In this chapter, we bridge cognitive neuroscience, neuroeconomics, and social psychology to examine the issue of human prosociality and cooperation. In the first section, we review literature from several fields to describe common experimental tasks used to measure human cooperation. In the second section, we review the dominant theoretical models that have been used to characterize cooperative decision-making, as well as the brain regions implicated in cooperation. Building on work in neuroeconomics, we suggest that a value-based account may provide the most powerful understanding of the psychology and neuroscience of group cooperation. In the third and fourth sections, we review the role of individual differences and social context in shaping the mental processes that underlie cooperation. Finally, we consider gaps in the literature and offer directions for future research on the social neuroscience of cooperation. We suggest that this multilevel approach provides a more comprehensive understanding of the mental and neural processes that underlie the decision to cooperate with others.

Measuring Cooperation

Cooperation involves any action in which one individual incurs a cost in order to benefit others (Rand & Nowak, 2013). These costs and benefits can range from primary reinforcers (e.g., food, drugs, sex) to secondary reinforcers (e.g., wealth, status, fame). Critically, cooperative acts are not always selfless; sometimes we help others at a cost to obtain rewards in the future. For instance, you may be motivated to tip a bartender not only to reward attentive service but to continue receiving excellent service in the future. For this reason, some researchers distinguish between pure or altruistic cooperation (i.e., when current or future rewards are ignored) and strategic cooperation (i.e., when future rewards motivate the cooperative act; Camerer & Fehr, 2004; Gintis, 2014). Cooperative acts can be pure, strategic, or a mixture of both. As a result, researchers go to great lengths to disambiguate these motives (Camerer & Fehr, 2004).
To better understand the motives that underlie cooperation and how they are studied, we briefly review four measures of cooperation.

**Social dilemmas** The most common approach to studying cooperation involves the use of social dilemmas, and perhaps the most widely used measure of cooperation is the prisoner’s dilemma (PD) game. In the PD, two players are each given the choice to either defect (D) or cooperate (C). This game has been popularized on the British game show Golden Balls because it creates a tension in which the fates of two players are tied together. In the standard, symmetric version of the game, both players receive payoff R(eward) if both choose C, payoff P(unishment) if both choose D, and payoffs T(emptation) or S(ucker) if one defects and the other cooperates, respectively. Thus, the hierarchical payoff structure is $T > R > P > S$. As in the legal system, there is a strong temptation not to be a sucker. In the PD each player can maximize individual profit by choosing D, regardless of what the other player chooses. In other words, outcome DD is the unique Nash equilibrium of the game and the prediction for fully rational and selfish players. However, the cooperative outcome, CC, maximizes their collective profit. This feature—that the players are always worse off if both defect compared to cooperate, but each is individually better off by defecting—is what makes the PD a social dilemma (Dawes, 1980). Pitting self-interest against collective interest captures the dynamic at play in countless real-world social decisions, from negotiating nuclear arms agreements to sharing research ideas.

Since decisions are typically made simultaneously, anonymous one-shot PDs (i.e., one round only) are used to measure pure cooperation in both players. In contrast, the iterated PD, in which players play multiple rounds with one another, measures strategic cooperation since players’ decisions may affect expectations for subsequent choices. In addition, people cooperate strategically when their choices are made public, and players can select partners known to be cooperative (Barclay & Willer, 2007; Feinberg, Willer, & Schultz, 2014). Despite understanding that defecting is in one’s best self-interest, decades of evidence from both iterated and one-shot versions of the PD reveal that people willingly cooperate—even with complete strangers.

1 Invented in 1950 by Merrill Flood and Melvin Dresher, while working at the RAND Corporation (no known relation to Dave Rand, who is cited throughout this chapter) as part of the research investigating the use of game theory to inform nuclear strategy.

To understand cooperation in groups with more than two players, researchers employ the public goods game (PGG). In this game, players choose between contributing their endowment to a collective pool (i.e., maximizing joint payoffs) or free riding, in which they keep their own endowment while also reaping the benefits of others’ contributions (i.e., maximizing individual payoffs in the short term). The PGG has a similar incentive structure to the PD and is sometimes suggested to be a generalization of it (Rand & Nowak, 2013). The PGG inherits many properties of the PD (e.g., anonymous one-shot games index pure cooperation) since contributing and free riding are group-based analogs of cooperating and defecting. Similar to the findings in the PD, evidence reveals that in typical variants of the PGG, people donate, on average, in 60% of the trials. However, because the PGG also inherits properties of group psychology, important differences can emerge (Dawes, 1980). For instance, contributions in iterated PGGs routinely diminish over time (Andreoni, 1988), whereas those in the PD do not. This may be due to the diffusion of responsibility or absence of direct reciprocity in the PGG, in which punishing one free rider equally penalizes the entire group. PGGs may also be particularly sensitive to other aspects of group psychology, such as norms concerning promise keeping (Bicchieri, 2002) and social identity (Kramer & Brewer, 1984). Furthermore, the PGG likely provides superior ecological validity to the PD since the most pressing real-world cooperative dilemmas, like climate change or science reform, involve more than two people (Camerer, 2011).

Social dilemmas sometimes include additional dimensions, such as introducing reinforcement or punishment opportunities (Fehr & Gächter, 2002; Kelley, 2003), confronting reputational concerns (Milinski, Semmann, & Krambeck, 2002), or manipulating the framing of the game (Van Lange et al., 2013). For instance, framing a social dilemma as a “community game” can double rates of cooperation compared to framing it as a “Wall Street game,” likely due to activating norms associated with those contexts (Liberman, Samuels, & Ross, 2004). Moreover, introducing opportunities for reward and punishment almost always boosts contributions (Andreoni, Harbaugh, & Vesterlund, 2002; Dreber, Rand, Fudenberg, & Nowak, 2008; Fehr & Gächter, 2002). These factors appear to alter the value people place on the decision to cooperate.

**Bargaining games** Another measure of cooperation comes from bargaining games in which responsiveness...
to fairness norms can be assessed. In the ultimatum game (UG; Güth, Schmittberger, & Schwarze, 1982), two players take the role of either proposer or responder. The proposer is given some endowment \( E \) and must offer the responder some amount \( O \) (which may be zero). The responder can either accept or reject the offer. If the offer is accepted, the responder receives \( O \), and the proposer keeps the remainder (\( E - O \)). If the offer is rejected, neither player receives anything. From an economically rational standpoint, responders should accept any offer since some money is better than no money. However, it has been repeatedly observed across cultures that responders will reject offers that are considered unfair according to local norms (Camerer & Fehr, 2004; Henrich et al., 2005), which is typically anything below 20% of the endowment. By rejecting the offer, people are signaling their willingness to forgo their own profit to punish a transgressor who violated fairness norms—harming both parties. Thus, a degree of cooperation is normally required to ensure a fair offer is accepted.\(^3\)

To capture pure prosociality, a modified UG is used in which the responder is not given the option to reject the proposer’s offer (Kahneman, Knetsch, & Thaler, 1986)—known as the dictator game (DG). In this game, the experimenter endows a sum of money to the dictator, who can then decide how much to give to the receiver. True to its name, the receiver has no bargaining power in the DG and has no choice but to accept the initial offer from the dictator. Surprisingly, dictators nevertheless make offers in these one-sided games, revealing just how altruistic people can be. This is the case even when the experimenter ensures complete anonymity between the two players, providing a measure of pure prosociality for the dictator since there is no opportunity to reciprocate or punish an unfair split. These games provide some evidence for the tendency of humans to cooperate under a wide variety of conditions.

Models of Cooperation

Models of prosocial behavior make assumptions about the underlying mental computations that guide people toward self-interest or cooperation. In the following sections, we contrast three such models of cooperation. The first two are based on a dual-process account that casts intuitive and deliberative processes as competing for control in cooperative behavior. The third offers a single-process framework from neuroeconomics that emphasizes the role of valuation circuits. We briefly review each approach and argue that social and cognitive neuroscience might prove fruitful for arbitrating between these different models.

Intuition versus deliberation One of the most ubiquitous frameworks in psychology is the dual-process model, which posits that the mind can be carved into two core systems: intuition (i.e., fast, automatic, and unconscious processes) and deliberation (i.e., slow, controlled, and rational processes; Chaiken & Trope, 1989; Evans & Stanovich, 2013; Kahneman, 2011). Research in social neuroscience has attempted to map neural systems onto intuition and deliberation (Cohen, 2005; Satpute & Lieberman, 2006). For instance, patients with ventromedial prefrontal cortex (vmPFC) or amygdala damage presented with blunted affective processing (Bechara, 2000), whereas damage to the dorsolateral prefrontal cortex (dLPFC) impaired deliberative processes, like working memory, reasoning, and self-regulation (Barbey, Koenigs, & Grafman, 2013). The dissociations between these systems have been seen by several scholars as further evidence for dual-process models. In psychology, these models have been used to explain a wide range of phenomena, including stereotypes (Devine, 1989), persuasion (Petty & Cacioppo, 1986), and moral judgment (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). More recently, competing dual-process models of cooperation have proven reminiscent of old philosophical debates regarding humanity’s intrinsic benevolence (Rousseau, 1754) versus the need for institutions to restrain our greedy impulses (Hobbes, 1650).

The most prominent dual-process models of cooperation have argued that prosocial decisions stem primarily from intuition (Rand et al., 2014; Zaki & Mitchell, 2013). For instance, the social heuristics hypothesis (Rand et al., 2014) makes three core assumptions: (1) rational self-interested agents should never cooperate in anonymous one-shot games; (2) cooperation stems from errone-prone intuitions, whereas self-interest stems from more corrective deliberation; and (3) experimentally boosting reliance on intuition (vs. deliberation) should only result in increased or static cooperation. In their words, “Deliberation only ever reduces cooperation in social dilemmas ... or has no effect ... but never increases social-dilemma cooperation” (Bear, Kagan, & Rand, 2017). According to this view, cooperation is frequently rational—but people develop error-prone heuristics to cooperate even when it would be irrational.

Support for the social heuristics hypothesis comes from a mix of behavioral and neural evidence. The most important behavioral evidence comes from experiments showing that people are slower to make
self-interested choices compared to cooperative choices in both the one-shot PD and PGG (Everett, Ingbretsen, Cushman, & Cikara, 2017; Rand et al., 2012). Moreover, putting people under time pressure increases cooperation rates (Rand, Greene, & Nowak, 2012). However, a recent international replication effort came up with mixed support for this key finding, suggesting that the behavioral evidence in support of the social heuristics hypothesis may be weaker than previously thought (Bouwmeester et al., 2017; but see also Rand, 2017). Recent functional magnetic resonance imaging (fMRI) studies found that greater dlPFC activation was associated with decisions that prioritize selfish gain over another’s pain (FeldmanHall et al., 2013), while reduced dlPFC functional activity and volume were associated with more generosity in a dictator game, which together suggest a link between deliberation and self-interest (Fernmin et al., 2016; Yamagishi et al., 2016). Those findings are in line with dual-process models in general and the social heuristics hypothesis in particular.

This perspective has proven particularly provocative and controversial because it contrasts with more traditional prosocial restraint models, whereby cooperation primarily stems from the deliberate restraint of our selfish impulses (Achtziger, Alós-Ferrer, & Wagner, 2015; Lohse, 2016; Martinsson, Myrseth, & Wollbrant, 2012). That is, some argue that humans’ unique capacity for self-reflection (i.e., compared to other primates) provides a critical avenue to promote prosocial behavior (Stevens & Hauser, 2004). Moreover, prosocial restraint models are supported by evidence that depleting cognitive resources impairs helping behavior (DeWall et al., 2008) and amplifies dishonesty (Mead, Baumeister, Gino, Schweitzer, & Ariely, 2009; but see Saraiva & Marshall, 2015). We recently found that patients with damage to the dlPFC showed impaired cooperation—and reductions in cooperation scaled with the scope of damage in this region (Wills et al., 2017). We found no such decrements for patients with damage to the vmPFC or the amygdala or other brain-damaged control patients. One limitation of this research area is that several preregistered attempts to replicate ego-depletion effects have found null or very small effect sizes—calling many findings in this literature into question. As such, the evidence behind these models has proven unconvincing to opposing camps.

A value-based approach to cooperation. A central approach to neuroeconomics has examined how value is represented in the human brain and used to guide decision-making. Instead of conceptualizing cooperation as arising from distinct, competing psychological systems, we argue that cooperation, and social preferences in general, should be situated within such a value-based decision framework. Central to this framework is the assumption, found in most economic and psychological theories of choice, that prior to deciding between one or several alternatives an organism determines the subjective value of each alternative. Subjective value allows comparisons between complex and qualitatively different alternatives by placing them on a common scale (Bartra, McGuire, & Kable, 2013; Levy & Glimcher, 2012; Rangel, Camerer, & Montague, 2008). Moreover, this approach allows for individual differences and contextual factors to shape the value of these alternatives. We provide an overview of this perspective, examine the underlying neural system involved in value computations, and describe how this might be fruitfully applied to the study of cooperation.

The field of neuroeconomics has focused on understanding how the brain computes the value of alternative actions during decisions, such as when people are forced to decide between engaging in self-interest or cooperation. The decision-making literature across topics has consistently found that brain activation in the orbitofrontal cortex or vmPFC, ventral striatum (VS), and posterior cingulate cortex increases with subjective value during choice tasks and while receiving monetary, primary, or social rewards (Bartra, McGuire, & Kable, 2013; Levy & Glimcher, 2012). This has been taken as evidence that representations of value are computed in these regions and used as a common currency to decide between different options (Grabenhorst & Rolls, 2011; Levy & Glimcher, 2012).

Recent studies suggest that a value-based framework better explains human cooperation than either dual-process account mentioned above. Prosocial intuition models argue that intuitive responses are shorter than deliberative ones. But from the perspective of value-based frameworks, response times are a function of the discriminability of alternatives: people make faster choices when deciding between very different values as opposed to similar values (Krajbich, Armel, & Rangel, 2010; Shadlen & Kiani, 2013). Thus, these models make competing predictions about cooperation. In one such experiment, participants played multiple PGGs with varying returns on money contributed (Krajbich, Bartling, Hare, & Fehr, 2015). In one condition, for each monetary unit contributed, each player received 50% back. In the other conditions, the multipliers were 30% (rewarding selfishness) and 90% (rewarding cooperation). Consistent with the value-based approach, the

4 Recall that in a PGG a player is always better off keeping the money rather than cooperating. In other words, the multiplier per monetary unit and player is always strictly less than 1.
relationship between reaction time and cooperation was determined by the reward structure: cooperation was fast when it was rewarded, and selfishness was fast when it was rewarded. In other words, cooperation decisions were fastest when the reward structure made the alternatives clear. These findings also highlight why researchers should be cautious when interpreting reaction time differences as evidence for intuition or deliberation.

A growing body of work in cognitive neuroscience also supports the value-based account of cooperation. Specifically, several studies have found that vmPFC activation relates to value-based quantities during cooperative decisions (Feldman-Hall, Dalgeish, Evans, & Mobbs, 2015; Hutcherson, Bushong, & Rangel, 2015; Zaki, Lopez, & Mitchell, 2014). During altruistic decision-making, for instance, the brain forms an overall value signal as a weighted sum of two quantities: the payoffs available for oneself and to a recipient (Hutcherson, Bushong, & Rangel, 2015). Both quantities were associated with activation in vmPFC during people’s choices, supporting the idea that vmPFC encodes the overall value of prosocial choices.

The notion that the vmPFC encodes the subjective value of cooperation is also supported by findings from a neuroimaging study conducted while people engaged in the PGG (Wills, Hackel, & Van Bavel, 2018). We found that vmPFC activity was greater when participants made choices aligned with their overall social preferences (i.e., when cooperative players made the decision to cooperate and selfish players made the decision to act selfishly). In contrast, dlPFC activity was associated with choices that went against players’ social preferences. Moreover, there was increased connectivity between vmPFC and dlPFC when people made cooperative decisions that violated social norms. In these cases the dlPFC may be needed to integrate value signals computed in the vmPFC (Domenech, Redoute, Koechling, & Dreher, 2017), as value-related signals in dlPFC activate after those in vmPFC (Sokol-Hessner, Hutcherson, Hare, & Rangel, 2012). Clarifying the connectivity between regions will likely be key to further arbitrating between the value-based model and competing frameworks (see figure 86.1).

There is growing research into the various psychological factors that modulate (i.e., suppress or amplify) value. After all, when constructing interventions to promote cooperation, it is vital to understand when and for whom cooperation is valued. For instance, interventions designed to block “deliberative self-interest” could fail—or even backfire—among those who do not intrinsically value cooperation and need to deliberate longer to fully consider the potential value of cooperation. Similarly, while efforts to deter “intuitive self-interest” could prevail under some circumstances, these same interventions might also reduce cooperation under contexts in which cooperation is strongly valued. Here we review two broad classes of these potential value modulators: (1) contextual factors and (2) individual differences.

**Contextual Factors**

Several contextual factors can influence cooperative decision-making by shaping social value. For instance, group norms have been known to boost compliance in perceptual judgments (Asch, 1951) and prosocial behavior (Cialdini, Reno, & Kallgren, 1990; Nook, Ong, Morelli, Mitchell, & Zaki, 2016). Evidence for cognitive neuroscience suggests that group norms also modulate relationship between reaction time and cooperation was determined by the reward structure: cooperation was fast when it was rewarded, and selfishness was fast when it was rewarded. In other words, cooperation decisions were fastest when the reward structure made the alternatives clear. These findings also highlight why researchers should be cautious when interpreting reaction time differences as evidence for intuition or deliberation.

A growing body of work in cognitive neuroscience also supports the value-based account of cooperation. Specifically, several studies have found that vmPFC activation relates to value-based quantities during cooperative decisions (Feldman-Hall, Dalgeish, Evans, & Mobbs, 2015; Hutcherson, Bushong, & Rangel, 2015; Zaki, Lopez, & Mitchell, 2014). During altruistic decision-making, for instance, the brain forms an overall value signal as a weighted sum of two quantities: the payoffs available for oneself and to a recipient (Hutcherson, Bushong, & Rangel, 2015). Both quantities were associated with activation in vmPFC during people’s choices, supporting the idea that vmPFC encodes the overall value of prosocial choices.

The notion that the vmPFC encodes the subjective value of cooperation is also supported by findings from a neuroimaging study conducted while people engaged in the PGG (Wills, Hackel, & Van Bavel, 2018). We found that vmPFC activity was greater when participants made choices aligned with their overall social preferences (i.e., when cooperative players made the decision to cooperate and selfish players made the decision to act selfishly). In contrast, dlPFC activity was associated with choices that went against players’ social preferences. Moreover, there was increased connectivity between vmPFC and dlPFC when people made cooperative decisions that violated social norms. In these cases the dlPFC may be needed to integrate value signals computed in the vmPFC (Domenech, Redoute, Koechling, & Dreher, 2017), as value-related signals in dlPFC activate after those in vmPFC (Sokol-Hessner, Hutcherson, Hare, & Rangel, 2012). Clarifying the connectivity between regions will likely be key to further arbitrating between the value-based model and competing frameworks (see figure 86.1).

There is growing research into the various psychological factors that modulate (i.e., suppress or amplify) value. After all, when constructing interventions to promote cooperation, it is vital to understand when and for whom cooperation is valued. For instance, interventions designed to block “deliberative self-interest” could fail—or even backfire—among those who do not intrinsically value cooperation and need to deliberate longer to fully consider the potential value of cooperation. Similarly, while efforts to deter “intuitive self-interest” could prevail under some circumstances, these same interventions might also reduce cooperation under contexts in which cooperation is strongly valued. Here we review two broad classes of these potential value modulators: (1) contextual factors and (2) individual differences.

**Contextual Factors**

Several contextual factors can influence cooperative decision-making by shaping social value. For instance, group norms have been known to boost compliance in perceptual judgments (Asch, 1951) and prosocial behavior (Cialdini, Reno, & Kallgren, 1990; Nook, Ong, Morelli, Mitchell, & Zaki, 2016). Evidence for cognitive neuroscience suggests that group norms also modulate relationship between reaction time and cooperation was determined by the reward structure: cooperation was fast when it was rewarded, and selfishness was fast when it was rewarded. In other words, cooperation decisions were fastest when the reward structure made the alternatives clear. These findings also highlight why researchers should be cautious when interpreting reaction time differences as evidence for intuition or deliberation.

A growing body of work in cognitive neuroscience also supports the value-based account of cooperation. Specifically, several studies have found that vmPFC activation relates to value-based quantities during cooperative decisions (Feldman-Hall, Dalgeish, Evans, & Mobbs, 2015; Hutcherson, Bushong, & Rangel, 2015; Zaki, Lopez, & Mitchell, 2014). During altruistic decision-making, for instance, the brain forms an overall value signal as a weighted sum of two quantities: the payoffs available for oneself and to a recipient (Hutcherson, Bushong, & Rangel, 2015). Both quantities were associated with activation in vmPFC during people’s choices, supporting the idea that vmPFC encodes the overall value of prosocial choices.

The notion that the vmPFC encodes the subjective value of cooperation is also supported by findings from a neuroimaging study conducted while people engaged in the PGG (Wills, Hackel, & Van Bavel, 2018). We found that vmPFC activity was greater when participants made choices aligned with their overall social preferences (i.e., when cooperative players made the decision to cooperate and selfish players made the decision to act selfishly). In contrast, dlPFC activity was associated with choices that went against players’ social preferences. Moreover, there was increased connectivity between vmPFC and dlPFC when people made cooperative decisions that violated social norms. In these cases the dlPFC may be needed to integrate value signals computed in the vmPFC (Domenech, Redoute, Koechling, & Dreher, 2017), as value-related signals in dlPFC activate after those in vmPFC (Sokol-Hessner, Hutcherson, Hare, & Rangel, 2012). Clarifying the connectivity between regions will likely be key to further arbitrating between the value-based model and competing frameworks (see figure 86.1).

There is growing research into the various psychological factors that modulate (i.e., suppress or amplify) value. After all, when constructing interventions to promote cooperation, it is vital to understand when and for whom cooperation is valued. For instance, interventions designed to block “deliberative self-interest” could fail—or even backfire—among those who do not intrinsically value cooperation and need to deliberate longer to fully consider the potential value of cooperation. Similarly, while efforts to deter “intuitive self-interest” could prevail under some circumstances, these same interventions might also reduce cooperation under contexts in which cooperation is strongly valued. Here we review two broad classes of these potential value modulators: (1) contextual factors and (2) individual differences.

**Contextual Factors**

Several contextual factors can influence cooperative decision-making by shaping social value. For instance, group norms have been known to boost compliance in perceptual judgments (Asch, 1951) and prosocial behavior (Cialdini, Reno, & Kallgren, 1990; Nook, Ong, Morelli, Mitchell, & Zaki, 2016). Evidence for cognitive neuroscience suggests that group norms also modulate relationship between reaction time and cooperation was determined by the reward structure: cooperation was fast when it was rewarded, and selfishness was fast when it was rewarded. In other words, cooperation decisions were fastest when the reward structure made the alternatives clear. These findings also highlight why researchers should be cautious when interpreting reaction time differences as evidence for intuition or deliberation.

A growing body of work in cognitive neuroscience also supports the value-based account of cooperation. Specifically, several studies have found that vmPFC activation relates to value-based quantities during cooperative decisions (Feldman-Hall, Dalgeish, Evans, & Mobbs, 2015; Hutcherson, Bushong, & Rangel, 2015; Zaki, Lopez, & Mitchell, 2014). During altruistic decision-making, for instance, the brain forms an overall value signal as a weighted sum of two quantities: the payoffs available for oneself and to a recipient (Hutcherson, Bushong, & Rangel, 2015). Both quantities were associated with activation in vmPFC during people’s choices, supporting the idea that vmPFC encodes the overall value of prosocial choices.

The notion that the vmPFC encodes the subjective value of cooperation is also supported by findings from a neuroimaging study conducted while people engaged in the PGG (Wills, Hackel, & Van Bavel, 2018). We found that vmPFC activity was greater when participants made choices aligned with their overall social preferences (i.e., when cooperative players made the decision to cooperate and selfish players made the decision to act selfishly). In contrast, dlPFC activity was associated with choices that went against players’ social preferences. Moreover, there was increased connectivity between vmPFC and dlPFC when people made cooperative decisions that
the neural substrates of subjective value (Nook & Zaki, 2015; Wills, Hackel, & Van Bavel, 2018), as well as systems implicated in conflict monitoring (Chang & Sanfey, 2013) and control (Knoch, Pasqual-Leone, Meyer, Trayer, & Fehr, 2006; Richeson et al., 2005). For example, disrupting the dlPFC has been shown to disrupt participants’ ability to act in accordance with fairness norms and reject unfair offers in ultimatum games (Knoch et al., 2006). Notably, participants still reported accurate valuations of the offers, suggesting a role of the dlPFC in integrating the outputs from valuation circuits.

Social psychologists distinguish between descriptive norms (i.e., how do others typically behave?) and injunctive norms (i.e., how should others behave?). Since there is strong evidence that descriptive norms influence cooperation (Kopelman, Weber, & Messick, 2002), the same is likely true for injunctive norms—especially since cooperation is often characterized as a moral imperative. Consider, for instance, an influential finding in which framing the PGG as “the community game” boosts cooperation significantly more than when it is called the “Wall Street game” (Liberman, Samuels, & Ross, 2004). Even when other players were expected to be selfish, those assigned to the community condition decided to cooperate nonetheless, suggesting that injunctive norms can bias moral behavior.

Social identity—a person’s sense of who they are based on their group membership—is another core social psychological construct that drives cooperation and conflict (Tajfel & Turner, 2001). For instance, cooperative decisions can be influenced by existing intergroup conflicts, such as race relations (Kubota, Li, Bar-David, Banaji, & Phelps, 2013) and political partisanship (Iyengar & Westwood, 2015), as well as by artificially created identities (Marcus-Newhall, Miller, Holtz, & Brewer, 1995). Social identity may drive cooperation because it connotes interdependence: people assume in-group members will reciprocate with one another (Yamigishi, 1992). There is also reason to believe that identity can change the value people place on in-group members and their outcomes. For example, one study found greater activation in the ventral striatum when participants observed in-group members receive rewards compared to out-group members, but only for participants who heavily identified with the in-group (Hackel, Zaki, & Van Bavel, 2017). Indeed, simply categorizing faces of in-group members activates the neural circuitry associated with valuation, including the amygdala, orbitofrontal cortex, and dorsal striatum (Van Bavel, Packer, & Cunningham, 2008). Thus, generating a shared group identity can induce cooperation by imbuing in-group members with value or increasing the expectations of future reward due to reciprocity.

**Individual Differences**

People differ in their tendency to cooperate, and these preferences tend to be stable over time (Volk, Thöni, & Ruigrok, 2012). Within PGGs, for instance, researchers have estimated that a substantial majority of people (50%–55%) are conditional cooperators (i.e., those who only cooperate when others cooperate), a sizable portion (23%–30%) are considered consistent free riders (Fischbacher, Gächter, & Fehr, 2001), and only a small percentage (5%–10%) fall into the category of consistent contributors who always cooperate (Weber & Murnighan, 2008). Some measures, such as the Social Value Orientation measure, are designed to capture these differences (see Van Lange, 1999). **Proselfs** are people who place a high value on their own rewards, whereas **prosicals** are people who place a high value on collective rewards. Research in the past decade has consistently found that prosicals are more inclined to cooperate in both one-shot and iterated games (Balliet, Parks, & Joireman, 2009). Thus, individual differences are robust predictors of cooperative (vs. selfish) behavior.

Critically, individual differences may determine which contextual factors steer cooperative decision-making. Take, for instance, consistent contributors, who are defined by their iconoclastic commitment to cooperating under any circumstance (i.e., even when everyone else in their group is free riding). There is evidence that the mere presence of these consistent contributors can boost cooperation in others by activating moral identities (Gill, Packer, & Van Bavel, 2013). That is, consistent contributors may provide a contextual cue that predominately boosts cooperation among individuals who consider generosity and fairness to be central features of their identity (Packer, Gill, Chu, & Van Bavel, 2018). In addition, there is evidence that experimentally invoking deliberation promotes cooperation, but only for people exhibiting prosocial tendencies (Mischkowski & Glöckner, 2015). Thus, individual differences can also predict which contextual factors are more or less likely to shape cooperative decision-making. More work should examine this interplay using neuroscientific methods to better understand how individual differences and context are integrated in the brain during decision-making.

**Future Directions**

**Attention** A key element of dynamic value-based cognition is the role of attention. By measuring participants’ fixations during simple economic choices, researchers have shown that attention to certain options influences decisions (Krajibich, Armel, & Rangel, 2010). These findings have been shown to also hold for more complicated...
value-based choices, such as those that are moral (Páraments, Balkenius, & Richardson, 2014). By tracking participants’ fixations and prompting them to make a choice only after sufficiently fixating on one option, researchers were even able to influence what choice participants made (Páraments et al., 2015). Moreover, one study found that value signals in the striatum and vmPFC were modulated by the relative value of fixated versus nonfixated food options (Lim, O’Doherty, & Rangel, 2011). Thus, visual attention influences valuation and alters prosocial behavior. In our view, integrating measures of attention and other sensory information into models of cooperative decision-making offers significant opportunities for understanding more about the underlying mental processes and potentially even designing effective interventions for increasing cooperation.

Learning A key element of value-based models is that people learn the value of different actions over time, whether through personal experience (FeldmanHall, Otto, & Phelps, 2018) or social observation (Haaker et al., 2017; Lindström, Haacker, & Olsson, 2018). Understanding this process may offer new insights into how people choose to cooperate.

Canonical models of reciprocity suggest that people form impressions of others’ generosity and tend to help those viewed as generous (Wedekind & Milinski, 2000). However, models of value learning in neuroscience suggest another route by which people may learn to cooperate with others. During cooperative interactions, people experience reward value—that is, the material benefits of the interaction. When receiving money from an interaction partner, people engage not only the neural regions associated with forming social impressions but also the neural regions associated with reward learning (e.g., the ventral striatum; Hackel, Doll, & Amodio, 2015). As a result, people learn to reciprocate not only with givers who frequently display generosity but also with givers who have greater wealth and thus provide larger rewards (Hackel & Zaki, 2018). Modeling how experience and feedback is integrated into value to guide future decisions is key to fully understanding cooperation. Although the evidence is currently sparse, value learning likely plays a similar role in shaping whether people contribute to collective goods in social dilemmas.

Conclusion Unlocking the secret to group cooperation is critical for solving social dilemmas ranging from climate change to public resource management to improving science. For this reason, the study of cooperation has attracted an enormous amount of attention in recent years. We believe that a value-based approach holds significant promise for understanding how different people in different contexts make cooperative decisions. This approach not only has an explanatory power that can generate important directions in learning and attention but offers to bridge a number of literatures under a common multilevel framework. This has important implications since models consistent with neural architecture should be privileged over models that are not biologically described, and theories that provide consistent evidence across multiple levels of analysis are most likely to provide a complete and enduring explanation of behavior (Wilson, 1998). If this approach can harness the collective intelligence of scientists and scholars from philosophy to neuroscience, it will allow them to cooperate on solving a long-standing scientific debate as well as some of the most pressing problems facing humanity.

Acknowledgments

This chapter was partially funded by a grant from the National Science Foundation to Jay J. Van Bavel (award #1349089) and from the Swedish Research Council to Philip Páraments (2016-06793).

REFERENCES


—0

—+1

WILLS ET AL.: THE SOCIAL NEUROSCIENCE OF COOPERATION 987
A function of context or difficulty? Journal of Neuroscience, 35, 5087–5088.