

## 5.2 | Affective and cognitive social processes in altruism

### 5.2.1 | Conceptual distinctions between mentalizing and empathy

Certain cognitive and affective processes play a more prominent role in social settings in which our decision affects the welfare of people around us. This is hardly surprising. To effectively align behaviors with individuals and social groups, we need to understand other peoples’ thoughts, feelings, and hidden goals. As humans, we are remarkably skilled in using these adaptive mental tools that can be broadly referred to as social cognition. Social cognition is not one uniform concept. It encompasses multiple psychological processes that enable us to navigate social settings successfully (Frith, 2008). Broadly speaking, social cognition includes processes used to decode and encode the social world, including information processing about oneself, other people, or social norms (Beer & Ochsner, 2006). Social neuroscience has been instrumental in identifying and delineating the neural underpinnings of various features of social cognition. Here, we will focus on two social processes that can guide altruistic decisions: empathy and mentalizing. Empathy refers to the ability to share the *emotions* of another person (e.g., suffering or joy) (“feeling with”) (De Vignemont & Singer, 2006; Decety & Jackson, 2004). Empathic concerns about others’ welfare have been proposed to be an evolutionary outcome of empathy (De Waal, 2008). Mentalizing refers to a *cognitive* process of inferring and reasoning about others’ mental states, such as their desires, beliefs, thoughts, or intentions (Frith & Frith, 2006; Premack & Woodruff, 1978). Mentalizing does not require affective involvement, distinguishing it from social emotions such as empathy (for a detailed discussion, see Singer & Tusche, 2014). To date, there is still a certain lack of agreement on the concepts and taxonomy (Schurz et al., 2020). Mentalizing is sometimes referred to as cognitive empathy, theory of mind, or perspective-taking (Koster-Hale & Saxe, 2013; Tusche et al., 2016; Warrier et al., 2018). Likewise, empathy is sometimes used as an umbrella term encompassing affective and cognitive processes that enable us to understand others (Zaki, 2017).

Despite the variation in terms and taxonomies, there is widespread agreement that mentalizing and empathy are involved in altruistic decision-making. Psychological theories such as the empathy-altruism hypothesis have long acknowledged the impact of these social processes (Batson, 2011). Seminal research showed that instructing people to take others’ perspectives—via cognitive mentalizing or affective empathy—increases altruistic behaviors (Batson, 2011; Oswald, 1996). Likewise, downregulating social emotions like empathy through moral disengagement (e.g., by ignoring or dehumanizing others in need) has been linked to antisocial and selfish behaviors (Bandura, 2016). The functional role of empathy and mentalizing in prosocial decision-making is intuitive. Putting ourselves in another person’s shoes enables us to understand what others go through. These processes can evoke social motives and alter the weight that individuals place on others’ welfare, providing a bridge to novel computational models of altruism.

### 5.2.2 | Separable neural core networks of empathy and mentalizing

Incorporating social processes in (neuro)computational models of altruism requires conceptual specificity. This is not trivial. As discussed earlier, our ability to understand others’ inner states—their thoughts, goals, and feelings—is a complex, multi-dimensional process. While empathy and mentalizing often work hand-in-hand in our daily lives, they can exert distinct effects on social behaviors (Preckel et al., 2018; Singer & Klimecki, 2014; Singer & Tusche, 2014; Tusche et al., 2016). Delineating these processes on the behavioral level can be challenging. Here, evidence from social neuroscience has provided crucial insights. Several meta-analyses suggest that empathy and mentalizing draw on partly dissociated networks in the brain. Empathy reliably recruits the AI and MCC (Figure 2) (Bellucci et al., 2020; Fan et al., 2011; Kurth et al., 2010;

Lamm et al., 2011) (for evidence on secondary brain networks and domain-specificity, see Ding et al., 2020; Jauniaux et al., 2019; Timmers et al., 2018). Notably, the AI and MCC are also recruited during first-hand experiences of emotional states. Evidence on shared neural codes for empathy for others and the first-hand experience of affective states (Corradi-Dell'Acqua et al., 2016) has interesting implications: to understand others' feelings, we rely on the representations in the brain evoked when we experience these feelings ourselves. Meta-analyses and reviews have also identified core brain networks activated during mentalizing. Inferring other's mental states frequently activates the TPJ, posterior superior temporal sulcus, temporal poles, and MPFC (Figure 2) (Bellucci et al., 2020; Bzdok et al., 2012; Krall et al., 2015; Mar, 2011; Molenberghs et al., 2016; Schurz et al., 2014; Schurz et al., 2017; Van Overwalle, 2009). Thus, while empathy and mentalizing are closely related on a conceptual and functional level, separable core networks in the brain suggest distinct processes with unique inputs into social behaviors (Bzdok et al., 2012; Singer & Tusche, 2014). This evidence closely matched evidence from neuroimaging studies on altruistic choice. Differential recruitment of empathy and mentalizing is linked to variance in altruistic behaviors across people and contexts. Neural substrates of empathy and mentalizing in *altruism* are consistent with meta-analytical findings on core brain networks of empathy and mentalizing reviewed above (Hare et al., 2010; Hein et al., 2010; Masten et al., 2011; Mathur et al., 2010; Morelli et al., 2014; Morishima et al., 2012; Rameson et al., 2012; Telzer et al., 2011; Waytz et al., 2012) (for a recent comparison of meta-analytical maps for empathy, mentalizing, and prosociality, see Bellucci et al., 2020).

### 5.2.3 | Delineating social inputs into altruistic decision-making

Early studies on social processes in altruistic choice often focused on one particular feature of social cognition (e.g., mentalizing). However, studying these processes in isolation limited researchers' ability to delineate process-specific inputs into altruistic choices. More nuanced assessment tools—partly informed by neuroimaging evidence—have also helped the field to overcome this challenge. Significant strides have been made in developing measures of distinct features of social cognition in the form of tasks and questionnaires (Adolphs & Tusche, 2017; Jordan et al., 2016; Kim & Hommel, 2019; Preckel et al., 2018). For example, the novel EmpaToM task simultaneously assesses empathy, mentalizing, compassion, and social meta-cognition (Breil et al., 2021; Kanske et al., 2015; Kanske et al., 2016). The task uses video-based scenarios as naturalistic, dynamic stimuli to evoke these processes and includes control conditions, overcoming prevalent methodological concerns in social neuroscience (Schilbach et al., 2013). Researchers can use these new paradigms to disentangle the contributions of empathy and mentalizing to altruistic decision-making.

For example, a recent neuroimaging study examined how distinct features of social cognition drive variance in altruistic choice (Tusche et al., 2016). The study combined a charitable donation task, the EmpaToM task, self-reports, formal models, and functional brain data (fMRI). The researchers identified three mental processes that drive variance in people's altruistic behavior: feeling with others in need (empathy), taking others' perspectives (mentalizing), and attention shifts (visually orienting towards relevant information). Formal models allowed quantifying the degree to which each individual relied on distinct processes during altruistic decision-making. The estimates of mathematical models revealed that the relative input of empathy and mentalizing on altruistic choices varied across people. In other words, some people relied heavily on affective empathic responses to guide their decisions. Other subjects were more likely to engage “cold” cognitive processes related to mentalizing (e.g., reason about others' needs, desires, or intentions) to guide their behavior. The study also linked variance in these two social processes to dissociable neural computations during the decision process. Brain responses in the AI (but not the TPJ) encoded empathy for beneficiaries during altruistic choice. Neural activation in the TPJ (but not AI) predicted the degree of mentalizing in the donation task. These findings are consistent with meta-analytic evidence on functional segregation of both social processes on the neural level (reviewed in Section 5.2.2). Notably, variance in the degree to which individuals used empathy and mentalizing to guide altruistic decisions generalized to social settings that do not require decisions. The latter was captured in behavior and brain responses in the EmpaToM task. The results suggest that people's general propensity to recruit empathy or to mentalize in social settings determines their contributions to altruistic decisions.

### 5.2.4 | Implications of deconstructing affective and cognitive social inputs into altruism

Findings such as these are essential for several reasons. First, understanding distinct processes that guide social behaviors contributes to our ability to answer a fundamental question: How can we make people more altruistic? Delineating

the factors and processes that drive variance in altruism across people and context is necessary to develop more effective means of increasing prosocial behaviors. For instance, the results by Tusche et al. (2016) suggest that improving the affective or cognitive capacity for understanding others (empathy or mentalizing) are two viable routes to increase human prosociality. A recent large-scale intervention study tested this prediction (Böckler, Tusche, Schmidt, et al., 2018). The study found that particularly mental interventions that cultivate social affective processes boosted altruistically motivated behaviors. This held true for a range of altruistic measures, including game-theoretical tasks (dictator game), hypothetical resource allocations (social value orientation), spontaneous helping, or donations to real-world charities. Evidence from neurally informed interventions can inform policymakers and the general public about how to increase global cooperation. Second, information on people's propensities to engage specific processes to guide altruistic choices (e.g., dispositional empathy and mentalizing) may allow selecting interventions that "fit" with the individual. Third, distinguishing process-specific computations is essential for developing complete theoretical and neuroscientific accounts of altruistic decision-making. For instance, meta-analytical evidence can shed light on brain areas reliably recruited during altruistic and strategic social decisions (Cutler & Campbell-Meiklejohn, 2019). However, they provide limited insights into the concrete variables (what) and computations performed in individual regions or neural networks (how). A computational framework of studies can complement meta-analytic evidence.

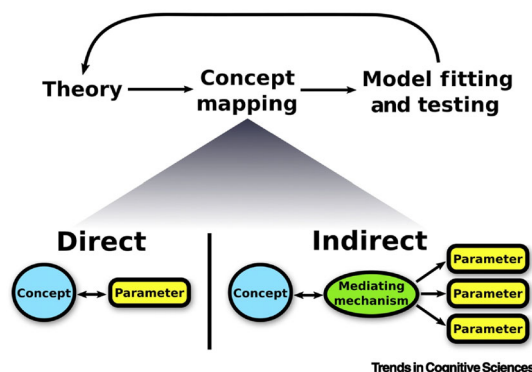
## 6 | SYNTHESIS AND FUTURE DIRECTIONS

### 6.1 | Integrating social affect and cognition into neurocomputational models of altruism

Significant strides have been made in research on human altruism. Fast-growing fields like neuroeconomics have pushed applications of a neurocomputational framework to understand social decision-making. Research in social neuroscience has started to unravel the impact of distinct facets of affective and cognitive social processes on prosocial behaviors. Integrating these two research lines provides an exciting path forward. We propose two tangible advancements.

First, a computational framework can help to reduce the ambiguity of concepts studied in social and affective research on altruistic choice. Despite significant progress, conceptual and neural components of social affect and cognition are still underspecified. Social processes relevant to altruism (e.g., empathy) represent complex, multilevel phenomena (e.g., the valence and arousal associated with an affective state). To date, we know little about *how* these components are encoded in the brain and, more importantly, contribute to decision-making. Mapping parameters of computational models on discrete components of the social process may offer crucial insights (Roberts & Hutcherson, 2019). This mapping can be direct, linking a specific model parameter to a concept, or indirect through a mediating psychological mechanism (Figure 3). This operationalization in a neurocomputational framework enables researchers to test predictions of the models, which in turn can inform theories (for a review on how computational modeling approaches like DDMs enable studies on affect, see Roberts & Hutcherson, 2019). Neurocomputational frameworks of social affect and cognition are still in their infancy. However, recent work in the domain of social learning (Lockwood & Klein-Flügge, 2020; Rosenthal et al., 2019) and strategic decision-making (Hill et al., 2017; Rusch et al., 2020) highlights the potential for neurocomputational approaches to study social processes in altruism.

This brings us to our second point. There is a wide agreement that multiple computational processes occur in parallel during altruistic decision-making. Our understanding of where these processes are computed in the brain has advanced significantly over the last decade. For instance, we highlighted several brain regions involved in value computation, cognitive control, and social processes like empathy or mentalizing in altruism (Figure 2). We also reviewed prior evidence on the neural underpinnings of key variables that guide value computations during altruistic choice (e.g., gains for oneself or others). How these components are *integrated* in the brain to produce coherent behaviors is less established (Suzuki & O'Doherty, 2020). Examining patterns of connectivity between brain areas that encode distinct choice-relevant computations may shed light on this question. Simply put, brain areas involved in altruistic choice do not act in isolation. They are embedded in *interconnected* networks. There is a trend in neuroimaging research to move away from narrow localization towards analyzing distributed brain networks. Suppose we aim to probe how other-regard is integrated into altruistic decision-making. Researchers can examine connectivity patterns between brain regions that perform other-regarding computations (e.g., TPJ) and those believed to encode the integrated subjective value of available choice options (e.g., VMPFC, Figure 2) (Hare et al., 2010; Park et al., 2017). Several analysis tools exist



**FIGURE 3** Mapping affective science concepts to estimates of computational models. Reprinted from Roberts and Hutcherson (2019) (Fig. 1), Copyright 2019, with permission from Elsevier

to examine functional connectivity patterns in the brain (e.g., psycho-physiological interaction analysis [PPI], Friston et al., 1997; dynamical causal modeling [DCM], Friston et al., 2003). Meta-analytic evidence suggests that PPI represents a reliable methodological approach to examine functional integration in the brain (Smith et al., 2016). Likewise, empirical evidence highlights the test–retest reliability of the DCM approach to study connectivity patterns in the brain (Frässle et al., 2015). One significant advantage of DCM is that it allows inferences about the *directionality* of the connectivity (e.g., from brain area A to area B). Functional and structural properties of neural networks can also be linked to estimates of formal models of social preferences. This approach has been shown to reveal social motives that guide altruistic decisions. For example, in a study that used DCM, functional coupling from the MCC to AI has been linked to empathy-driven altruistic motivations (modified dictator game) (Hein et al., 2016). Positive connectivity from the AI to VS has been linked to prosocial decisions driven by reciprocity motives. Reciprocity in this context refers to the motivation to respond in kind (i.e., the desire or expectation that a generous behavior will be returned). In other words, the results suggest that distinct social motives have different neurophysiological representations in the brain at the level of functional networks (Hein et al., 2016). These results echo our earlier argument: while resulting behaviors (generous choice) look alike, underlying social motives can be revealed through a multi-disciplinary computational framework. More generally, the combination of computational modeling, neuroimaging, and connectivity analysis will likely advance studies on how distinct computations are *integrated* in the brain to guide behaviors (for a general discussion beyond altruism, see Suzuki & O’Doherty, 2020). This approach may also inform us about how network configurations change due to situational or dispositional differences in empathy and mentalizing in altruism (or other key computational variables).

## 6.2 | Neurocomputational models of altruism across the lifespan

Other-regarding behaviors emerge during infancy (Dunfield et al., 2011), and lifespan changes in childhood and adolescence have inspired a good deal of research (for an overview, see Eisenberg et al., 2007). Only recently, the field has started to examine age-related changes in altruism in late adulthood. Understanding other-regard in the elderly is essential for one apparent reason: global populations continue to grow older. By 2050, one in six people may be aged 65 or older (Kamiya et al., 2020). Consequently, changes in social preferences in late adulthood have significant social and economic consequences. Promising behavioral evidence suggests that we may become more prosocial as we age (for a recent overview, see Mayr & Freund, 2020, but see Bailey et al., 2020; Rieger & Mata, 2015; Wiepking & James, 2013). This effect holds when researchers control for differences in wealth across age groups (Kettner & Waichman, 2016). For example, charitable giving and volunteering increase across adulthood up to 70 years (Freund & Blanchard-Fields, 2014). While intriguing, these findings do not tell us *why* and *how* other-regard changes across the adult lifespan. We argue that an interdisciplinary, computational framework is uniquely suited to provide answers to these questions.

Preliminary research on altruism in the elderly draws on various measures like donations (Bekkers & Wiepking, 2011), surveys (Bekkers, 2010), and economic games (e.g., dictator game) (Engel, 2011; Kettner & Waichman, 2016; Matsumoto et al., 2016; Rosi et al., 2019) (for a review of age-related changes in economic games, see

Lim & Yu, 2015). However, studies combining the perspectives and analysis tools from neuroscience, psychology, and behavioral economics are still rare. To illustrate the potential of a multi-level approach, we turn to the example of a recent study that bridged this gap. The results suggest that reduced reward activity in the brain in response to self-gains and increased reward activity to others' gains may underlie age-related changes in altruism (Hubbard et al., 2016). In other words, neural evidence suggests that the elderly may genuinely care more about others' well-being. We propose that incorporating formal models can provide even more insight into other-regard. For instance, formal models could quantify age-related changes in contributions of gains for oneself and others and link these estimates of model parameters to the brain's functional and structural properties. Model-based approaches also allow researchers to delineate the role of distinct social motives (e.g., maximizing others' gain vs. fairness). A recent behavioral study combined data from an economic game and computational modeling to examine age-related differences in other-regarding motives (Cho et al., 2020). The study used formal models (Dufwenberg & Kirchsteiger, 2004; Fehr & Schmidt, 1999) to delineate how young and older adults take intention- and outcome-based fairness into consideration during social decision-making. The parameter estimates of formal models suggest that older adults focus more on fair outcomes to guide their decisions and less on other's intentions. These findings explain *why* observable behaviors change as we grow older. Specifically, the results illuminate age-related changes in the relative importance of choice features and motives. In sum, we propose that an interdisciplinary, neurocomputational framework can advance our understanding of age-related changes in altruism.

Social neuroscience offers another window into lifespan changes of altruism and why the elderly may genuinely care more about others' welfare. Popular accounts suggest that the motivation to make strong emotional connections with others increases in older people (socioemotional selectivity theory; Carstensen et al., 2003). Consequently, researchers have examined emotional processes relevant to altruism throughout adulthood. This includes the emotional consequences of helping others (Bjälkebring et al., 2016) and emotional precursors of social decisions like empathy. Older individuals report greater empathy and empathic concern for others than their middle-aged and young counterparts (Sun et al., 2018; Sze et al., 2012), which partly accounts for age-related increases in prosocial behavior (Sze et al., 2012) (for a nuanced review on age-related changes in facets of empathy and mentalizing, see Beadle & De la Vega, 2019). These findings fit into a growing body of evidence that distinct facets of social cognition age differently. Empathy seems to be intact in old age, and empathic concern for others' well-being is even elevated (Reiter et al., 2017; Wieck & Kunzmann, 2015). Other components such as mentalizing or meta-cognition decline in late adulthood (Reiter et al., 2017; for evidence on age-differences when inferring others' intentions, see Reiter et al., 2021). Neuroimaging evidence on the aging brain provides insights into the neurobiological underpinning of these differential trajectories of social processes in late adulthood and decision-making (for reviews, see Beadle & De la Vega, 2019; Lighthall, 2020). Research on this topic is still in its infancy. Preliminary evidence suggests that core brain areas involved in affective processing seem to maintain their structural integrity during healthy aging (Mather, 2012). In light of this evidence, it would seem plausible that older adults rely more heavily on affective processes to guide altruistic decisions. Consistent with this notion, empathy-inducing messages increased altruism in a dictator game in the elderly more than in younger adults (Beadle et al., 2015). In sum, neuroimaging studies, together with formal models of altruism, are uniquely suited to elucidate the origins of process-specific inputs into social decisions in the elderly.

## 7 | CHALLENGES AND RECOMMENDATIONS

Below we outline potential pitfalls and practical advice regarding four topics pertinent to a multidisciplinary, computational framework of altruism: (1) challenges of interdisciplinary research, (2) implementing a computational model (focusing on popular open-source software solutions), (3) theoretical and methodological properties of models (focusing on model validity and parameter recovery), and (4) potential methodological issues regarding the data used to fit a model (focusing on psychometric properties of task-based observations and estimates of computational models).

### 7.1 | Interdisciplinarity

So far, we have highlighted the benefits, insights, and future directions of an interdisciplinary, computational framework of altruism and social motives. However, such an approach is not without challenges. Differences in conceptualization and terminology across disciplines can hamper advances in altruism research (Clavien & Chapuisat, 2013; El

Mallah, 2020; West et al., 2007). For instance, scientists from different disciplines might use the *same* term to refer to different facets of a complex concept or use *different* terms to refer to the same phenomenon (West et al., 2007). Finding a common “language” is not easy and requires commitment (and time). Moreover, limited insights into measures and methodological approaches of other research fields can lead to misconceptions of results. A prime example of this potential pitfall is reverse inference in neuroscientific research (Poldrack, 2006). The same brain area can support different mental processes and functions. Reverse inference refers to situations in which the engagement of a *particular* process (e.g., empathy) is inferred from the mere activation of a specific brain area (e.g., the AI, see Figure 2). Serra (2020)<sup>4</sup> provides an overview of this issue from the viewpoints of economics and neuroeconomics (for a discussion focusing on social neuroscience, see Alós-Ferrer, 2018). We advocate that linking brain data to interpretable parameters of formal models can help provide conceptual clarity and facilitate the interpretation of brain responses. Research disciplines also differ in their attitudes regarding the acceptability of deceptive research practices. Deception can take different forms (e.g., related to project goals or hypotheses). Let us consider the example of the actual presence of other people in experimental paradigms in which subjects (supposedly) interact with others. The logistical complexity and costs of recruiting numerous individuals to a laboratory session can be substantial. This is particularly true for game-theoretical paradigms with many repetitions (usually dozens or hundreds of trials, often with different interaction partners across repetitions). These repetitions are vital for many applications of computational models but can generate practical challenges for researchers. Deceptive methods such as engaging members of the research team or computer scripts (mimicking other people's behavior) represent pragmatic approaches and increase experimental control. This practice is widely accepted in psychology but not economics (Krasnow et al., 2020; Serra, 2020). A primary practical concern is that subjects' suspicion of deception may alter the behavior of interest. Recent evidence indicates that suspicion of deception does not adversely affect subsequent behavior (Krasnow et al., 2020). However, it illustrates how differences in research practices can complicate crosstalk across research fields. Empirical probes into the validity of underlying concerns are one vital step to overcome this challenge.

## 7.2 | Implementing a model: selection of a software solution

How can researchers actually implement a computational framework in their work? Given the focus of our review, we use DDMs to address this question. Below, we provide a practical path for scientists interested in adopting this family of computational models and highlight points that deserve critical reflection. Several open-source software packages exist for diffusion model data analysis (Box 1). Each of these software tools implements vital steps of the computational analysis approach: load the data, build a model, fit a model to data, generate methods for assessing model fit, and simulations, among others (we discuss model fit and simulations in Section 7.3). Tutorials and example datasets provide guidance and practical advice for each step. Readily available, user-friendly analysis tools lower the barriers for computational applications in studies on altruism and social neuroscience more generally. However, they also present researchers with an interesting challenge: choosing the most appropriate software package for their research question and data.

Settling on one software solution—designed to implement particular models in a specific way—requires careful theoretical and methodological deliberation. For example, some users may prefer a Bayesian approach to implement their model (for a general discussion of the advantages of Bayesian data analysis and tutorial, see Kruschke, 2014). Differences in underlying assumptions can also guide researchers' selection. Let us consider the choice between a non-hierarchical and hierarchical model solution (Box 1). Non-hierarchical approaches fit models separately to each subject (assuming that they are entirely independent of each other) or for the whole group (assuming that all subjects are identical). Hierarchical (Bayesian) models, on the other hand, support the simultaneous estimation of individual subject parameters and the group distribution that they are drawn from. In other words, the parameters of a model are estimated at different hierarchical levels. Participants within a group are assumed to be similar—but not identical—to each other (Wiecki et al., 2013; for evidence of the risks of ignoring hierarchical data structure, see Boehm et al., 2018). This assumption has practical consequences. Hierarchical models require fewer data per subject or experimental condition. Thus, when the number of observations per individual is limited, hierarchical models may appeal to potential users (for a tutorial to fit hierarchical Bayesian models, see Lin & Strickland, 2020). Research on the recommended number of trials for diffusion modeling analysis is ongoing (Lerche et al., 2017). We encourage potential users to examine current evidence before applying a specific computational model or, if applicable, when designing a research project. Other practical considerations may guide the choice of a software solution. For instance, researchers may be more comfortable

**Box 1 Open-source software packages for drift diffusion analysis**

Toolboxes for (non-hierarchical) models:

- DMAT (Diffusion Model Analysis Toolbox; Vandekerckhove & Tuerlinckx, 2008)
- Fast-dm-30 (Voss et al., 2015)
- EZ-Diffusion (Wagenmakers et al., 2007)
- ChaRT (Chandrasekaran & Hawkins, 2019)

Toolboxes for hierarchical Bayesian models:

- HDDM (hierarchical drift diffusion model; Wiecki et al., 2013)
- DMC (dynamic models of choice, Heathcote et al., 2019)

with the language of a toolbox (e.g., Python, Matlab, or R). Alternatively, they may care about the efficiency with which models can be built, simulated, and fitted, or base their choice on the usability and handling of the programs.

A detailed discussion of the advantages of available programs is beyond this paper's scope. As a starting point, Table 2 in Shinn et al. (2020) offers a brief comparative outline of popular DDM toolboxes. Here, we highlight one potential selection criteria: the flexibility of the software package to adapt to diverse data, paradigms, and models. Earlier, we outlined a DDM with a limited set of parameters (see Figure 1b). However, DDMs can be extended to model a variety of experimental paradigms and hypotheses. Generalized drift diffusion models promise a useful framework to encapsulate extensions for flexible model designs (for a recent software package, see Shinn et al., 2020). This software solution may be valuable for projects suitable for exploring new model mechanisms or involving innovative experimental designs. Researchers interested in applying different computational models (beyond “traditional” DDMs) may also explore recent toolboxes like the DMC, ChaRT (Box 1), and hBayesDM package (hierarchical Bayesian modeling of decision-making tasks) (Ahn et al., 2017). They support a broad range of models applicable to various decision-making paradigms, including social exchange tasks (like economic games), inhibition tasks, or reversal learning tasks. Many packages also generate trial-by-trial estimates (e.g., HDDM, hBayesDM). This feature allows for seamless integration with model-based analyses of brain data to identify neural underpinnings of latent variables. Researchers interested in applying computational models to brain data will likely embrace analysis approaches that provide this feature.

In sum, selecting a software solution can be based on various theoretical, methodological, and practical considerations. We urge researchers to consult the documentation and tutorials to learn more about a software program's specific analytical approach and assumptions. This is an essential step to determine the appropriateness of the approach for their project. Our second point is more practical in nature. Most open-source toolboxes do not require a background in mathematics or advanced programming skills. Often users execute single lines of code to implement individual steps of the modeling approach. Well-designed tutorials provide hands-on training on their implementation. Nevertheless, some familiarity with programming will likely facilitate computational modeling and increase researchers' comfort levels. To more experienced users, technical programming skills provide further advantages. Open-source software packages allow researchers to inspect, modify, and extend the underlying functions and tailor them to their projects' specific needs. We encourage interested researchers to take advantage of open education resources on popular languages like R or Python.<sup>5</sup> Moreover, online and summer schools on computational modeling can support building basic technical skills and increasing the comfort level with computational analysis approaches.<sup>6</sup>

### 7.3 | Model validation and model selection

Suppose that the appropriate analysis tool is selected, and a model is implemented successfully. Researchers now face the challenge of determining the validity and utility of the model. One crucial question is whether the model provides an adequate account of the observed behavior. In other words, does the model ‘fit’ the empirical observations that researchers want to explain (e.g., the choices, reaction times, or neural responses)? Let us consider the example of the



diffusion model analysis. At its core, it represents the search for a set of estimates for the free parameters of a model (e.g., Figure 2b) that yields a close match between predicted and observed data (Voss et al., 2015). If the estimated parameters account for the observed events, the model is validated (but see below for other important considerations). However, if the fit is inadequate, researchers should be cautious in drawing conclusions based on the estimated parameters of the model (for a general introduction to good practices in cognitive modeling, see Heathcote et al., 2015). Open-source toolboxes (e.g., Box 1) offer different means to quantify the fit between predicted and observed data for the parameter search. In addition to statistical means to assess model fit, they also provide plotting functions for visual checks. These visualizations can be extremely helpful to identify systematic deviations from the actual data (e.g., for specific individuals, experimental conditions, or types of events). If these misfits are consequential for the model (e.g., fail to account for crucial experimental conditions), researchers will likely consider alternative models or an iterative process of refining a model (e.g., see Heathcote et al., 2019). This brings us to the topic of model selection.

In essence, model selection requires researchers to evaluate whether the model performs better than *alternative* models. Choosing between competing models is not trivial. Models can differ on many levels. For instance, models vary in what they deem essential and, consequently, in their complexity. Simply put, models include essential features and omit “unnecessary” details. Imagine a choice between a simple model (characterized by a smaller number of variables and free parameters) and a more complex model. Simpler models are usually preferable. This notion is based on concerns that the risk for overfitting can increase with higher complexity (Vandekerckhove et al., 2015). Overfitting refers to situations in which a model describes a particular dataset very well (including noise) but fails to generalize to new data. Even a perfect model fit is insufficient if the model does not account for other observations that supposedly capture the same cognitive process. One common methodological approach to test for a model's ability to generalize to “new” data is cross-validation (for an introduction, see Berrar, 2019; for a tutorial in R, see Song et al., 2021). Alternatively, popular metrics for model comparisons like the Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwarz, 1978) assess the goodness of fit in relation to the complexity of a model (effectively penalizing models with more free parameters; for discussion of differences of AIC, BIC, and other selection methods, and recommendations for their respective use, see Aho et al., 2014; Evans, 2019). Researchers will select a model with more parameters if it provides a substantially better account of the observed data—despite this penalty—than a model with fewer parameters. Pitt and Myung (2002) provide a comprehensive overview of criteria for model selection (including a more nuanced description of complexity beyond the number of parameters). A discussion of model complexity in diffusion modeling can be found in Lerche and Voss (2016). We urge researchers to use methods like AIC and BIC as one helpful tool—among many—when choosing between competing models. While the use and interpretation of these metrics seem straightforward, they ought to be considered together with other methodological and theoretical properties of a model.

Another method to assess model performance is to evaluate how well simulations can recover the estimated parameters of a model (parameter recovery; for review of the broader issue of model recovery and model falsification, see Palminteri et al., 2017). Here, researchers *start* with a set of estimated parameters of the model. Based on these known parameter values, simulations generate new synthetic data. The model is then fit to the synthetic data (as if they had been collected in an actual research project), yielding a new set of parameters. Researchers can now assess the match between these “recovered” parameter estimates and the ones used to simulate the data in the first place. Researchers could consider simplifying the model if a model fails to recover known parameter values (i.e., eliminate free parameters). Both model fit and parameter recovery will depend on numerous details of a project. For example, the recoverability of parameters systematically varies as a function of the number of trials and subjects (e.g., Pedersen & Frank, 2020). Ongoing research efforts on these factors promise practical recommendations for researchers interested in applying computational models (e.g., regarding choosing the minimum sample size and trials required to estimate parameters reliably). Simulations can also be helpful during the preparation of a research project or the development of a new paradigm. For example, they can indicate whether researchers have “enough” data to estimate the parameters of a model reliably.

Theoretical considerations are equally important when assessing the utility of candidate models. We will summarize some widely accepted recommendations (for details, see Fum et al., 2007). For starters, models should be plausible. In other words, variables and assumptions of a model should be justifiable in light of psychological or biological evidence. Researchers may also deliberate on features of the theoretical and practical utility of a model. Models based on general cognitive theories—covering a broad range of phenomena—are preferable over models designed to account for specific behaviors observed with certain constrained paradigms. Moreover, the predictions made by different models can guide model selection. If faced with multiple models that describe the observed behavior, models that make bold, surprising,

or risky predictions may ultimately be more valuable to researchers. Predictions made by a model should also be precise (making it easy to falsify them). In sum, researchers applying computational models to their research will likely encounter many different candidate models explaining and reproducing their empirical data. Here, criteria related to a model's theoretical utility provide a viable path forward.

## 7.4 | Psychometric properties of measures

We want to highlight one final methodological challenge that broadly relates to the question: what data should researchers collect? Earlier, we introduced various research fields that traditionally rely on different measures to study altruism. The psychometric properties of some of these measures have recently been questioned. These concerns raise a critical question for neurocomputational research on altruism and social motives: are the measures *valid* (capture what we intend to measure?) and *reliable* (stable under repeated measures)? Especially the reliability of task-based observations has come under scrutiny. Evidence from different domains points towards poorer reliability of task-based measures than for questionnaires (Enkavi et al., 2019; Frey et al., 2017). The weaker reliability of many observational behavioral measures is likely due to methodological and conceptual reasons (for discussion, see Dang et al., 2020; Hedge et al., 2018).

Given our advocacy for the utility of game-theoretical tasks, this issue has important implications for computational models of social preferences. For instance, poor stability of task-based observations over repeated measurements questions their utility for capturing individual differences or lifespan changes in social preferences. How reliable are behavioral observations in the dictator game and computational parameters estimated from observed behaviors? The answer to this question will likely depend on details of the projects (e.g., number of choice trials, the delay between repeated measures, or characteristics and size of the subject sample). Providing concrete recommendations about acceptable psychometric properties is therefore difficult. Moreover, research on the reliability of parameters of formal models is still in its infancy. Preliminary evidence regarding the psychometric properties of diffusion model parameters seems promising (Lerche & Voss, 2017; Schubert et al., 2016; Shahar et al., 2019; von Krause et al., 2021). Increased availability of (dictator game) data—and estimates from computational models—in the context of open science initiatives like OSF (<https://osf.io/>) will likely advance our understanding of psychometric properties of behavioral measures, their dependence on features of the project, and facilitate practical recommendations for researchers. Preregistration of studies using computational modeling offers another avenue to address the crisis of confidence facing various measures of social preferences. Box 2 outlines guidelines and unique challenges for computational modelers when preregistering their work plans (for general guidelines for preregistration, see Nosek et al., 2019).

There are causes for cautious optimism. First, recent findings suggest high internal consistency ( $\alpha = 0.91$ ) and moderate test–retest reliability for the dictator game ( $r = 0.60$ ) (McAuliffe et al., 2018). Second, initial evidence suggests satisfying reliability for the main parameters of diffusion models ( $r_s > 0.70$ , Lerche & Voss, 2017). The reliability of computational parameters (DDM) also did not significantly differ from “raw” observed behaviors (response times) in self-regulation tasks (Enkavi et al., 2019). Given the interpretability of DDM parameters, Enkavi et al. (2019) conclude that researchers may therefore prefer using computational parameters. Third, where possible, adopting a latent variable approach presents a way to improve the psychometric properties of individual measures. Multiple noisy measures can be integrated into latent variables. Researchers can use data-driven approaches (e.g., principal component analysis, exploratory factor analysis) or previously identified latent factors (Böckler, Tusche, Schmidt, et al., 2018; McAuliffe et al., 2019; Peysakhovich et al., 2014). Initial evidence suggests that latent variables of altruism are stable over time ( $r_s = 0.77$ – $0.90$  over nine months; Böckler, Tusche, Schmidt, et al., 2018). The results mirror evidence of more robust latent factor scores in other domains (e.g., latent state–trait model of drift rate parameters, Schubert et al., 2016; self-regulation, Eisenberg et al., 2019; risk-preferences, Frey et al., 2017). This notion has important implications. Latent variables may be more suitable for trait-like assessments of individual differences in altruism that generalize across time and settings. Specific behavioral measures may be more appropriate for studying decision processes as they unfold or contexts that enhance or diminish certain social motives (e.g., fairness, other-regard) (for a similar argument, see Dang et al., 2020).

A latent variable approach also addresses another potential concern: different measures may capture different facets of a complex psychological construct. For instance, self-reports and behavior observed in common tasks seem to represent distinct (but interrelated) components of prosociality (Böckler, Tusche, Schmidt, et al., 2018) (for initial evidence on discriminant, convergent, and weak ecological validity of common altruistic measures like the dictator game, see

**Box 2 Getting started with preregistration for computational modelers**

Comprehensive theoretical and practical guidelines on preregistration:

- “Robust modeling in Cognitive Science”—Lee et al. (2019)
- “Robust standards in Cognitive Science”—Crüwell et al. (2019)
- “Preregistration in complex contexts: A preregistration template for the application of cognitive models”—Crüwell and Evans (2019)

Diverse discussions on best practices and preregistration (e.g., points of agreement and disagreement, practical extensions, or theoretical considerations):

- Back-to-back special issues of Computational Brain and Behavior
- Blog post

Popular Repositories commonly used:

- OSF—most popular across research disciplines
- AsPredicted—mainly used within psychology
- AEA—mainly used within economics

Böckler et al., 2016; Galizzi & Navarro-Martinez, 2019; Peysakhovich et al., 2014; Yamagishi et al., 2013). Evidence linking different measures with specific components of a complex construct does not negate the value of multi-disciplinary frameworks. It merely emphasizes the need for researchers to consider the validity of potential measurement tools to align them with the specific psychological construct of interest. It also stresses that different measures cannot be used interchangeably without careful inspection of whether they measure similar things.

We focused on psychometric properties of behavioral choice tasks like the dictator game and estimates of diffusion models. However, other measures commonly used to study altruism are also subject to critical evaluation of their psychometric properties. This includes questions regarding the (ecological) validity of social cognition tasks (e.g., EmpaToM task, see Hildebrandt et al., 2021; Kanske et al., 2015) or the reliability of neuroimaging data (Elliott et al., 2020; Fröhner et al., 2019; Noble et al., 2021). Consequently, our main recommendation concerns a broad array of measures used to study altruism and social motives. Researchers need to pay attention to the psychometric properties of the measures used to fit computational models. Faulty, incomplete, or imprecise data will impact the utility of formal models. Sophisticated computational analysis tools cannot substitute careful task design, thoughtful selection, and appropriate use of the measures used to capture the concepts of interest.

## 8 | CONCLUSIONS

Despite these caveats, we argue that the benefits of an interdisciplinary, computational approach to altruism outweigh its challenges. Formal modeling approaches of altruism are potent tools in studying social motives of behavior and their neural underpinnings in the brain. They yield refined insights into the mechanisms of altruistic decision-making and provide researchers with powerful ways to test competing theories. We showcase exciting ways in which the intersection of social neuroscience and computational models can complement each other and advance our understanding of social behavior. Drawing on examples of DDMs, we highlight how a cross-disciplinary, neurocomputational approach can disentangle the specific processes that drive variance in altruism across people and contexts. These models enable researchers to empirically test why, how, and when individuals will behave altruistically towards others (or not). They also allow identifying unifying mechanisms that drive various types of decisions. It provides crucial insights into whether dedicated mental processes and brain systems guide (pro)social choices. Interdisciplinary fields like neuroeconomics and social neuroscience have started to embrace cross-disciplinary computational approaches.

Yet, to date, many research questions remain unanswered. How are various components of the decision process integrated in the brain to produce coherent behaviors? How do social preferences and altruistic choices change across the lifespan (especially during late adulthood)? We advocate for the utility of an interdisciplinary, computational approach to address these questions and develop a unified framework of altruistic and social behavior. We also made practical recommendations regarding these models and their potential pitfalls to open a path for researchers interested in adopting computational methods in their research. In sum, we argue that multi-disciplinary perspectives and computational modeling have advanced our collective understanding of other-regard and social decisions that affect the people around us.

## ACKNOWLEDGMENTS

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) (funding reference number RGPIN-2019-04329), the Canada Foundation for Innovation, John R. Evans Leaders Fund (CFI JELF project 38812), and the NIMH Conte Center (2P50 MH094258). We thank Ian D. Roberts and Caroline J. Charpentier for helpful comments on sections of this manuscript.

## CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## AUTHOR CONTRIBUTIONS

**Lisa M. Bas:** Conceptualization; methodology; validation; visualization; writing - original draft; writing-review & editing. **Anita Tusche:** Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; supervision; validation; visualization; writing - original draft; writing-review & editing.

## ORCID

Anita Tusche  <https://orcid.org/0000-0003-4180-8447>

Lisa M. Bas  <https://orcid.org/0000-0003-0376-0059>

## ENDNOTES

- <sup>1</sup> Comprehensive overview and historic perspective on behavioral research on social emotions and cognition in human prosociality.
- <sup>2</sup> Discusses applications of experimental games in studies on social decision-making beyond altruism (e.g., intergroup polarization and conflict, cross-cultural differences in cooperation and norm enforcement, computational models related to the formation and updating of social preferences and beliefs).
- <sup>3</sup> Beginner-friendly, pragmatic, and details-oriented introduction on how to relate computational models to data (including behavioral choice, eye-tracking, neuroimaging data).
- <sup>4</sup> Accessible introduction to challenges and limitations of inter-disciplinary approaches at the cross-sections of economics, neuroscience, and modeling approaches.
- <sup>5</sup> Free online courses on programming languages and computational modeling are offered by Coursera, <https://www.coursera.org/>; Datacamp provides video tutorials and coding challenges on R, Python, and analysis approaches, <https://www.datacamp.com/>; Codecademy offers introduction tutorials on R and Python, among others (<https://www.codecademy.com/catalog/all>).
- <sup>6</sup> For example, <https://neuromatch.io/academy/>.

## RELATED WIREs ARTICLES

[Neural, cognitive, and evolutionary foundations of human altruism](#)

[Neural mechanisms of decision-making in aging](#)

[Enhancing Models of Social and Strategic Decision Making with Process Tracing and Neural Data](#)

## REFERENCES

- Adolphs, R. (2010). Conceptual challenges and directions for social neuroscience. *Neuron*, *65*(6), 752–767.
- Adolphs, R., & Tusche, A. (2017). From faces to prosocial behavior: Cues, tools, and mechanisms. *Current Directions in Psychological Science*, *26*(3), 282–287. <https://doi.org/10.1177/0963721417694656>
- Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry*, *1*, 24–57.
- Aho, K., Derryberry, D., & Peterson, T. (2014). Model selection for ecologists: The worldviews of AIC and BIC. *Ecology*, *95*(3), 631–636. <https://doi.org/10.1890/13-1452.1>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.
- Akerlof, G. A., & Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, *19*(1), 9–32.
- Alós-Ferrer, C. (2018). A review essay on social neuroscience: Can research on the social brain and economics inform each other? *Journal of Economic Literature*, *56*(1), 234–264.
- André, K., Bureau, S., Gautier, A., & Rubel, O. (2017). Beyond the opposition between altruism and self-interest: Reciprocal giving in reward-based crowdfunding. *Journal of Business Ethics*, *146*(2), 313–332.
- Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–753.
- Bailey, P. E., Brady, B., Ebner, N. C., & Ruffman, T. (2020). Effects of age on emotion regulation, emotional empathy, and prosocial behavior. *The Journals of Gerontology: Series B*, *75*(4), 802–810.
- Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. Worth Publishers.
- Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology*, *107*(3), 393–413.
- 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.
- Batson, C. D. (2011). *Altruism in humans*. Oxford University Press.
- Batson, C. D., & Powell, A. A. (2003). Altruism and prosocial behavior. *Handbook of Psychology*, *5*, 463–484.
- Beadle, J. N., & De la Vega, C. E. (2019). Impact of aging on empathy: Review of psychological and neural mechanisms. *Frontiers in Psychiatry*, *10*, 331.
- Beadle, J. N., Sheehan, A. H., Dahlben, B., & Gutchess, A. H. (2015). Aging, empathy, and prosociality. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *70*(2), 213–222.
- Beer, J. S., & Ochsner, K. N. (2006). Social cognition: a multi level analysis. *Brain Research*, *1079*(1), 98–105.
- Bekkers, R. (2010). Who gives what and when? A scenario study of intentions to give time and money. *Social Science Research*, *39*(3), 369–381.
- Bekkers, R., & Wiepking, P. (2011). Who gives? A literature review of predictors of charitable giving part one: Religion, education, age and socialisation. *Voluntary Sector Review*, *2*(3), 337–365.
- Bellucci, G., Camilleri, J. A., Eickhoff, S. B., & Krueger, F. (2020). Neural signatures of prosocial behaviors. *Neuroscience & Biobehavioral Reviews*, *118*, 186–195.
- 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.
- Bénabou, R., & Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, *126*(2), 805–855.
- Berrar, D. (2019). Cross-Validation. In S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Eds.), *Encyclopedia of bioinformatics and computational biology* (pp. 542–545). Academic Press. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>
- Bethlehem, R. A., Allison, C., van Andel, E. M., Coles, A. I., Neil, K., & Baron-Cohen, S. (2017). Does empathy predict altruism in the wild? *Social Neuroscience*, *12*(6), 743–750.
- Bjälkebring, P., Västfjäll, D., Dickert, S., & Slovic, P. (2016). Greater emotional gain from giving in older adults: Age-related positivity bias in charitable giving. *Frontiers in Psychology*, *7*, 846.
- Böckler, A., Tusche, A., Schmidt, P., & Singer, T. (2018). Distinct mental trainings differentially affect altruistically motivated, norm motivated, and self-reported prosocial behaviour. *Scientific Reports*, *8*(1), 13560–13560. <https://doi.org/10.1038/s41598-018-31813-8>
- 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>
- Böckler, A., Tusche, A., & Singer, T. (2018). The structure of human prosociality revisited: Corrigendum and addendum to Böckler, Tusche, and Singer (2016). *Social Psychological and Personality Science*, *9*(6), 754–759.
- Boehm, U., Marsman, M., Matzke, D., & Wagenmakers, E.-J. (2018). On the importance of avoiding shortcuts in applying cognitive models to hierarchical data. *Behavior Research Methods*, *50*(4), 1614–1631. <https://doi.org/10.3758/s13428-018-1054-3>
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765.
- Bolenz, F., Reiter, A. M., & Eppinger, B. (2017). Developmental changes in learning: Computational mechanisms and social influences. *Frontiers in Psychology*, *8*, 2048.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166–193.

- Breil, C., Kanske, P., Pittig, R., & Böckler, A. (2021). A revised instrument for the assessment of empathy and theory of mind in adolescents: Introducing the EmpaToM-Y. *Behavior Research Methods*, 1–15. <https://doi.org/10.3758/s13428-021-01589-3>
- Bretschneider, U., & Leimeister, J. M. (2017). Not just an ego-trip: Exploring backers' motivation for funding in incentive-based crowdfunding. *The Journal of Strategic Information Systems*, 26(4), 246–260.
- Bruch, E., & Feinberg, F. (2017). Decision-making processes in social contexts. *Annual Review of Sociology*, 43, 207–227.
- Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, 217(4), 783–796.
- Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.
- Carlson, R. W., & Crockett, M. J. (2018). The lateral prefrontal cortex and moral goal pursuit. *Current Opinion in Psychology*, 24, 77–82.
- Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, 27(2), 103–123.
- Chandrasekaran, C., & Hawkins, G. E. (2019). ChaRTr: An R toolbox for modeling choices and response times in decision-making tasks. *Journal of Neuroscience Methods*, 328, 108432.
- Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3), 817–869.
- 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.
- Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1), 431–457.
- Cheong, J. H., Jolly, E., Sul, S., & Chang, L. J. (2017). Computational models in social neuroscience. *Computational Models of Brain and Behavior*, 229–244. <https://doi.org/10.3758/s13428-021-01589-3>.
- Cho, I., Song, H.-J., Kim, H., & Sul, S. (2020). Older adults consider others' intentions less but allocentric outcomes more than young adults during an ultimatum game. *Psychology and Aging*, 35(7), 974–980.
- Clavien, C., & Chapuisat, M. (2013). Altruism across disciplines: One word, multiple meanings. *Biology and Philosophy*, 28(1), 125–140.
- Clithero, J. A., & Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, 9(9), 1289–1302.
- Corradi-Dell'Acqua, C., Ronchi, R., Thomasson, M., Bernati, T., Saj, A., & Vuilleumier, P. (2020). Deficits in cognitive and affective theory of mind relate to dissociated lesion patterns in prefrontal and insular cortex. *Cortex*, 128, 218–233.
- Corradi-Dell'Acqua, C., Tusche, A., Vuilleumier, P., & Singer, T. (2016). Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. *Nature Communications*, 7(1), 1–12.
- Crockett, M. J. (2016). How formal models can illuminate mechanisms of moral judgment and decision making. *Current Directions in Psychological Science*, 25(2), 85–90.
- Crüwell, S., & Evans, N. J. (2019). *Preregistration in complex contexts: A preregistration template for the application of cognitive models*.
- Crüwell, S., Stefan, A. M., & Evans, N. J. (2019). Robust standards in cognitive science. *Computational Brain & Behavior*, 2(3), 255–265.
- Cutler, J., & Campbell-Meiklejohn, D. (2019). A comparative fMRI meta-analysis of altruistic and strategic decisions to give. *NeuroImage*, 184, 227–241.
- Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, 24(4), 267–269.
- De Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, 10(10), 435–441.
- De Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59, 279–300.
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), 71–100.
- Declerck, C., & Boone, C. (2015). *Neuroeconomics of prosocial behavior: The compassionate egoist*. Academic Press.
- Diederich, A., & Busemeyer, J. R. (2012). Computational modeling. In *APA handbook of research methods in psychology, Vol. 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 387–405). American Psychological Association. <https://doi.org/10.1037/13620-021>
- Ding, R., Ren, J., Li, S., Zhu, X., Zhang, K., & Luo, W. (2020). Domain-general and domain-preferential neural correlates underlying empathy towards physical pain, emotional situation and emotional faces: An ALE meta-analysis. *Neuropsychologia*, 137, 107286.
- Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., & Penner, L. A. (2017). *The social psychology of prosocial behavior*. Psychology Press.
- Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2), 268–298.
- Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, 16(3), 227–247.
- Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, 10(1), 1–13.
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2007). Prosocial development. In *Handbook of child psychology* (Vol. 3). John Wiley & Sons.
- El Mallah, S. (2020). Conceptualization and measurement of adolescent prosocial behavior: Looking back and moving forward. *Journal of Research on Adolescence*, 30, 15–38.
- Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychological Science*, 31(7), 792–806.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14(4), 583–610. <https://doi.org/10.1007/s10683-011-9283-7>

- Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test-retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences of the United States of America*, 116(12), 5472–5477. <https://doi.org/10.1073/pnas.1818430116>
- Evans, N. J. (2019). Assessing the practical differences between model selection methods in inferences about choice response time tasks. *Psychonomic Bulletin & Review*, 26(4), 1070–1098.
- Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, 35(3), 903–911.
- 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.
- Fehr, E., & Rockenbach, B. (2004). Human altruism: Economic, neural, and evolutionary perspectives. *Current Opinion in Neurobiology*, 14(6), 784–790.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.
- Filkowski, M. M., Cochran, R. N., & Haas, B. W. (2016). Altruistic behavior: Mapping responses in the brain. *Neuroscience and Neuroeconomics*, 5, 65–75.
- Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, 67, 641–666.
- Forstmann, B. U., Wagenmakers, E.-J., Eichele, T., Brown, S., & Serences, J. T. (2011, 2011/06/01/). Reciprocal relations between cognitive neuroscience and formal cognitive models: Opposites attract? *Trends in Cognitive Sciences*, 15(6), 272–279. <https://doi.org/10.1016/j.tics.2011.04.002>
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3), 347–369.
- Frässle, S., Stephan, K. E., Friston, K. J., Steup, M., Krach, S., Paulus, F. M., & Jansen, A. (2015). Test-retest reliability of dynamic causal modeling for fMRI. *NeuroImage*, 117, 56–66.
- Freund, A. M., & Blanchard-Fields, F. (2014). Age-related differences in altruism across adulthood: Making personal financial gain versus contributing to the public good. *Developmental Psychology*, 50(4), 1125–1136.
- Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3(10), e1701381.
- Friston, K. J., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6(3), 218–229.
- Friston, K. K., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4), 1273–1302.
- Frith, C. D. (2008). Social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1499), 2033–2039.
- Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4), 531–534.
- Fröhner, J. H., Teckentrup, V., Smolka, M. N., & Kroemer, N. B. (2019). Addressing the reliability fallacy in fMRI: Similar group effects may arise from unreliable individual effects. *NeuroImage*, 195, 174–189.
- Fum, D., Del Missier, F., & Stocco, A. (2007). The cognitive modeling of human behavior: Why a model is (sometimes) better than 10,000 words. *Cognitive Systems Research*, 8(3), 135–142.
- Gabay, A. S., Radua, J., Kempton, M. J., & Mehta, M. A. (2014). The ultimatum game and the brain: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 47, 549–558.
- Gaesser, B., Shimura, Y., & Cikara, M. (2020). Episodic simulation reduces intergroup bias in prosocial intentions and behavior. *Journal of Personality and Social Psychology*, 118(4), 683–705.
- Galizzi, M. M., & Navarro-Martinez, D. (2019). On the external validity of social preference games: A systematic lab-field study. *Management Science*, 65(3), 976–1002.
- Genevsky, A., Västfjäll, 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.
- Genevsky, A., Yoon, C., & Knutson, B. (2017). When brain beats behavior: Neuroforecasting crowdfunding outcomes. *Journal of Neuroscience*, 37(36), 8625–8634.
- Giudici, G., Guerini, M., & Rossi-Lamastra, C. (2018). Reward-based crowdfunding of entrepreneurial projects: The effect of local altruism and localized social capital on proponents' success. *Small Business Economics*, 50(2), 307–324.
- Gold, J. I., & Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience*, 23(2), 632–651.
- Guala, F., & Mittone, L. (2010). Paradigmatic experiments: The dictator game. *The Journal of Socio-Economics*, 39(5), 578–584.
- Hackel, L. M., & Amodio, D. M. (2018). Computational neuroscience approaches to social cognition. *Current Opinion in Psychology*, 24, 92–97.
- Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*, 316(5831), 1622–1625.
- Hare, T. A., Camerer, C. F., Knoeplfle, 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.
- Heathcote, A., Brown, S. D., & Wagenmakers, E.-J. (2015). An introduction to good practices in cognitive modeling. In *An introduction to model-based cognitive neuroscience* (pp. 25–48). Springer.

- Heathcote, A., Lin, Y.-S., Reynolds, A., Strickland, L., Gretton, M., & Matzke, D. (2019). Dynamic models of choice. *Behavior Research Methods*, *51*(2), 961–985.
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. <https://doi.org/10.3758/s13428-017-0935-1>
- Heekeren, H., Marrett, S., & Ungerleider, L. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*(6), 467–479.
- Hein, G., Morishima, Y., Leiberg, S., Sul, S., & Fehr, E. (2016). The brain's functional network architecture reveals human motives. *Science*, *351*(6277), 1074–1078.
- Hein, G., Silani, G., Preuschhoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149–160.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., & Ensminger, J. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*(6), 795–815.
- Hildebrandt, M. K., Jauk, E., Lehmann, K., Maliske, L., & Kanske, P. (2021). Brain activation during social cognition predicts everyday perspective-taking: A combined fMRI and ecological momentary assessment study of the social brain. *NeuroImage*, *227*, 117624. <https://doi.org/10.1016/j.neuroimage.2020.117624>
- Hill, C. A., Suzuki, S., Polania, R., Moisa, M., O'doherty, J. P., & Ruff, C. C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature Neuroscience*, *20*(8), 1142–1149.
- Hubbard, J., Harbaugh, W. T., Srivastava, S., Degras, D., & Mayr, U. (2016). A general benevolence dimension that links neural, psychological, economic, and life-span data on altruistic tendencies. *Journal of Experimental Psychology: General*, *145*(10), 1351–1358.
- Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1). Sinauer Associates Sunderland.
- Hui, B. P. H., Ng, J. C. K., Berzaghi, E., Cunningham-Amos, L. A., & Kogan, A. (2020). Rewards of kindness? A meta-analysis of the link between prosociality and well-being. *Psychological Bulletin*, *146*(12), 1084–1116. <https://doi.org/10.1037/bul0000298>
- Hutcherson, C., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, *87*(2), 451–462.
- Insko, C. A., & Schopler, J. (2013). *Experimental social psychology: Text with illustrative readings*. Academic Press.
- Jauniaux, J., Khatibi, A., Rainville, P., & Jackson, P. L. (2019). A meta-analysis of neuroimaging studies on pain empathy: Investigating the role of visual information and observers' perspective. *Social Cognitive and Affective Neuroscience*, *14*(8), 789–813. <https://doi.org/10.1093/scan/nsz055>
- Joiner, J., Piva, M., Turrin, C., & Chang, S. W. (2017). Social learning through prediction error in the brain. *NPJ Science of Learning*, *2*(1), 1–9.
- Jordan, M. R., Amir, D., & Bloom, P. (2016). Are empathy and concern psychologically distinct? *Emotion*, *16*(8), 1107–1116.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of Business*, *59*, S285–S300.
- Kamiya, Y., Mun Sim Lai, N., & Schmid, K. (2020). *World population ageing 2020 highlights: Living arrangements of older persons*. Retrieved from [https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Nov/undesapd\\_2020\\_world\\_population\\_ageing\\_highlights.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Nov/undesapd_2020_world_population_ageing_highlights.pdf)
- Kanske, P., Böckler, A., Trautwein, F.-M., Parianen Lesemann, F. H., & Singer, T. (2016). Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive and Affective Neuroscience*, *11*(9), 1383–1392.
- Kanske, P., Böckler, A., Trautwein, F.-M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and theory of mind. *NeuroImage*, *122*, 6–19. <https://doi.org/10.1016/j.neuroimage.2015.07.082>
- Kelley, H. H., Holmes, J. G., Kerr, N. L., Reis, H. T., Rusbult, C. E., & Van Lange, P. A. (2003). *An atlas of interpersonal situations*. Cambridge University Press.
- Kettner, S. E., & Waichman, I. (2016). Old age and prosocial behavior: Social preferences or experimental confounds? *Journal of Economic Psychology*, *53*, 118–130.
- Kim, D., & Hommel, B. (2019). Social cognition 2.0: Toward mechanistic theorizing. *Frontiers in Psychology*, *10*, 2643.
- Kononov, A., & Krajbich, I. (2019). Over a decade of neuroeconomics: What have we learned? *Organizational Research Methods*, *22*(1), 148–173.
- Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, *79*(5), 836–848.
- Krajbich, I., Adolphs, R., Tranel, D., Denburg, N. L., & Camerer, C. F. (2009). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *Journal of Neuroscience*, *29*(7), 2188–2192.
- Krajbich, I., Hare, T., Bartling, B., Morishima, Y., & Fehr, E. (2015). A common mechanism underlying food choice and social decisions. *PLoS Computational Biology*, *11*(10), e1004371.
- Krall, S. C., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P. T., Eickhoff, S. B., Fink, G. R., & Konrad, K. (2015). The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Structure and Function*, *220*(2), 587–604.
- Krasnow, M. M., Howard, R. M., & Eisenbruch, A. B. (2020). The importance of being honest? Evidence that deception may not pollute social science subject pools after all. *Behavior Research Methods*, *52*(3), 1175–1188.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. 2nd editions, Academic Press.



- Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, *214*(5–6), 519–534.
- 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.
- Kwak, Y., & Huettel, S. A. (2016). Prosocial reward learning in children and adolescents. *Frontiers in Psychology*, *7*, 1539.
- Kwak, Y., Pearson, J., & Huettel, S. A. (2014). Differential reward learning for self and others predicts self-reported altruism. *PLoS One*, *9*(9), e107621.
- 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.
- Larney, A., Rotella, A., & Barclay, P. (2019). Stake size effects in ultimatum game and dictator game offers: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *151*, 61–72.
- Le, B. M., Impett, E. A., Lemay, E. P., Jr., Muise, A., & Tskhay, K. O. (2018). Communal motivation and well-being in interpersonal relationships: An integrative review and meta-analysis. *Psychological Bulletin*, *144*(1), 1–25.
- Lee, D., & Seo, H. (2016). Neural basis of strategic decision making. *Trends in Neurosciences*, *39*(1), 40–48.
- Lee, M. D., Criss, A. H., Devezzer, B., Donkin, C., Etz, A., Leite, F. P., Matzke, D., Rouder, J. N., Trueblood, J. S., & White, C. N. (2019). Robust modeling in cognitive science. *Computational Brain & Behavior*, *2*(3), 141–153.
- Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious [methods]. *Frontiers in Psychology*, *7*(1324), 1–14. <https://doi.org/10.3389/fpsyg.2016.01324>
- Lerche, V., & Voss, A. (2017, May). Retest reliability of the parameters of the Ratcliff diffusion model. *Psychological Research*, *81*(3), 629–652. <https://doi.org/10.1007/s00426-016-0770-5>
- Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, *49*(2), 513–537.
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, *66*(1), 799–823. <https://doi.org/10.1146/annurev-psych-010213-115043>
- Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, *1*(3), 593–622.
- Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.
- Lighthall, N. R. (2020). Neural mechanisms of decision-making in aging. *WIREs Cognitive Science*, *11*(1), e1519.
- Lim, K. T. K., & Yu, R. (2015). Aging and wisdom: Age-related changes in economic and social decision making. *Frontiers in Aging Neuroscience*, *7*, 120.
- Lin, Y.-S., & Strickland, L. (2020). Evidence accumulation models with R: A practical guide to hierarchical Bayesian methods. *The Quantitative Methods for Psychology*, *16*(2), 133–153.
- Lockwood, P. L., Apps, M. A., & Chang, S. W. (2020). Is there a ‘social’ brain? Implementations and algorithms. *Trends in Cognitive Sciences*, *24*, 802–813.
- Lockwood, P. L., Apps, M. A., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(35), 9763–9768.
- Lockwood, P. L., & Klein-Flügge, M. (2020). Computational modelling of social cognition and behaviour—A reinforcement learning primer. *Social Cognitive and Affective Neuroscience*, *1*, 1–11.
- Love, B. C. (2015). The algorithmic level is the bridge between computation and brain. *Topics in Cognitive Science*, *7*(2), 230–242.
- Luo, J. (2018). The neural basis of and a common neural circuitry in different types of pro-social behavior [review]. *Frontiers in Psychology*, *9*(859), 1–17. <https://doi.org/10.3389/fpsyg.2018.00859>
- Ma, Y., Wang, C., & Han, S. (2011). Neural responses to perceived pain in others predict real-life monetary donations in different socioeconomic contexts. *NeuroImage*, *57*(3), 1273–1280.
- Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, *62*, 103–134.
- 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.
- Mather, M. (2012). The emotion paradox in the aging brain. *Annals of the New York Academy of Sciences*, *1251*, 33–49.
- Mathur, V. A., Harada, T., Lipke, T., & Chiao, J. Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *NeuroImage*, *51*(4), 1468–1475.
- Matsumoto, Y., Yamagishi, T., Li, Y., & Kiyonari, T. (2016). Prosocial behavior increases with age across five economic games. *PLoS One*, *11*(7), e0158671.
- Mayr, U., & Freund, A. M. (2020). Do we become more prosocial as we age, and if so, why? *Current Directions in Psychological Science*, *29*(3), 248–254.
- McAuliffe, W. H., Forster, D. E., Pedersen, E. J., & McCullough, M. E. (2018). Experience with anonymous interactions reduces intuitive cooperation. *Nature Human Behaviour*, *2*(12), 909–914.
- McAuliffe, W. H., Forster, D. E., Pedersen, E. J., McCullough, M. E., & Baumert, A. (2019). Does cooperation in the laboratory reflect the operation of a broad trait? *European Journal of Personality*, *33*(1), 89–103.
- McClintock, C. G., & Van Avermaet, E. (1982). Social values and rules of fairness: A theoretical perspective. In *Cooperation and helping behavior* (pp. 43–71). Elsevier.

- McNamara, J. M., & Leimar, O. (2010). Variation and the response to variation as a basis for successful cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1553), 2627–2633.
- Miletić, S., Boag, R. J., & Forstmann, B. U. (2020). Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia*, 136, 107261.
- Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, 65, 276–291.
- Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., & Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(42), 15623–15628.
- Morelli, S. A., Rameson, L. T., & Lieberman, M. D. (2014). The neural components of empathy: Predicting daily prosocial behavior. *Social Cognitive and Affective Neuroscience*, 9(1), 39–47.
- 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.
- Murnighan, J. K., & Wang, L. (2016). The social world as an experimental game. *Organizational Behavior and Human Decision Processes*, 136, 80–94.
- Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision making*, 6(8), 771–781.
- Nelson, S. K., Layous, K., Cole, S. W., & Lyubomirsky, S. (2016). Do unto others or treat yourself? The effects of prosocial and self-focused behavior on psychological flourishing. *Emotion*, 16(6), 850–861.
- Neumann, L. J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (Vol. 60). Princeton University Press.
- Noble, S., Scheinost, D., & Constable, R. T. (2021). A guide to the measurement and interpretation of fMRI test–retest reliability. *Current Opinion in Behavioral Sciences*, 40, 27–32.
- Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van't Veer, A. E., & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, 23(10), 815–818.
- Olsson, A., Knapska, E., & Lindström, B. (2020). The neural and computational systems of social learning. *Nature Reviews Neuroscience*, 21(4), 197–212.
- Oswald, P. A. (1996). The effects of cognitive and affective perspective taking on empathic concern and altruistic helping. *The Journal of Social Psychology*, 136(5), 613–623.
- Overgaauw, S., Jansen, M., & de Bruijn, E. R. (2020). Self-centered or other-directed: Neural correlates of performance monitoring are dependent on psychopathic traits and social context. *Cortex*, 129, 199–210.
- Palmeri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, 21(6), 425–433.
- Park, S. Q., Kahnt, T., Dogan, A., Strang, S., Fehr, E., & Tobler, P. N. (2017). A neural link between generosity and happiness. *Nature Communications*, 8(1), 1–10.
- Pärnamets, P., Shuster, A., Reinero, D. A., & Van Bavel, J. J. (2020). A value-based framework for understanding cooperation. *Current Directions in Psychological Science*, 29(3), 227–234. <https://doi.org/10.1177/0963721420906200>
- Pedersen, M. L., & Frank, M. J. (2020). Simultaneous hierarchical Bayesian parameter estimation for reinforcement learning and drift diffusion models: A tutorial and links to neural data. *Computational Brain & Behavior*, 3, 458–471.
- Pereda, M., Brañas-Garza, P., Rodríguez-Lara, I., & Sánchez, A. (2017). The emergence of altruism as a social norm. *Scientific Reports*, 7(1), 1–10.
- Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a ‘cooperative phenotype’ that is domain general and temporally stable. *Nature Communications*, 5, 4939. <https://doi.org/10.1038/ncomms5939>
- Pisor, A. C., Gervais, M. M., Purzycki, B. G., & Ross, C. T. (2020). Preferences and constraints: The value of economic games for studying human behaviour. *Royal Society Open Science*, 7(6), 192090.
- Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, 6(10), 421–425.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59–63.
- Preckel, K., Kanske, P., & Singer, T. (2018). On the interaction of social affect and cognition: Empathy, compassion and theory of mind. *Current Opinion in Behavioral Sciences*, 19, 1–6. <https://doi.org/10.1016/j.cobeha.2017.07.010>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28(1), 363–392.
- Rameson, L. T., Morelli, S. A., & Lieberman, M. D. (2012). The neural correlates of empathy: Experience, automaticity, and prosocial behavior. *Journal of Cognitive Neuroscience*, 24(1), 235–245.
- Rand, D. G., Newman, G. E., & Wurzbacher, O. M. (2015). Social context and the dynamics of cooperative choice. *Journal of Behavioral Decision Making*, 28(2), 159–166.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281.

- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367.
- Reiter, A. M., Diaconescu, A. O., Eppinger, B., & Li, S.-C. (2021). Human aging alters social inference about others' changing intentions. *Neurobiology of Aging*, *103*, 98–108.
- Reiter, A. M., Kanske, P., Eppinger, B., & Li, S.-C. (2017). The aging of the social mind-differential effects on components of social understanding. *Scientific Reports*, *7*(1), 1–8.
- Reuter, M., & Montag, C. (2016). *Neuroeconomics*. Springer.
- Rieger, M., & Mata, R. (2015). On the generality of age differences in social and nonsocial decision making. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *70*(2), 200–212.
- Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, *62*, 23–48.
- Roberts, I. D., & Hutcherson, C. A. (2019). Affect and decision making: Insights and predictions from computational models. *Trends in Cognitive Sciences*, *23*(7), 602–614.
- Rosenthal, I. A., Hutcherson, C. A., Adolphs, R., & Stanley, D. A. (2019). Deconstructing theory-of-mind impairment in high-functioning adults with autism. *Current Biology*, *29*(3), 513–519.e516. <https://doi.org/10.1016/j.cub.2018.12.039>
- Rosi, A., Nola, M., Lecce, S., & Cavallini, E. (2019). Prosocial behavior in aging: Which factors can explain age-related differences in social-economic decision making? *International Psychogeriatrics*, *31*(12), 1747–1757. <https://doi.org/10.1017/S1041610219000061>
- Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562.
- Rusch, T., Steixner-Kumar, S., Doshi, P., Spezio, M., & Gläscher, J. (2020). Theory of mind and decision science: Towards a typology of tasks and computational models. *Neuropsychologia*, *146*, 107488.
- Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, *2*(4), 293–302.
- Ryu, S., Park, J., Kim, K., & Kim, Y.-G. (2020). Reward versus altruistic motivations in reward-based crowdfunding. *International Journal of Electronic Commerce*, *24*(2), 159–183.
- Samuelson, L. (2016). Game theory in economics and beyond. *Journal of Economic Perspectives*, *30*(4), 107–130.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience 1. *Behavioral and Brain Sciences*, *36*(4), 393–414.
- Schubert, A.-L., Frischkorn, G. T., Hagemann, D., & Voss, A. (2016). Trait characteristics of diffusion model parameters. *Journal of Intelligence*, *4*(3), 7.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34.
- Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2020). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*, 293–327. <https://doi.org/10.1037/bul0000303>
- Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. *Human Brain Mapping*, *38*(9), 4788–4805.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.
- Scott, J. (2000). Rational choice theory. In *Understanding contemporary society: Theories of the present* (Vol. 129, pp. 671–685). Sage Publications.
- Serra, D. (2020). *Neuroeconomics: Reliable, scientifically legitimate and useful knowledge for economists?*. Montpellier, France: CEE-M, University of Montpellier.
- Sescousse, G., Caldú, X., Segura, B., & Dreher, J.-C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *37*(4), 681–696.
- Shahar, N., Hauser, T. U., Moutoussis, M., Moran, R., Keramati, M., & Dolan, R. J. (2019). Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Computational Biology*, *15*(2), e1006803.
- Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *eLife*, *9*, e56938.
- Shuster, A., & Levy, D. J. (2020). Contribution of self-and other-regarding motives to (dis) honesty. *Scientific Reports*, *10*(1), 1–11.
- Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, *24*(18), R875–R878.
- Singer, T., & Tusche, A. (2014). Understanding others: Brain mechanisms of theory of mind and empathy. In *Neuroeconomics* (pp. 513–532). Elsevier.
- Smith, D. H. (1981). Altruism, volunteers, and volunteerism. *Journal of Voluntary Action Research*, *10*(1), 21–36.
- Smith, D. V., Gseir, M., Speer, M. E., & Delgado, M. R. (2016). Toward a cumulative science of functional integration: A meta-analysis of psychophysiological interactions. *Human Brain Mapping*, *37*(8), 2904–2917.
- Song, Q. C., Tang, C., & Wee, S. (2021). Making sense of model generalizability: A tutorial on cross-validation in R and Shiny. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920947067.

- Soutschek, A., Ruff, C. C., Strombach, T., Kalenscher, T., & Tobler, P. N. (2016). Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Science Advances*, 2(10), e1600992.
- 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.
- 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 of the United States of America*, 112(25), 7851–7856.
- Sun, B., Luo, Z., Zhang, W., Li, W., & Li, X. (2018). Age-related differences in affective and cognitive empathy: Self-report and performance-based evidence. *Aging, Neuropsychology, and Cognition*, 25(5), 655–672.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- 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.
- Sze, J. A., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (2012). Greater emotional empathy and prosocial behavior in late life. *Emotion*, 12(5), 1129–1140.
- Tankersley, D., Stowe, C. J., & Huettel, S. A. (2007). Altruism is associated with an increased neural response to agency. *Nature Neuroscience*, 10(2), 150–151.
- 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.
- Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin*, 146(1), 30–90.
- Timmers, I., Park, A. L., Fischer, M. D., Kronman, C. A., Heathcote, L. C., Hernandez, J. M., & Simons, L. E. (2018). Is empathy for pain unique in its neural correlates? A meta-analysis of neuroimaging studies of empathy. *Frontiers in Behavioral Neuroscience*, 12, 289.
- Tusche, A., Böckler, A., Kanske, P., Trautwein, F.-M., & Singer, T. (2016). Decoding the charitable brain: Empathy, perspective taking, and attention shifts differentially predict altruistic giving. *The Journal of Neuroscience*, 36(17), 4719–4732. <https://doi.org/10.1523/jneurosci.3392-15.2016>
- Tusche, A., & Hutcherson, C. (2018). Cognitive regulation alters social and dietary choice by changing attribute representations in domain-general and domain-specific brain circuits. *eLife*, 7, e31185.
- Vandekerckhove, J., Matzke, D., & Wagenmakers, E. J. (2015). Model comparison and the principle. *The Oxford handbook of computational and mathematical psychology*, 300–318. USA: Oxford University Press.
- van Dijk, E., & De Dreu, C. K. (2020). Experimental games and social decision making. *Annual Review of Psychology*, 72, 415–438.
- Van Lange, P. A. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, 77(2), 337–349.
- Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, 30(3), 829–858.
- Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, 40(1), 61–72.
- Vanyukov, P. M., Hallquist, M. N., Delgado, M., Szanto, K., & Dombrovski, A. Y. (2019, 2019/08/01). Neurocomputational mechanisms of adaptive learning in social exchanges. *Cognitive, Affective, & Behavioral Neuroscience*, 19(4), 985–997. <https://doi.org/10.3758/s13415-019-00697-0>
- von Krause, M., Radev, S. T., Voss, A., Quintus, M., Egloff, B., & Wrzus, C. (2021). Stability and change in diffusion model parameters over two years. *Journal of Intelligence*, 9(2), 26.
- Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, 6, 336.
- Wagenmakers, E.-J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14(1), 3–22.
- Warrier, V., Grasby, K. L., Uzefovsky, F., Toro, R., Smith, P., Chakrabarti, B., Khadake, J., Mawbey-Adamson, E., Litterman, N., & Hottenga, J.-J. (2018). Genome-wide meta-analysis of cognitive empathy: Heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Molecular Psychiatry*, 23(6), 1402–1409.
- Waytz, A., Zaki, J., & Mitchell, J. P. (2012). Response of dorsomedial prefrontal cortex predicts altruistic behavior. *Journal of Neuroscience*, 32(22), 7646–7650.
- West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2), 415–432.
- Wieck, C., & Kunzmann, U. (2015). Age differences in empathy: Multidirectional and context-dependent. *Psychology and Aging*, 30(2), 407–419.
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in Neuroinformatics*, 7, 14.
- Wiepking, P., & James, R. N. (2013). Why are the oldest old less generous? Explanations for the unexpected age-related drop in charitable giving. *Ageing & Society*, 33(3), 486–510.
- Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, 8, e49547.

- Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A., Inukai, K., Tanida, S., & Kiyonari, T. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes*, *120*(2), 260–271.
- Zaki, J. (2017). Moving beyond stereotypes of empathy. *Trends in Cognitive Sciences*, *21*(2), 59–60.
- Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience*, *15*(6), 695–707.
- Zhao, K., & Smillie, L. D. (2015). The role of interpersonal traits in social decision making: Exploring sources of behavioral heterogeneity in economic games. *Personality and Social Psychology Review*, *19*(3), 277–302.
- Zinchenko, O. (2019). Brain responses to social punishment: A meta-analysis. *Scientific Reports*, *9*(1), 1–8.

**How to cite this article:** Tusche, A., & Bas, L. M. (2021). Neurocomputational models of altruistic decision-making and social motives: Advances, pitfalls, and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science*, *12*(6), e1571. <https://doi.org/10.1002/wcs.1571>