Bias and Control in Social Decision-Making

Paloma Díaz-Gutiérrez, Sonia Alguacil, and María Ruz

Abstract Social decisions are crucial in our life. Many of these include interactions between agents in scenarios of varying complexity, where trust and cooperation are essential and multiple sources of information influence our choices. In this chapter we review the contributions from social neuroscience to understanding the sources of bias and control mechanisms in social decisions, integrating insights from diverse methodologies and analyses. These biases include individual influences (both stable and transient) and other stimulus-driven factors, such as social stereotypes, emotion displays, or information regarding personality traits. This information modulates different stages of processing, with control-related influences playing crucial roles to override conflicts between automatic tendencies and goals.

Keywords Social neuroscience • Decision-making • Social bias • Control mechanisms • Neuroimaging

1 Introduction

Decisions of different complexity are a constant element in our life. Both simple and more thoughtful and relevant choices share the need of processing different options to choose the action that best fulfills our goals [1]. As social beings, a large part of our decisions involves other people, so that we must take into account information about others and predict their likely behavior. Accordingly, trust and cooperation are central factors in social interactions [2–4]. However, our supposedly rational decisions are fairly influenced by several factors, or biases, which generate predispositions to behave in certain ways [5–7]. The evidence to date shows that these biases are entrained not only with late decision stages related to value or response processing [8] but also with early stages of perception [9]. In addition, the

P. Díaz-Gutiérrez • S. Alguacil • M. Ruz (✉)
Mind, Brain and Behavior Research Center, University of Granada, 18071 Granada, Spain

Department of Experimental Psychology, University of Granada, 18071 Granada, Spain
e-mail: pdiaz@ugr.es; salguacil@ugr.es; mruz@ugr.es

need to arbitrate among these different and complex action tendencies to make optimal decisions calls for strategic control mechanisms.

Several disciplines, such as psychology or economics, seek to understand the role of these biases on social decision-making, and the way control mechanisms are recruited to channel their influence. In this respect, social neuroscience is an innovative discipline that addresses such questions by studying the neural underpinnings of relevant phenomena, focusing on where, when, and how they take place in the brain [10, 11]. The goal of the current chapter is to provide a comprehensive overview of such contributions, integrating insights from diverse methodologies and analysis strategies [8, 12–14].

In the following sections, we first describe the methodology employed in social neuroscience to study the factors that influence social decisions. Then, we present evidence about the different sources of bias in these scenarios, which derive from individual factors and from the stimuli we perceive. Thereupon, we review how these influences are regulated by control mechanisms. Lastly, we offer some conclusions and future directions.

2   Methodological Tools

Research in social neuroscience combines various behavioral methods with modern neuroimaging techniques [15]. On the one hand, several studies rely on the use of interactive games from the field of experimental economics and classic game theory. These paradigms have been often used to derive normative descriptions of how people make economic and trust decisions while interacting with others [16]. However, the reasons for such normative behaviors can be better understood if we know their underlying sources. In this sense, the mechanisms underlying the departures from rationality that people often display in these settings can be explored at the behavioral level by paradigms developed in the field of psychology and at the neural level by modern noninvasive neuroimaging tools. Hence, this mixture of approaches promotes the explanation of human behavior at normative, mechanistic, and neural levels, which complement and nurture each other [17].

Among the tools developed in behavioral game theory, the *ultimatum game* [18] is a very popular task to study the response of people to fairness. Here one player acts as the proposer, choosing how to divide a certain amount of money. The other player, the responder, decides whether to accept or reject the offer. In the first case, both players earn their split, whereas if the responder rejects the offer, neither of them gains anything from that interaction. Reciprocation behavior has been extensively studied with the *trust game* [16]. In this case, one player (the investor) decides whether or not to share an amount of money with another partner (the trustee). If shared, this money is multiplied and transferred to the trustee, who then gets to decide whether to reciprocate or not. In the first case, both earn half of the total money, but if there is no reciprocation, the investor loses the initial sum. In this scenario, the best strategy rests with the mutual cooperation between players. The

prisoner's dilemma [19] is similar, but here both players choose to trust the other one or not, and payoffs depend on both decisions. In addition, some studies have developed online versions of these tasks [20], whereas others have tried to implement cooperation settings in more realistic scenarios (e.g., the *apple game*; [21]).

Several paradigms developed in different fields of psychology are designed to study the mechanisms, or processes, underlying human behavior and choices. For example, the field of social psychology has developed several tasks to explore implicit biases, such as prejudice [22]. Among these, the *implicit association test* (IAT; [23]) is frequently used to explore how people associate social dimensions (e.g., gender, race) with different attributes (e.g., women are emotional vs. men are logical), which ultimately reflect automatic manifestations of prejudice. Similarly, implicit prejudices are often revealed in *sequential evaluative priming tasks*, where, for instance, participants view targets preceded by prime stimuli referring to social categories (e.g., white and black faces) and classify them as "pleasant" or "unpleasant." A variant is the *weapon identification task*, which assesses racial prejudice by asking participants to categorize guns and tools after the presentation of white and black face primes [24]. In addition, other studies use words or facial displays to assess how people form first impressions (e.g., [25]) or associate different social categories depending on their shared stereotypes (e.g., [22]). In addition, moral dilemmas [26, 27], where people have to judge the moral acceptability of behaviors in complex scenarios, are used to explore how personal dispositions or induced analytical tendencies influence moral evaluations.

These behavioral paradigms offer an integrated knowledge of the different phenomena influencing our social choices at different stages of processing. Social neuroscience adds neuroimaging methods to study the neural underpinnings of these decisions. This provides a better understanding of the sources of type of information relevant for social behavior and allows analyzing the commonalities and differences between social and nonsocial phenomena [15]. Among these neuroimaging techniques, electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) are the ones most frequently used to study brain activity noninvasively.

EEG, given its temporal precision, allows tracking how different cognitive processes operate in time [28]. This technique provides information about the stages of processing (e.g., perception, decision, or motor output) at which the phenomena of interest take place. Complementarily, the good spatial resolution of magnetic resonance imaging (MRI) makes it an optimal choice to explore the neural regions underlying all these processes. Additionally, functional near-infrared spectroscopy (fNIRS) measures hemodynamic activity as functional MRI (fMRI), but facilitates more natural experimental settings as it is a portable device, at the expense of lower spatial resolution compared with fMRI [29].

These methods are combined with different analytical strategies, which integrate traditional univariate with multivariate approaches adopted from machine learning. While classic univariate methods compare activation between experimental conditions for each voxel (unit of measurement in MRI), multi-voxel pattern analyses (MVPA; [30]) allow studying how information is encoded in patterns of neural activity across several voxels. Furthermore, representational similarity analysis [31]

relates the structure of neural patterns with each other and also with behavioral data, offering information about the nature of representations in different brain regions and their relation to different psychological theories (e.g., [32]).

Altogether, these new approaches open new avenues to further the understanding of how biasing social information is coded in the brain and the reason for their pervasive effects in our interpersonal behavior.

3 Bias in Social Decision-Making

Influences on social decision-making stem from different sources. On the one hand, individual factors or states impact how we process information, which can bias our decisions. On the other hand, the perception of certain features in other people may also be associated with different action tendencies, judgments, or attributes, impacting how we perceive and behave toward others.

3.1 Bias in the Observer

The individual factors that influence choices include stable personal characteristics (such as gender, age, prosociality, or permanent brain lesions) and contextual, nonpermanent factors (such as induced emotional states). Below we address them in turn.

Beginning with stable factors, gender has been linked to differences in social decision-making in several studies. For example, women seem to make more ethical decisions in certain social scenarios (e.g., [33]). However, altruistic behavior for each gender seems to depend on the expensiveness of the cooperation, which generates different contexts for each of them. Thus, women are more altruistic when it is most costly, whereas the opposite happens for men [34]. Gender differences in moral decisions may also be modulated by emotional empathy [35]. In this case, gender seems to influence our empathic responses to noncooperative partners. For instance, Singer et al. [36] observed that empathic responses to the pain of others, as measured in fronto-insular regions and the anterior cingulate cortex (ACC; see Fig. 1 for a visualization of the brain areas), were reduced in males when observing unfair players receiving painful electric shocks. The brain of male participants in the same situation also showed increased activation in regions related to the reward system, such as the ventral striatum, which was interpreted by the authors as a sense of "revenge." Note that this study is one of many examples of how the introduction of measurements of brain activity adds evidence that helps to understand the mechanisms underlying biases in human social behavior.

Age is another factor that has been related to differences in social decisions. At a young age, children's cooperative behavior is already dependent on the agent they are interacting with, as they are more generous with friends than non-friends and
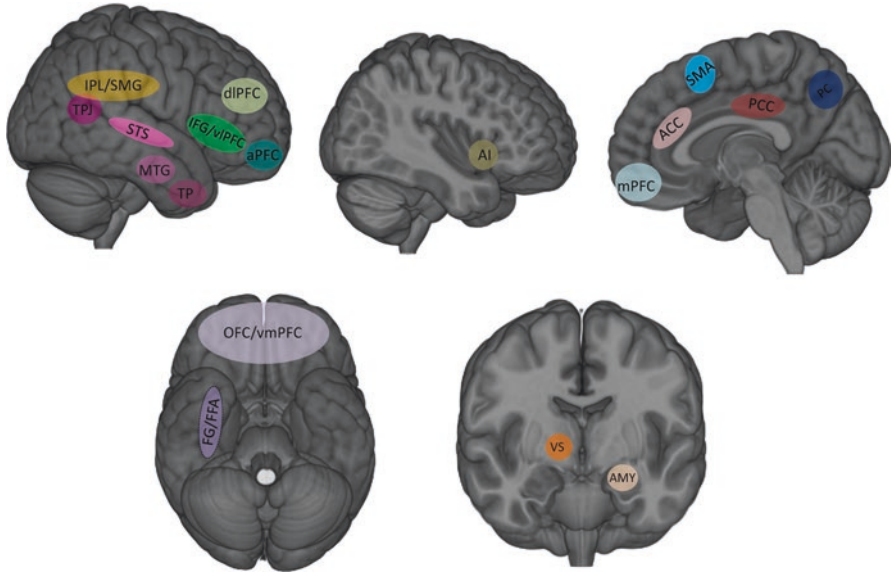
Fig. 1 Display of approximate location of the brain areas mentioned throughout the chapter. *IPL/SMG* inferior parietal lobe/supramarginal gyrus, *TPJ* temporoparietal junction, *STS* superior temporal sulcus, *MTG* middle temporal gyrus, *TP* temporal pole, *dlPFC* dorsolateral prefrontal cortex, *IFG/vlPFC* inferior frontal gyrus/ventrolateral prefrontal cortex, *aPFC* anterior prefrontal cortex, *AI* anterior insula, *mPFC* medial prefrontal cortex, *ACC* anterior cingulate cortex, *SMA* supplementary motor area, *PCC* posterior cingulate cortex, *PC* precuneus, *FG/FFA* fusiform gyrus/fusiform face area, *OFC/vmPFC* orbitofrontal cortex/ventromedial prefrontal cortex, *VS* ventral striatum, *AMY* amygdala

show cooperative tendencies toward strangers when there is no high cost involved [37]. As age increases, children attribute more positive feelings to cooperating with other children [38]. On the other hand, adults seem to be more generous than younger people in economic decisions [39]. Similarly, Rosen et al. [35] also observed that adults made more moral choices than younger participants, but this effect was again mediated by empathy, as the gender case presented above. In addition, Harlé and Sanfey [40] showed that older people appear to be more sensitive to unfairness, with higher rejection rates to unfair offers than younger participants. This unfairness effect was related to activation increases in dorsolateral prefrontal cortex (dlPFC) and decreased in the anterior insula (AI) for adults, compared to young participants. This pattern suggests higher reliance on goal maintenance and less emotional processing due to norm violation with age. However, these effects do not seem to be consistent, as manifested by Lim and Yu [41], who reviewed related literature and observed that the existing evidence proves heterogeneous and does not offer certainty about age differences in prosocial behavior.

Furthermore, individual social preferences or personal concerns for other people, such as altruism, envy, fairness, reciprocity, or inequity aversion, are another source of influence in decisions (e.g., [42]). Individual preferences have also been studied

under the name of social value orientation (SVO; [43, 44]), via different tools—e.g., *decomposed games*, the *ring measure*, *social orientation choice cards*, or characteristics space theory [45]. Within the predominant SVO framework, several studies have tried to distinguish between self-oriented ("proself") and other-oriented ("prosocial") participants and how these individual differences affect cooperation tendencies. While proself subjects show increased calculating and strategizing tendencies, prosocials tend to follow social norms and have moral considerations for others, making more cooperative choices [46, 47]. Also, at a neural level, the brain of prosocials has been linked to increased activation in the precuneus, superior temporal sulcus, and medial prefrontal cortex (mPFC), showing also that this pattern correlates with increased cooperation decisions [48].

Social biases also appear in neuropsychological conditions involving damage in brain areas related to social processing. For instance, temporal lobe epilepsy patients exhibit social functioning deficits [49] in, for example, basic and complex theory of mind processes, which have an impact on social decisions. Amygdala-damaged patients display higher cooperation rates, especially when interacting with untrustworthy partners [50]. This pattern could reflect a deficit in the integration of different social signals that takes place in the amygdala, which would disable proper indications to guide successful social interactions. Moreover, utilitarian judgments in moral dilemmas increase in patients with ventromedial prefrontal cortex (vmPFC) lesions, which has been taken as evidence for the role of this region in the representation of the emotional value of stimuli [51, 52]. Frontotemporal dementia patients also show altered social decisions, with increased impulsiveness and risky behavior, which could be partly related to damage in the orbitofrontal cortex (OFC; [53]). Additionally, these patients make more utilitarian choices in moral dilemmas, which seems to be related to theory of mind deficits [54]. For example, during social bargaining, they manifest altered prosociality and punishment behavior, due to a failure to incorporate information about the perspective of others [55].

Apart from individual factors, a large part of the literature on biases employs experimental settings to induce transient mental states in the agent. A cornerstone source of influences on decisions is the framing effect, which refers to how decisions are affected by the way the scenario is presented [56]. For instance, working with moral dilemmas, De Martino et al. [57] showed that when the problem was framed in a "gain" context, participants tended to choose the safe option, whereas in a "loss" situation they chose the risky alternative to a higher extent. In similar scenarios, positive framing in moral dilemmas has been associated with risk aversion choices, accompanied by increased activation in a cluster involving the ACC and the vmPFC compared to negative framing [58]. Conversely, risk-seeking behavior under negative framing of social cues has been related to activation in the inferior frontal gyrus (IFG; [59]).

Furthermore, the induction of mood states is also a tool frequently used to explore how incidental emotions bias our choices. The affect infusion model [60] claims that incidental emotions prime mood-congruent dispositions, positing that behavioral effects in decision-making tasks depend on the participants' mood [61]. For instance, positive moods prime positive information and have been related to growing confidence, friendliness, and cooperative tendencies during interpersonal

interactions. In this context, positive moods lead to a greater joint gain seeking, interpersonal trust behavior, and cooperative choices [62–64] and also generate an increased preference for avoiding losses (e.g., [65]). Similarly, social reward can serve as another bias in cooperative behavior, as people tend to act more generously when they know they are being watched. When feeling observed, people want to be socially acknowledged about their behavior, which itself constitutes a larger social reinforcement associated with greater activation in the ventral striatum [4].

Conversely, negative moods can have different effects. In economic games, whereas sad affection has been associated with generous behavior [66], it has also been related to a decrement of acceptance of unfair offers, which could be the reflection of a mood-congruent framing for negative outcomes [61]. At a neural level, this bias has been related to increased activity in the bilateral AI, which was thought to mediate between mood and choices. It was also accompanied by higher activity in the dorsal ACC (dACC) for unfair offers, indicating a possible affective conflict. According to the affect infusion model, negative moods would induce a sensitive disposition to detect social violations. This negative mood appears to be coupled with lower activity in the reward system (e.g., ventral striatum) to fair offers. Additionally, "harm to save" dilemmas tend to induce negative emotions such as fear or disgust, each of these biasing participants' response toward different responses. When participants experience fear, they show deontological bias (do nothing), while disgust seems to enhance utilitarian responses (e.g., kill one in order to save five; [67]). Moreover, the application of emotion regulation strategies can also modulate behavioral and neural responses during social decisions. This regulation has been associated with the involvement of the IFG, temporoparietal junction (TPJ), and the AI [68]. The implementation of downregulation (a more positive interpretation) entails higher acceptance rates for unfair offers, while upregulation (a more negative interpretation) elicits more rejection decisions [69].

As we have just described, a variety of individual factors bias people's choices in social decisions. Nonetheless, external factors, mostly originated from the agents we interact with, also exert varying degrees of influence on our choices, as shown in the next section.

3.2 Bias in the Stimuli

Biases in social decision-making also stem from different features of the stimuli we perceive. These choices frequently involve perception and social categorization, as well as the generation of expectations. Faces often provide rich information in these contexts, such as the gender, social group, emotion, and trustworthiness of the people we interact with. This information is highly valuable to generate expectations about others to guide successful decisions. Below, we will first focus on the mechanisms underlying social judgments about other people and then examine how emotion displays and personal information bias choices.

3.2.1 Social Categorization

When we first interact with others, we tend to form impressions about how they are, what they like, or how we expect them to behave, which is a case of social categorization. We form an initial idea of others very quickly, based on the information we can gather in a few milliseconds [70]. These rapid impressions have been related to activity in the posterior cingulate cortex (PCC) and the amygdala [25], both involved in social cognition. The amygdala has also been studied in connection with other regions in terms of the context-relevant representation of social stimuli, especially faces. Its ventrolateral region belongs to a network specialized in social perception [71], in connection with sensory regions of the temporal lobe – the superior temporal sulcus, the temporal pole, the fusiform gyrus, and the OFC.

Categorization judgments are closely related to stereotypes and expectations [72, 73]. Some of such stereotypes refer to biases related to gender, as people tend to assign attributes and internal dispositions differently to women and men. Regions related to evaluative processes and representation of knowledge [74], such as the vmPFC and the amygdala, together with the supramarginal gyrus and the middle temporal gyrus, seem to be at the basis of these judgments. Additionally, contextual influences on face categorization appear mediated by retrosplenial and prefrontal cortices [75].

Furthermore, some biases relate to racial stereotypes. Traditionally, the amygdala has been set as a racial prejudice marker [76], showing higher activation in participants facing a member of a racial outgroup. This involvement has been explained appealing to different roles: activity in this structure could act as a marker for a threat of an outgroup, as an indicator of fear of being considered prejudicial, or as a motivational response [24]. It has also been suggested that the amygdala may be in charge of the representation of relevant social information, while the striatum, which participates in the computation of valence, would represent these stereotypes to guide decisions toward positive interactions and trust behavior with the racial outgroup [77]. Moreover, the AI has been related to negative reactions to a disliked racial outgroup when it has been rewarded. However, this region has also been linked to empathy toward the ingroup [24]. On the other hand, neural representations in the OFC seem to underlie affect-based judgments depending on race, while neural patterns in the anterior portion of the PFC (aPFC) differentiate stereotype-based judgments [78].

At a perceptual level, race influences visual face processing and attention at early stages [79–81]. In this regard, Tortosa et al. [81] observed larger amplitude in the N170 during the processing of black versus white faces, a negative potential related to face encoding [82], which seems related to implicit racial bias [83]. The variations in this potential seem to be originated in early visual processing in the fusiform gyrus [84]. In addition, different studies have also reported varying neural patterns in the fusiform face area depending on the race of faces (e.g., [85]) and how these differences may rely on implicit racial bias [86].

Race bias additionally acts at the decision point. For example, some reports show higher punishment to members of one's own racial group, because they, unlike

outsiders, are expected to cooperate [87]. Moreover, others (e.g., [88]) have observed that participants offer more money and show increased trust toward white versus black partners. However, Tortosa et al. [81] observed that Caucasian participants cooperated more with black than white partners while presenting implicit race bias, which may be explained by participants' desire to counteract their implicit biases.

Interestingly, Stolier and Freeman [22] have recently shown how different social categories are entangled with each other, in the sense that one category activates stereotypes shared with another. Even more, employing novel representational similarity analyses, the authors suggested that the stereotypes related to different categories represented in the OFC modulate activity in earlier visual processing areas of the fusiform gyrus. This results in a greater perceptual similarity between representations of faces sharing the same stereotypes, even if they are of different gender or race. According to the dynamic interactive model [9, 89], social perception is highly dynamic, based on an interactive system in which bottom-up perceptual information activates categorization, which in turn activates stereotypes. Additionally, top-down factors, such as expectations or goals, can modulate lower processing stages in a dynamic fashion.

3.2.2 Emotional Expressions and Trustworthiness

A large part of judgments about others is related to the emotions we perceive in them. Emotional expressions are rapidly processed, even in the absence of awareness (e.g., [90]). In this way, emotional displays have a significant effect on trustworthiness judgments (e.g., [91]), friendliness, or dominance [92], given that they provide information that can be used to decode the intentions, beliefs, and desires of others in social scenarios [93].

Positive expressions tend to induce trust and cooperation [94, 95], whereas negative emotions are associated with uncooperative behavior [16]. However, these emotional expressions may not have the same interpretation in all contexts [93, 96, 97]. For instance, de Melo et al. [98] found that, after mutual cooperation, happiness increased cooperation expectations, whereas in noncooperative scenarios, smiles decreased such expectations. Alternatively, when partners defected, their positive expressions could be considered redundant to their behavior, thus not affecting cooperation expectations. Conversely, when people consider their partner's emotions, anger expressions can induce cooperative decisions (e.g., [99]).

These biases are not only reflected in the type of decision participants make but also in the time they need to make up their minds. Some studies have found "emotional conflict" effects, where participants take longer to choose an option contrary to the automatic response elicited by ignored and non-predictive emotions. For instance, responses in a trust game were slower when emotion and identity information did not lead to the same responses, even when participants were told to ignore these emotions [100]. Moreover, responses are also slowed down when emotions predict consequences opposite to their "natural consequences" [101, 102]. In this scenario, when emotional expressions are predictive of their natural consequences, activity

increases in the precuneus [101], a region associated with the representation of personal information [103] and trust in cooperative scenarios [104].

Furthermore, facial expressions seem to be associated with trustworthiness judgments along a continuum, where untrustworthy faces are linked to anger expressions, whereas trustworthiness is related to happiness [92]. These trust judgments correlate with amygdala activity, as this region presents a higher response to untrustworthy agents [105]. Interestingly, such behavioral and neural sensitivity to trustworthiness may occur even with no perceptual awareness [106, 107]. In this regard, several studies have shown that trustworthiness can indeed impact our decisions in different ways. During trust and economic games, people manifest higher cooperation rates and acceptance of offers from trustworthy agents [108, 109]. People invest more money with partners who have been rated as trustworthy even when there is no objective relationship between ratings and actual behavior [110]. Moreover, rejection of offers based on facial trustworthiness correlates with activity in the OFC, and its functional connectivity with the AI correlates with individual rejection decisions from untrustworthy partners [108].

3.2.3   Personal Information

In certain cases, interactions among strangers take place at distance, without physical information about others. Nonetheless, even in these cases we can obtain information about them that may bias our decisions, even if this knowledge is unrelated to their actual behavior. In this regard, initial research showed that positive and negative moral information about others influence decisions and reduce reliance on feedback for learning [111].

First, we can assume several characteristics when interacting with people who are familiar to us. Thus, closeness with partners is associated with higher trustworthiness judgments and cooperation decisions, accompanied by higher response in the striatum and mPFC when friends reciprocate [112]. Also, striatal activity seems related to reputation learning of agents with different closeness [112]. Yet, there are situations where we need to make decisions involving unknown people, which is a frequent scenario in experimental settings. In this regard, Hackel et al. [113] showed that the striatum supports feedback-based instrumental learning, integrating different sources of social information, while vmPFC activation correlates with behavioral decisions according to trait-learned information about generosity during social exchanges.

Moreover, our choices can also be influenced by knowledge about our partners' personal characteristics. For instance, participants reject more offers from partners associated with negative descriptions compared to those described by positive information [114]. These influences are stronger when offers are unfair, as well as in uncertain contexts [115]. Negative descriptions of partners compared to positive ones increase the amplitude of the medial frontal negativity (MFN), a potential associated with the emotional evaluation of negative outcomes [116]. However, this negative polarity is reversed when unfair offers come from a friend, a scenario that

is also associated with fewer rejection rates [117]. This may reflect that personal information about the partners, as well as social distance, bias the evaluation of *objective* offers differently, making them look more adverse when the partners are associated with negative personal information. In addition, previous information can influence competence expectations, related to choices whether to continue or not a social interaction with a specific partner [21].

In conclusion, several individual factors carry a heavy impact on social decision-making. In nonnatural controlled scenarios, these sources of bias can also be evaluated through the manipulation of motivational and emotional elements in the experimental setting. In these contexts, biases relate to stereotypes built on the characteristics of others, which are represented at several stages that take place during the analysis of perceptual and social representations about others. To avoid such information when it conflicts with internal goals, control mechanisms become essential.

4 Control Mechanisms During Social Decision-Making

Adaptive social interactions need control mechanisms to regulate actions in scenarios where biases conflict with short- or long-term goals. Here we review part of the evidence on the functioning of these mechanisms. Our focus is on regulation mechanisms involved in economic and moral decisions as well as in contexts where automatic responses must be controlled or our expectations clash with other agents' behavior.

A large part of the biases reviewed so far are studied in relation to the control mechanisms that steer the organism toward context-appropriate actions. For instance, in a classic study, Sanfey et al. [8] employed the ultimatum game to explore reactions to unfair offers. Here, they observed increased activity to unfair offers in the ACC, a region related to conflict of different types, which suggests the existence of interference between emotional reactions and the monetary goals of the task. In addition, they also observed a trade-off between the activity of the AI and the dlPFC to unfair offers. Specifically, the activation in the insula was larger than in the dlPFC when unfair offers were rejected, which may reflect the negative reactions associated with unfairness. On the other hand, activity in the dlPFC was higher when unfair offers were accepted, supporting the function of this region in the control of social behavior.

Similarly, Knoch et al. [118] showed that the disruption of the dlPFC reduces rejection rates to unfair offers. In this vein, Baumgartner et al. [119], observed that dlPFC and vmPFC activity, as well as their effective communication, was needed to make costly normative choices, that is, to reject offers and, therefore, lose earnings. However, the role of these regions in social decisions is not clear, as other studies have shown that people with damage to the vmPFC seem more likely to reject unfair offers [51]. These divergent results might be explained by different dynamics in the PFC in healthy participants and neurological patients, and they also suggest the importance of the communication between these prefrontal regions to regulate social behaviors.

In moral dilemmas, reasoning processes also influence our choices [120]. For example, performing the *cognitive reflection task* induces a decrement of confidence in intuition, related to an increment in utilitarian judgments [120]. These utilitarian decisions have been associated with activity in the dlPFC, inferior parietal lobe, and PCC [121]. In addition, the disruption of activity of the dlPFC after transcranial magnetic stimulation increases utilitarian choices [122]. Taken together, these data add support to the role of the PFC in overcoming emotional reactions in moral scenarios [123].

Ochsner and Gross [124] proposed the mediation of two routes in this control. The dorsal PFC, which has been related to orientation to task context and goals, would be in charge of changing stimuli-emotional response associations. On the other hand, the ventral PFC would maintain the representation of the emotional value of stimuli according to the context and jointly with the OFC and would impact emotional reactions through its reciprocal connections with the nucleus accumbens and the amygdala. These regions would, in turn, modulate the representation of relevant information in higher control areas (e.g., PFC, OFC). In addition, rational behavior in framing tasks, in which decisions are not influenced by framing effects, is correlated with enhanced activation of the OFC and vmPFC [57].

Moreover, biases that derive from stereotypes of prejudice toward others can also be modulated by control. Top-down processes can attenuate this influence [76] through the detection of conflict between goals and biases by the dACC and implementation by the PFC of domain-general control. In addition, the mPFC and the rostral ACC seem to be in charge of more specific representations of social cues to orient regulatory processes to suppress behavior opposite to social norms [24]. Another type of control-demanding situations takes place when facing emotional conflict during decision-making, where need to route the emotional information displayed by faces and attend to the relevant information to fulfill our goals. The resolution of emotional interference has been associated with the activation of the IFG [125] and top-down modulation of the amygdala by the ACC [126].

In social decisions, control mechanisms are recruited when we hold expectations about other people that are not matched by their actual behavior. In this regard, when emotional expressions do not predict their "natural consequences" (happiness = cooperation, anger = no cooperation), there is an increment in the N1 potential, related to attentional processes [102]. Moreover, when emotional displays interfere with identity expectations, Alguacil et al. [127] observed an early conflict effect during face processing, associated with higher amplitude in N170 potential, associated with structural encoding of faces. Later stages linked to response selection were also affected, as reflected by increments in the amplitude of the P3 potential.

The violation of expectations, when we need to overcome the automatic response associated with the expectations induced by emotional expressions, has been shown to engage activity in the PFC, ACC, and AI [101]. This study also observed different coupling of the ACC depending on the level of conflict. While in low-conflict contexts the ACC showed greater interaction with the precuneus and the vmPFC, high conflict was associated with greater coupling with control-related regions, such as the supplementary motor area and the middle region of the cingulate cortex. This

agrees with data indicating that emotional conflict engages the increment of task-relevant information processing, including high-level areas involved in non-emotional tasks [128].

Furthermore, the ability to respond accordingly to previous expectations, even in the presence of behavior which conflicts with that information, seems to be in charge of the ventrolateral prefrontal cortex (vlPFC). In this regard, Fouragnan et al. [129] observed that deactivations in the ventral striatum when trust was violated were functionally correlated with vlPFC activation. Therefore, the vlPFC modulated striatal activity to orient decisions to match expectations when these conflicted with the observed behavior. In addition, [111] observed increased activation in the ACC when participants offered responses that contradicted previous information they held about their partners.

In addition, research in the field of cognitive control suggests the existence of two different networks linked to control. The frontoparietal and cingulo-opercular networks act at different timescales to orient our behavior according to our goals [130]. The increase of activation in these networks has been traditionally associated with a deactivation in the default mode network (DMN), which has been interpreted as an indicator of this network's absence of functionality during difficult tasks [131]. Interestingly, although these mechanisms have been more extensively studied in nonsocial contexts, Cáceda et al. [132] have observed similar neural patterns related to prosocial behavior. These authors reported that enhanced intrinsic connectivity between the salience and the central executive networks (insula/ACC and dlPFC/posterior parietal cortex, respectively) predicted increased cooperation decisions. Moreover, multivariate approaches have been employed to explore control mechanisms encoding the response in social scenarios. For example, Hollmann et al. [133] employed real-time MVPA to explore control mechanisms during economic decisions. They observed that participants' decisions to reject the offers could be decoded in the AI and lateral portion of the PFC (lPFC). Taken together, results add further evidence to the need of control mechanisms to successful social functioning.

In this section, we have reviewed how control mechanisms are recruited to overcome interference. Such conflictive situations tend to arise from contradictions related to personal information or from the incompatibility between personal goals and non-appropriate or automatic responses, which may appear very early in time. Through coordinated activations, frontal regions participate in the evaluation of stimuli and expectations, and they also contribute to maintain neural representations of relevant goals to flexibly adjust behavior.

5   Final Remarks and Conclusions

We have reviewed some of the contributions of social neuroscience to the understanding of the sources of bias in social decisions. We introduced the methodologies that allow the study of the behavior and neural underpinnings of these phenomena,

reviewed internal and external sources of biases, and considered the control mechanisms engaged during conflictive situations.

Altogether, the evidence underscores the relevance of the amygdala and the vmPFC in the integration of emotional and social signals relevant to guide our behavior in social scenarios. The amygdala may enhance processing of social relevant stimuli, while the vmPFC has been related to the representation of other's intentions. Furthermore, positive mood seems to foster cooperation through the reward system (e.g., striatum), which is also in charge of reputational learning according to observed behavior. Conversely, negative states engage areas associated with conflict and the emotional value of negative outcomes, such as the AI or the ACC. Moreover, the OFC appears crucial to represent expectations, especially based on stereotypical information. Interestingly, these expectations also dictate representations in lower-level regions, such as the fusiform cortex, which suggests the importance of top-down modulations in the representation of social information.

The evidence suggests the presence of common pathways of biases on perception and on decisions. For instance, Amodio [24] proposed a neural circuitry for stereotyping, which included mainly the vmPFC, amygdala, AI, and OFC. As we have seen above, these regions also are involved in other biasing contexts. The mPFC has been associated with prosocial dispositions as well as with the representation of a partner's personal traits. Its ventral part also seems relevant for the integration of emotional stimuli in moral dilemmas, including framing effects, as well as in categorization processes [74, 108]. Moreover, the amygdala is necessary to regulate interpersonal trust and facial categorizations [25, 50]. The AI appears to be involved in the emotional evaluation of negative outcomes, which can be guided by negative mood, prejudice, or trustworthiness [61, 101]. Furthermore, the OFC seems to be in charge of representing expectations of others based on stereotypes and emotions [22, 108] and to guide adaptive behavior in social contexts [53]. In addition to these areas, the ventral striatum has a central role in reward-related processing, learning in social scenarios about the valence of the interactions, and fostering interactions associated with positive outcomes [112, 113].

As regards control mechanisms, the evidence points to the relevance of regions such as the ACC, the AI, and several regions of the PFC to maintain goals and suppress deviant responses and to modulate regions involved in social processes to increase attention to the task. Furthermore, data suggest that these regions work at a network level, where frontoparietal and cingulo-insular networks seem to foster prosocial behavior. This highlights the importance of control mechanisms in cooperative scenarios, not only to overcome automatic or undesirable responses but also to behave adaptively in our social environment. Crucially, the data shows that there is important similarity between control mechanisms involved during social decisions and those that have been extensively studied in nonsocial domains (e.g., [134]).

Likewise, in this review we have presented some evidence noting the relevance for cooperative behavior of some regions associated with the DMN, which comprises areas such as the mPFC, precuneus, PCC, angular gyrus, and some temporal areas. This network has been considered until very recently as functionally inactive

during effortful tasks, being involved in mind wondering and self-referential processes [131]. However, recent data seem to indicate that the DMN encodes task-relevant information, even in complex settings and nonsocial tasks [135, 136]. In addition, it has also been related to socials tasks [137, 138] and emotional engaging in social interactions [129]. Unraveling the processes underlying this network is a field of intensive current research (e.g., [139]).

Taking all this into consideration, the use of different methodologies turns crucial to understand how social information is represented in the brain and how different mechanisms coordinate with each other to regulate human social activity and orient our behavior toward goals. Given the complexity of social scenarios, more realistic paradigms are being developed to be implemented in laboratories, in more natural settings [140]. In this regard, the use of methodologies, such as fNIRS in social scenarios (e.g., [141]), may be an interesting approach to study the influences on social decisions in real life.

Social neuroscience is an interdisciplinary and vibrant field. It incorporates methodologies from complementary fields to generate a description of the variety of factors that can influence our interactions and how the different biases operate from early to late stages of processing. In this context, social decisions are key to understand interpersonal exchanges, which are crucial in our life. These processes are important to analyze group dynamics, social perception, or how rational decisions such as economic ones are modulated by different factors. Furthermore, this field may aid to develop interventions for patients with some sort of neural damage that affects their social functioning. Finally, social contexts can extend our knowledge about how our brain works in a large diversity of scenarios filled with rich social stimuli, where decisions take place. Hence, current research efforts provide a comprehensive view of the mechanisms underlying core processes in our daily social life.

References

1. Sanfey AG. Social decision-making: insights from game theory and neuroscience. Science. 2007;318(5850):598–602. https://doi.org/10.1126/science.1142996.
2. Adolphs R, Anderson D. Social and emotional neuroscience. Curr Opin Neurobiol. 2013;23(3):291–3. https://doi.org/10.1016/j.conb.2013.04.011.
3. Fehr E, Gachter S. Cooperation and punishment in public goods experiments. Am Econ Rev. 2000;90(4):980–94. https://doi.org/10.1257/aer.90.4.980.
4. Ruff CC, Fehr E. The neurobiology of rewards and values in social decision making. Nat Rev Neurosci. 2014;15(8):549–62. https://doi.org/10.1038/nrn3776.
5. Frith CD, Frith U. Implicit and explicit processes in social cognition. Neuron. 2008;60(3):503–10. https://doi.org/10.1016/j.neuron.2008.10.032.
6. van Kleef GA. How emotions regulate social life. Curr Dir Psychol. 2009;18(3):184–8. https://doi.org/10.1111/j.1467-8721.2009.01633.x.
7. Dunne S, O'Doherty JP. Insights from the application of computational neuroimaging to social neuroscience. Curr Opin Neurobiol. 2013;23(3):387–92. https://doi.org/10.1016/j.conb.2013.02.007.
8. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. The neural basis of economic decision-making in the ultimatum game. Science. 2003;300(5626):1755–8. https://doi.org/10.1126/science.1082976.

9. Stolier RM, Freeman JB. Functional and temporal considerations for top-down influences in social perception. Psychol Inq. 2016;27(4):352–7. https://doi.org/10.1080/10478 40X.2016.1216034.

10. Cacioppo JT, Berntson GG. Social psychological contributions to the decade of the brain. Doctrine of multilevel analysis. Am Psychol. 1992;47(8):1019–28.

11. Ochsner KN, Lieberman M. The emergence of social cognitive neuroscience. Am Psychol. 2001;56(9):717–34.

12. Fehr E, Fischbacher U, Kosfeld M. Neuroeconomic foundations of trust and social preferences neuroeconomic foundations of trust and social preferences. Am Econ Rev. 2005;95(2):346–51. https://doi.org/10.1257/000282805774669736.

13. Sanfey AG, Chang LJ. Multiple systems in decision making. Ann N Y Acad Sci. 2008;1128:53–62. https://doi.org/10.1196/annals.1399.007.

14. Leotti LA, Delgado MR. The value of exercising control over monetary gains and losses. Psychol Sci. 2014;25(2):596–604. https://doi.org/10.1177/0956797613514589.

15. Berkman ET, Cunningham WA, Lieberman MD. Research methods in social and affective neuroscience. In: Reis HT, Judd CM, editors. Handbook of research methods in personality and social psychology. 2nd ed. New York: Cambridge University Press; 2014. p. 123–58.

16. Camerer CF. Behavioral game theory: experiments in strategic interactions. Princeton: Princeton University Press; 2003.

17. Ruz M, Acero JJ, Tudela P. What does the brain tell us about the mind? Psicológica. 2006;29:149–57.

18. Güth W, Schmittberger R, Schwarze B. An experimental analysis of ultimatum bargaining. J Econ Behav Organ. 1982;3(4):367–88.

19. Sally D. Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. Ration Soc. 1995;7(1):58–92.

20. Gilam G, Lin T, Raz G, Azrielant S, Fruchter E, Ariely D, et al. Neural substrates underlying the tendency to accept anger-infused ultimatum offers during dynamic social interactions. NeuroImage. 2015;120:400–11. https://doi.org/10.1016/j.neuroimage.2015.07.003.

21. Heijne A, Sanfey AG. How social and nonsocial context affects stay/leave decision-making: the influence of actual and expected rewards. PLoS One. 2015;10(8):e0135226. https://doi.org/10.1371/journal.pone.0135226.

22. Stolier RM, Freeman JB. Neural pattern similarity reveals the inherent intersection of social categories. Nat Neurosci. 2016;19(6):795–7. https://doi.org/10.1038/nn.4296.

23. Greenwald AG, Mcghee DE, Schwartz JLK. Measuring individual differences in implicit cognition. J Pers Soc Psychol. 1998;74(6):1464–80.

24. Amodio DM. The neuroscience of prejudice and stereotyping. Nat Rev. 2014;15(10):670–82. https://doi.org/10.1038/nrn3800.

25. Schiller D, Freeman JB, Mitchell JP, Uleman JS, Phelps EA. A neural mechanism of first impressions. Nat Neurosci. 2009;12(4):508–14. https://doi.org/10.1038/nn.2278.

26. Greene J, Haidt J. How (and where) does moral judgement work? Trends Cogn Sci. 2002;6(12):517–23.

27. Greene JD, Morelli SA, Lowenberg K, Nystrom LE, Cohen JD. Cognitive load selectively interferes with utilitarian moral judgment. Cognition. 2008;107(3):1144–54. https://doi.org/10.1016/j.cognition.2007.11.004.

28. Luck SJ. An introduction to the event-related potential technique. Cambridge, MA: The MIT Press; 2005.

29. Cutini S, Basso Moro S, Bisconti S. Functional near infrared optical imaging in cognitive neuroscience: an introductory review. J Near Infrared Spectrosc. 2012;20:75–92. https://doi.org/10.1255/jnirs.969.

30. Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 2001;293(5539):2425–30. https://doi.org/10.1126/science.1063736.

31. Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis – connecting the branches of systems neuroscience. Front Syst Neurosci. 2008;2:4. https://doi.org/10.3389/neuro.06.004.2008.

32. Tamir DI, Thornton MA, Contreras JM, Mitchell JP. Neural evidence that three dimensions organize mental state representation: rationality, social impact, and alence. Proc Natl Acad Sci U S A. 2016;113(1):194–9. https://doi.org/10.1073/pnas.1511905112.

33. Glover SH, Bumpus MA, Sharp GF, Munchus GA. Gender differences in ethical decision making. Women Manag Rev. 2002;17(5):217–27. https://doi.org/10.1108/09649420210433175.

34. Andreoni J, Vesterlund L. Which is the fair sex? Gender differences in altruism. Q J Econ. 2001;116(1):293–312. https://doi.org/10.1162/003355301556419.

35. Rosen JB, Brand M, Kalbe E. Empathy mediates the effects of age and sex on altruistic moral decision making. Front Behav Neurosci. 2016;10:67. https://doi.org/10.3389/fnbeh.2016.00067.

36. Singer T, Seymour B, Doherty JPO, Stephan KE, Dolan RJ, Frith CD. Empathic neural responses are modulated by the perceived fairness of others. Nature. 2006;439(7075):466–9. https://doi.org/10.1038/nature04271.

37. Moore C. Fairness in children's resource allocation depends on the recipient. Psychol Sci. 2009;20(8):944–8. https://doi.org/10.1111/j.1467-9280.2009.02378.x.

38. Weller D, Hansen Lagattuta K. Helping the in-group feels better: children's judgments and emotion attributions in response to prosocial dilemmas. Child Dev. 2013;84(1):253–68. https://doi.org/10.1111/j.1467-8624.2012.01837.x.

39. Bailey PE, Ruffman T, Rendell PG. Age-related differences in social economic decision making: the ultimatum game. J Gerontol Ser B Psychol Sci Soc Sci. 2012;68(3):356–63. https://doi.org/10.1093/geronb/gbs073.

40. Harlé KM, Sanfey AG. Social economic decision-making across the lifespan: an fMRI investigation. Neuropsychologia. 2012;50(7):1416–24. https://doi.org/10.1016/j.neuropsychologia.2012.02.026.

41. Lim KTK, Yu R. Aging and wisdom: age-related changes in economic and social decision making. Front Aging Neurosci. 2015;7:120. https://doi.org/10.3389/fnagi.2015.00120.

42. Fehr E, Schmidt KM. The economics of fairness, reciprocity and altruism – experimental evidence and new theories. In: Kolm SC, Ythier JM, editors. Handbook of the economics of giving, altruism and reciprocity. Amsterdam: Elsevier; 2006. p. 615–91.

43. Murphy R, Ackermann K. A review of measurement methods for social preferences. ETH Zurich Chair of decision theory and behavioral game theory, working paper. http://vlab.ethz.ch/svo/SVO_rev_paper.pdf.

44. Murphy RO, Ackermann KA, Handgraaf MJJ. Measuring social value orientation. Judgm Decis Mak. 2011;6(8):771–81. https://doi.org/10.2139/ssrn.1804189.

45. Au WT, Kwong JY. Measurements and effects of social-value orientation in social dilemmas. In: Suleiman R, Budescu DV, Fischer I, Messick DM, editors. Contemporary research on social dilemmas. New York: Cambridge University Press; 2004. p. 71–98.

46. Bogaert S, Boone C, Declerck C, Bogaert Boone C, Declerck CHS. Social value orientation and cooperation in social dilemmas: a review and conceptual model. Br J Soc Psychol. 2008;47(3):453–80. https://doi.org/10.1348/014466607X244970.

47. Emonds G, Declerck CH, Boone C, Vandervliet EJM, Parizel PM. Comparing the neural basis of decision making in social dilemmas of people with different social value orientations, a fMRI study. J Neurosci Psychol Econ. 2011;4(1):11–24. https://doi.org/10.1037/a0020151.

48. Emonds G, Declerck CH, Boone C, Seurinck R, Achten R. Establishing cooperation in a mixed-motive social dilemma. An fMRI study investigating the role of social value orientation and dispositional trust. Soc Neurosci. 2014;9(1):10–22. https://doi.org/10.1080/17470919.2013.858080.

49. Wang WH, Shih YH, Yu HY, Yen DJ, Lin YY, Kwan SY, et al. Theory of mind and social functioning in patients with temporal lobe epilepsy. Epilepsia. 2015;56(7):1117–23. https://doi.org/10.1111/epi.13023.

50. Koscik TR, Tranel D. The human amygdala is necessary for developing and expressing normal interpersonal trust. Neuropsychologia. 2011;49(4):602–11. https://doi.org/10.1016/j.neuropsychologia.2010.09.023.

51. Koenigs M, Young L, Adolphs R, Tranel D, Cushman F, Hauser M, et al. Damage to the prefrontal cortex increases utilitarian moral judgements. Nature. 2007;446(7138):908–11. https://doi.org/10.1038/nature05631.

52. Moll J, de Oliveira-Souza R. Moral judgments, emotions and the utilitarian brain. Trends Cogn Sci. 2007;11(8):319–21. https://doi.org/10.1016/j.tics.2007.06.001.

53. Gleichgerrcht E, Ibanez A, Roca M, Torralva T, Manes F. Decision-making cognition in neurodegenerative diseases. Nat Rev Neurol. 2010;6(11):611–23. https://doi.org/10.1038/nrneurol.2010.148.

54. Gleichgerrcht E, Torralva T, Roca M, Pose M, Manes F. The role of social cognition in moral judgment in frontotemporal dementia. Soc Neurosci. 2011;6(2):113–22. https://doi.org/10.1080/17470919.2010.506751.

55. O'Callaghan C, Bertoux M, Irish M, Shine JM, Wong S, Spiliopoulos L, et al. Fair play: social norm compliance failures in behavioral variant frontotemporal dementia. Brain. 2016;139(1):204–16. https://doi.org/10.1093/brain/awv315.

56. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science. 1981;211(4481):453–8.

57. De Martino B, Kumaran D, Seymour B, Dolan RJ. Frames, biases, and rational decision-making in the human brain. Science. 2006;313(5787):684–7. https://doi.org/10.1126/science.1128356.

58. Wang XT, Rao L, Zheng H. Neural substrates of framing effects in social contexts: a meta-analytical approach. Soc Neurosci. 2016;28:1–12. https://doi.org/10.1080/17470919.2016.1165285.

59. Zheng H, Wang XT, Zhu L. Framing effects: behavioral dynamics and neural basis. Neuropsychologia. 2010;48(11):3198–204. https://doi.org/10.1016/j.neuropsychologia.2010.06.031.

60. Forgas JP. Mood and judgment: the affect infusion model (AIM). Psychol Bull. 1995;117(1):39–66.

61. Harlé KM, Chang LJ, van't Wout M, Sanfey AG. The neural mechanisms of affect infusion in social economic decision-making: a mediating role of the anterior insula. NeuroImage. 2012;61(1):32–40. https://doi.org/10.1016/j.neuroimage.2012.02.027.

62. Forgas J. On feeling good and getting your way: mood effects on negotiator cognition and bargaining strategies. J Pers Soc Psychol. 1998;74(3):565–77.

63. Mislin A, Williams LV, Shaughnessy BA. Motivating trust: can mood and incentives increase interpersonal trust? J Behav Exp Econ. 2015;58:11–9. https://doi.org/10.1016/j.socec.2015.06.001.

64. Rand DG, Kraft-Todd G, Gruber J. The collective benefits of feeling good and letting go: positive emotion and (dis) inhibition interact to predict cooperative behavior. PLoS One. 2015;10(1):1–12. https://doi.org/10.1371/journal.pone.0117426.

65. Nygren TE, Isen AM, Taylor PJ, Dulin J. The influence of positive affect on the decision rule in risk situations: focus on outcome (and especially avoidance of loss) rather than probability. Organ Behav Hum Decis Process. 1996;66(1):59–72.

66. Tan HB, Forgas JP. When happiness makes us selfish, but sadness makes us fair: affective influences on interpersonal strategies in the dictator game. J Exp Soc Psychol. 2010;46(3):571–6. https://doi.org/10.1016/j.jesp.2010.01.007.

67. Szekely RD, Miu AC. Incidental emotions in moral dilemmas: the influence of emotion regulation. Cognit Emot. 2015;29(1):64–75. https://doi.org/10.1080/02699931.2014.895300.

68. Grecucci A, Giorgetta C, Bonini N, Sanfey AG. Reappraising social emotions: the role of inferior frontal gyrus, temporo-parietal junction and insula in interpersonal emotion regulation. Front Hum Neurosci. 2013;7:523. https://doi.org/10.3389/fnhum.2013.00523.

69. Grecucci A, Giorgetta C, Van't Wout M, Bonini N, Sanfey AG. Reappraising the ultimatum: an fMRI study of emotion regulation and decision making. Cereb Cortex. 2013;23(2):399–410. https://doi.org/10.1093/cercor/bhs028.

70. Bar M, Neta M, Linz H. Very first impressions. Emotion. 2008;6(2):269–78. https://doi.org/10.1037/1528-3542.6.2.269.

71. Bickart KC, Dickerson BC, Feldman Barrett L. The amygdala as a hub in brain networks that support social life. Neuropsychologia. 2014;63:235–48. https://doi.org/10.1016/j.neuropsychologia.2014.08.013.

72. Tajfel H, Billig MG, Bundy RP, Flament C. Social categorization and intergroup behavior. Eur J Soc Psychol. 1971;1(2):149–78. https://doi.org/10.1002/ejsp.2420010202.

73. Fiske ST. Stereotyping, prejudice, and discrimination. In: Gilbert DT, Fiske ST, Lindzey G, editors. The handbook of social psychology. New York: McGraw-Hill; 1998. p. 357–411.

74. Quadflieg S, Turk DJ, Waiter GD, Mitchell JP, Jenkins AC, Macrae CN. Exploring the neural correlates of social stereotyping. J Cogn Neurosci. 2009;21(8):1560–70. https://doi.org/10.1162/jocn.2009.21091.

75. Freeman JB, Ma Y, Barth M, Young SG, Han S, Ambady N. The neural basis of contextual influences on face categorization. Cereb Cortex. 2015;25(2):415–22. https://doi.org/10.1093/cercor/bht238.

76. Frith CD, Frith U. How we predict what other people are going to do. Brain Res. 2006;1079(1):36–46. https://doi.org/10.1016/j.brainres.2005.12.126.

77. Stanley DA, Sokol-Hessner P, Fareri DS, Perino MT, Delgado MR, Banaji MR, et al. Race and reputation: perceived racial group trustworthiness influences the neural correlates of trust decisions. Philos Trans R Soc B Biol Sci. 2012;367(1589):744–53. https://doi.org/10.1098/rstb.2011.0300.

78. Gilbert SJ, Swencionis JK, Amodio DM. Evaluative vs. trait representation in intergroup social judgments: distinct roles of anterior temporal lobe and prefrontal cortex. Neuropsychologia. 2012;50(14):3600–11. https://doi.org/10.1016/j.neuropsychologia.2012.09.002.

79. Ito T, Urland G. Race and gender on the brain. J Pers Soc Psychol. 2003;85(4):616–26. https://doi.org/10.1037/0022-3514.85.4.616.

80. Ofan RH, Rubin N, Amodio DM. Seeing race: N170 responses to race and their relation to automatic racial attitudes and controlled processing. J Cogn Neurosci. 2011;23(10):3153–61. https://doi.org/10.1162/jocn_a_00014.

81. Tortosa M, Lupiáñez J, Ruz M. Race, emotion and trust: an ERP study. Brain Res. 2013;1494:44–55. https://doi.org/10.1016/j.brainres.2012.11.037.

82. Bentin S, Allison T, Pruce A, Perez E, Mccarthy G. Electrophysiological studies of face perception in humans. J Cogn Neurosci. 1996;8(6):551–65.

83. Ibáñez A, Gleichgerrcht E, Hurtado E, Gonzalez R, Haye A, Manes F. Early neural markers of implicit attitudes: N170 modulated by intergroup and evaluative contexts in IAT. Front Hum Neurosci. 2010;4:188. https://doi.org/10.3389/fnhum.2010.00188.

84. Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends Cogn Sci. 2000;4(6):223–33. https://doi.org/10.1016/S1364-6613(00)01482-0.

85. Contreras JM, Banaji MR, Mitchell JP. Multivoxel patterns in fusiform face area differentiate faces by sex and race. PLoS One. 2013;8(7):e69684. https://doi.org/10.1371/journal.pone.0069684.

86. Brosch T, Bar-David E, E a P. Implicit race bias decreases the similarity of neural representations of black and white faces. Psychol Sci. 2013;24(2):160–6. https://doi.org/10.1177/0956797612451465.

87. Mendoza SA, Lane SP, Amido DM, Amodio DM. For members only: ingroup punishment of fairness norm violations in the ultimatum game. Soc Psychol Personal Sci. 2014;5(6):662–70. https://doi.org/10.1177/1948550614527115.

88. Stanley DA, Sokol-Hessner P, Banaji MR, Phelps EA. Implicit race attitudes predict trustworthiness judgments and economic trust decisions. Proc Natl Acad Sci U S A. 2011;108(19):7710–5. https://doi.org/10.1073/pnas.1014345108.

89. Freeman JB, Ambady N. A dynamic interactive theory of person construal. Psychol Rev. 2011;118(2):247–79. https://doi.org/10.1037/a0022327.

90. Adolphs R. Perception and emotion: how we recognize facial expressions. Curr Dir Psychol Sci. 2006;15(5):222–6. https://doi.org/10.1111/j.1467-8721.2006.00440.x.

91. Oosterhof NN, Todorov A. The functional basis of face evaluation. Proc Natl Acad Sci U S A. 2008;105(32):11087–92. https://doi.org/10.1073/pnas.0805664105.

92. Todorov A, Said CP, Engell AD, Oosterhof NN. Understanding evaluation of faces on social dimensions. Trends Cogn Sci. 2008;12(12):455–60. https://doi.org/10.1016/j.tics.2008.10.001.

93. G a v K. The emerging view of emotion as social information. Soc Personal Psychol Compass. 2010;4(5):331–43. https://doi.org/10.1111/j.1751-9004.2010.00262.x.

94. Scharlemann JPW, Eckel CC, Kacelnik A, Wilson RK. The value of a smile: game theory with a human face. J Econ Psychol. 2001;22(5):617–40. doi: https://ora.ox.ac.uk/objects/uuid:60f12fb9-4fea-4f5f-9837-7d1982010b76.

95. Mussel P, Göritz AS, Hewig J. The value of a smile: facial expression affects ultimatum-game responses. Judgm Decis Mak. 2013;8(3):1–5.

96. Hareli S, Hess U. What emotional reactions can tell us about the nature of others: an appraisal perspective on person perception. Cognit Emot. 2010;24(1):128–40. https://doi.org/10.1080/02699930802613828.

97. Ibañez A, Kotz SA, Barrett L, Moll J, Ruz M. Situated affective and social neuroscience. Front Hum Neurosci. 2014;8:547. https://doi.org/10.3389/fnhum.2014.00547.

98. de Melo CM, Carnevale PJ, Read SJ, Gratch J. Reading people's minds from emotion expressions in interdependent decision making. J Pers Soc Psychol. 2014;106(1):73–88. https://doi.org/10.1037/a0034251.

99. van Kleef GA, De Dreu CKW, Manstead ASR. The interpersonal effects of anger and happiness in negotiations. J Pers Soc Psychol. 2004;86(1):57–76. https://doi.org/10.1037/0022-3514.86.1.57.

100. Alguacil S, Tudela P, Ruz M. Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions. Psicológica. 2015;36(2):309–35. http://www.redalyc.org/articulo.oa?id=16941182006.

101. Ruz M, Tudela P. Emotional conflict in interpersonal interactions. NeuroImage. 2011;54(2):1685–91. https://doi.org/10.1016/j.neuroimage.2010.08.039.

102. Ruz M, Madrid E, Tudela P. Interactions between perceived emotions and executive attention in an interpersonal game. Soc Cogn Affect Neurosci. 2013;8(7):838–44. https://doi.org/10.1093/scan/nss080.

103. Gobbini MI, Haxby JV. Neural systems for recognition of familiar faces. Neuropsychologia. 2007;45(1):32–41. https://doi.org/10.1016/j.neuropsychologia.2006.04.015.

104. Fett AKJ, Shergill SS, Joyce DW, Riedl A, Strobel M, Gromann PM, et al. To trust or not to trust: the dynamics of social interaction in psychosis. Brain. 2012;135(3):976–84. https://doi.org/10.1093/brain/awr359.

105. Todorov A, Baron SG, Oosterhof NN. Evaluating face trustworthiness: a model based approach. Soc Cogn Affect Neurosci. 2008;3(2):119–27. https://doi.org/10.1093/scan/nsn009.

106. Todorov A, Pakrashi M, Oosterhof NN. Evaluating faces on trustworthiness after minimal time exposure. Soc Cogn. 2009;27(6):813–33. https://doi.org/10.1521/soco.2009.27.6.813.

107. Freeman JB, Stolier RM, Ingbretsen ZA, Hehman EA. Amygdala responsivity to high-level social information from unseen faces. J Neurosci. 2014;34(32):10573–81. https://doi.org/10.1523/JNEUROSCI.5063-13.2014.

108. Kim H, Choi M-J, Jang I-J. Lateral OFC activity predicts decision bias due to first impressions during ultimatum games. J Cogn Neurosci. 2012;24(2):428–39. https://doi.org/10.1162/jocn_a_00136.

109. Rezlescu C, Duchaine B, Olivola CY, Chater N. Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. PLoS One. 2012;7(3):e34293. https://doi.org/10.1371/journal.pone.0034293.

110. van't Wout M, Sanfey AG. Friend or foe: the effect of implicit trustworthiness judgments in social decision-making. Cognition. 2008;108(3):796–803. https://doi.org/10.1016/j.cognition.2008.07.002.

111. Delgado MR, Frank RH, Phelps EA. Perceptions of moral character modulate the neural systems of reward during the trust game. Nat Neurosci. 2005;8(11):1611–8. https://doi.org/10.1038/nn1575.

112. Fareri DS, Chang LJ, Delgado MR. Computational substrates of social value in interpersonal collaboration. J Neurosci. 2015;35(21):8170–80. https://doi.org/10.1523/JNEUROSCI.4775-14.2015.

113. Hackel LM, Doll BB, Amodio DM. Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. Nat Neurosci. 2015;18(9):1233–5. https://doi.org/10.1038/nn.4080.

114. Gaertig C, Moser A, Alguacil S, Ruz M. Social information and economic decision-making in the ultimatum game. Front Neurosci. 2012;6(July):1–8. https://doi.org/10.3389/fnins.2012.00103.

115. Ruz M, Moser A, Webster K. Social expectations bias decision-making in uncertain inter-personal situations. PLoS One. 2011;6(2):e15762. https://doi.org/10.1371/journal.pone.0015762.

116. Moser A, Gaertig C, Ruz M. Social information and personal interests modulate neural activity during economic decision-making. Front Hum Neurosci. 2014;8(February):31. https://doi.org/10.3389/fnhum.2014.00031.

117. Campanhã C, Minati L, Fregni F, Boggio PS. Responding to unfair offers made by a friend: neuroelectrical activity changes in the anterior medial prefrontal cortex. J Neurosci. 2011;31(43):15569–74. https://doi.org/10.1523/JNEUROSCI.1253-11.2011.

118. Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. Science. 2006;314(5800):829–32. https://doi.org/10.1126/science.1129156.

119. Baumgartner T, Knoch D, Hotz P, Eisenegger C, Fehr E. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. Nat Neurosci. 2011;14(11):1468–74. https://doi.org/10.1038/nn.2933.

120. Paxton JM, Ungar L, Greene JD. Reflection and reasoning in moral judgment. Cogn Sci. 2012;36(1):163–77. https://doi.org/10.1111/j.1551-6709.2011.01210.x.

121. Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD. The neural bases of cognitive conflict and control in moral judgment. Neuron. 2004;44(2):389–400. https://doi.org/10.1016/j.neuron.2004.09.027.

122. Tassy S, Oullier O, Duclos Y, Coulon O, Mancini J, Deruelle C, et al. Disrupting the right prefrontal cortex alters moral judgement. Soc Cogn Affect Neurosci. 2012;7(3):282–8. https://doi.org/10.1093/scan/nsr008.

123. Frith C, Singer T. The role of social cognition in decision making. Philos Trans R Soc Lond Ser B Biol Sci. 2008;363(1511):3875–86. https://doi.org/10.1098/rstb.2008.0156.

124. Ochsner KN, Gross JJ. The cognitive control of emotion. Trends Cogn Sci. 2005;9(5):242–9. https://doi.org/10.1016/j.tics.2005.03.010.

125. Levens SM, Phelps EA. Insula and orbital frontal cortex activity underlying emotion interference resolution in working memory. J Cogn Neurosci. 2010;22(1978):2790–803. https://doi.org/10.1162/jocn.2010.21428.

126. Etkin A, Egner T, Peraza DM, Kandel ER, Hirsch J. Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. Neuron. 2006;51(6):871–82. https://doi.org/10.1016/j.neuron.2006.07.029.
127. Alguacil S, Madrid E, Espín AM, Ruz M. Facial identity and emotional expression as predictors during economic decisions. Cogn Affect Behav Neurosci. 2016:1–15. https://doi.org/10.3758/s13415-016-0481-9.
128. Egner T, Hirsch J. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat Neurosci. 2005;8(12):1784–90. https://doi.org/10.1038/nn1594.
129. Fouragnan E, Chierchia G, Greiner S, Neveu R, Avesani P, Coricelli G. Reputational priors magnify striatal responses to violations of trust. J Neurosci. 2013;33(8):3602–11. https://doi.org/10.1523/JNEUROSCI.3086-12.2013.
130. Dosenbach NUF, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. A dual-networks architecture of top-down control. Trends Cogn Sci. 2008;12(3):99–105. https://doi.org/10.1016/j.tics.2008.01.001.
131. Raichle M. The brain's default network. Ann N Y Acad Sci. 2015;8(38):433–47. https://doi.org/10.1146/annurev-neuro-071013-014030.
132. Cáceda R, James GA, Gutman DA, Kilts CD. Organization of intrinsic functional brain connectivity predicts decisions to reciprocate social behavior. Behav Brain Res. 2015;292:478–83. https://doi.org/10.1016/j.bbr.2015.07.008.
133. Hollmann M, Rieger JW, Baecke S, Lützkendorf R, Müller C, Adolf D, et al. Predicting decisions in human social interactions using real-time fMRI and pattern classification. PLoS One. 2011;6(10):e25304. https://doi.org/10.1371/journal.pone.0025304.
134. Duncan J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behavior. Trends Cogn Sci. 2010;14(4):172–9. https://doi.org/10.1016/j.tics.2010.01.004.
135. Crittenden BM, Mitchell DJ, Duncan J. Recruitment of the default mode network during a demanding act of executive control. elife. 2015;2015(4):e06481. https://doi.org/10.7554/eLife.06481.
136. González-García C, Arco JE, Palenciano AF, Ramírez J, Ruz M. Encoding, preparation and implementation of novel complex verbal instructions. NeuroImage. 2017;148:264–73. https://doi.org/10.1016/j.neuroimage.2017.01.037.
137. Mars RB, Neubert F-X, Noonan MP, Sallet J, Toni I, Rushworth MFS. On the relationship between the "default mode network" and the "social brain". Front Hum Neurosci. 2012;6:189. https://doi.org/10.3389/fnhum.2012.00189.
138. Li W, Mai X, Liu C. The default mode network and social understanding of others: what do brain connectivity studies tell us. Front Hum Neurosci. 2014;24(8):74. https://doi.org/10.3389/fnhum.2014.00074.
139. Margulies DS, Ghosh SS, Goulas A, Falkiewicz M, Huntenburg JM, Langs G, et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. Proc Natl Acad Sci. 2016;113(44):12574–9. https://doi.org/10.1073/pnas.1608282113.
140. Gilam G, Hendler T. With love, from me to you: embedding social interactions in affective neuroscience. Neurosci Biobehav Rev. 2016;68:690–701. https://doi.org/10.1016/j.neubiorev.2016.06.027.
141. Oliver D, Tachtsidis I, Hamilton AF. The role of parietal cortex in overimitation: a study with fNIRS. Soc Neurosci. 2017:1–12. https://doi.org/10.1080/17470919.2017.1285812.