

Valuing Others: Evidence from Economics, Developmental Psychology, and Neurobiology

Pablo Billeke, Patricia Soto-Icaza, Mauricio Aspé-Sánchez,
Verónica Villarroel, and Carlos Rodríguez-Sickert

Abstract Human social skills are widely studied among very different disciplines. In this chapter, we review, discuss, and relate evidence concerning the process of valuing others' perspectives, preferences, and behaviors from an economic, psychological, and neurobiological viewpoint. This process of valuing others (or other-regarding preferences) can be understood as weighing others' preferences to adapt our own behavior and achieve adequate social interaction. We first review economic research related to decision-making in social contexts, with emphasis on how decision-making has integrated other-regarding preferences into the decision-making algorithm. By means of social and developmental psychology research, we then review how social skills develop from identification to understanding others. Finally, we discuss the neurobiological mechanisms underlying social skills and social decision-making, focusing on those systems that can participate in processes of valuing others preferences. As a conclusion, we highlight five points that we believe an interdisciplinary approach should take into account. We thus intend to generate a starting point for building a more extensive explicatory bridge among the different disciplines that study complex human social behavior.

Keywords Neuroeconomics • Decision-making • Other-regarding preferences • Mentalization • Theory of mind • Social cognition • Interdisciplinary approach • Game theory

P. Billeke (✉) • M. Aspé-Sánchez • C. Rodríguez-Sickert
División de Neurociencias (NeuroCICS), Centro de Investigación en Complejidad Social,
Facultad de Gobierno, Universidad del Desarrollo,
Av. Las Condes 12461, Las Condes, Santiago 7590943, Chile
e-mail: pbilleke@udd.cl; maspes@udd.cl; carlosrodriguez@udd.cl

P. Soto-Icaza
Interdisciplinary Center of Neuroscience, Pontificia Universidad Católica de Chile,
Santiago, Chile
e-mail: pasoto@uc.cl

V. Villarroel
Centro de Investigación y Mejoramiento de la Educación (CIME), Facultad de Psicología,
Universidad del Desarrollo, Concepción, Chile
e-mail: vvillarroel@udd.cl

1 Introduction

We are an extremely social species; almost all of our behavior is related to other human beings. Currently, various disciplines deal with the problem of understanding human social behavior. However, few proposals that combine different approaches and findings have been elaborated. In this chapter, we discuss the evidence and research approaches from an array of disciplines related to the idea of how humans consider other preferences and behaviors during this decision-making process. We shall use the term “valuing others” to refer to the processes by which humans weigh the preferences and behaviors of others as to adapt or guide their behavior during social interactions. In the following pages, our endeavor will be to present and discuss the evidence from three research programs, namely, (1) economics research related to decision-making in social contexts, (2) social psychology research related to the development of mentalizing and perspective-taking skills, and (3) neuroscience research related to neuronal mechanisms underlying vicarious human behaviors.

The fundamental aim of this chapter is to show some of the current efforts to build an interdisciplinary understanding of social behavior instead of giving a global integrative approach. In order to build a fully interdisciplinary research programming between social science and neuroscience, the authors have established some basic bridges which are necessary to discuss and begin to build this understanding. Therefore, with the purpose of contributing to this global aim, we have structured this chapter in three sections. In the first one, we discuss how the approach from economics toward the social decision-making process has started to incorporate social preferences and how neuroscience approaches can contribute to improving the predictive ability of the behavioral model. In the second section, we review evidence from developmental psychology related to how human beings begin to understand and integrate the perspective of others into their own behavior and decisions. Finally, we discuss findings from social neuroscience and neuroeconomics related to the neurobiological mechanisms that underlie social decision-making, in order to suggest possible interdisciplinary approaches, and their possible pitfalls.

2 Behavioral Models of Human Conduct and the Black Box

In recent years, the emergence of subfields such as neuroeconomics and social neuroscience has driven the dialogue between behavioral economics and natural science. Especially, behavioral economics has relied on game theory an experimental paradigm for neuroscientists when studying complex social behavior inside the controlled settings of a laboratory. Likewise, to concurrently record or modulate brain activity—by means of techniques such as electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and transcranial magnetic stimulation

(TMS) (see below)—could shed a light on the cognitive mechanisms that underlie the behavior of experimental subjects and their reactions against the behavior of their fellow partners.

When there is a confluence of disciplines, the potential gains of combining both perspectives might be hampered by language barriers (e.g., jargon that is discipline specific) and incongruities between the widespread research practices within each discipline (e.g., the importance that is given to generality in contrast with parsimony or to prediction over explanation). In this section, we suggest three perspectives that can lead to a fruitful interdisciplinary interaction from the perspective of economics. We focus on (1) the neurophysiological foundations of behavioral models of social preferences, (2) general guidelines for modeling social behavior and social cognition, and (3) specific instantiations of neurophysiological variables within those behavioral models.

2.1 *Homo Behavioralis and the Influx of Ideas from Psychology and Other Disciplines*

When scholars from disciplines such as psychology or anthropology began to question the plausibility of the prevalent model of human agency in economics, the reply came from one of his most renowned representatives. Milton Friedman wrote his famous *Essays in Positive Economics* (1951), which strongly influenced future generations of economist researchers [1]. There he claimed that “the only relevant test of the validity of a hypothesis is the comparison of its predictions with experience.” Furthermore, Friedman argued that even if assumptions appear false or implausible, their empirical weakness should be tolerated if they lead to accurate predictions. When Friedman adds this second statement, not only can one infer that he was oblivious to the advances of neuroscience but also that the aim of Friedman and his fellow custodians was to keep the black box closed and to keep the *homo economicus* locked inside [2].

It is not that the members of the congregation for the Doctrine of the Economic Faith denied the existence of other drivers of human behavior beyond self-interest—e.g., altruism. Nor did they believe we are perfect optimizers. Their stance relied on an argument of parsimony: the benefits of generalizing the utility function to account for possible anomalies and produce more accurate predictions would be negligible against the loss of parsimony and tractability of adding new parameters to the utility function. The overwhelming amount evidence from laboratory and field experiments showed that this view on the trade-off between prediction power and parsimony was not accurate. The effort to correct this mistake was assumed by a new breed of “behavioral” economists. Indeed, one can say that there is nothing new in this approach. They are just continuing the enterprise launched by Adam Smith himself, as a moral philosopher, in *The Theory of Moral Sentiments* (1759) [3].

The first task undertaken by the behavioral squad was to upgrade the utility function so that these “anomalies” could be captured within an augmented utility function. Around psychological constructs, such as loss aversion and reference dependence, Kahneman and Tversky developed prospect theory [4]. While Kahneman, Tversky, and their followers focused on decision under uncertainty, and later on issues such as intertemporal inconsistency [5, 6], a separate group of behavioral economists reacted to the strong evidence against the self-interest hypothesis provided by experimental studies. These studies showed that agents do cooperate in social dilemmas such as trust games [7–11], public good games [12–14], even when cooperating is against their (material) self-interest. And, within bargaining games such as the ultimatum game [15–17], agents are willing to incur in material costs to avoid unfair outcomes and sanction free riders in collective action problems [18].

Taking their insights from social psychology, sociology, and anthropology, a family of models was produced within behavioral economics. These models, referred to as models of social (or other-regarding) preferences, can be either outcome based, e.g., models of inequity aversion [19, 20], or intention based, e.g., models that capture norms of both positive and negative reciprocity [21–23]. Cooperation in trust games was initially understood as the result of positive reciprocity (intention-based social preferences). The trustee is willing to spend resources to reward trust placed in him. On the other hand, rejection on the ultimatum game was initially understood as the result of inequity aversion. However, later studies provided evidence for a more complex structure of moral response. Trustees in a trust game are also motivated by outcome-based preferences [8], and rejection in the ultimatum game also involves negative reciprocity [24]. Furthermore, current studies show that the research on social preferences can also be extracted by the research produced in other areas of behavioral economics. For instance, time inconsistency can also affect the nature of social preferences [25].

To the extent that neuroeconomic studies have provided neurophysiological mechanisms for experimental anomalies and, thus, biological foundations for social preferences models, neuroeconomists were welcomed as part of the new tribe of behavioral economists but were not so well received by old-school orthodox economists who were still concerned with keeping the black box closed even for the new model of human agency: the *homo behavioralis* and its representation in an augmented utility function. For instance, it has been argued that neuroscience could not transform economics because what goes on inside the brain is irrelevant to the discipline. As if nothing had changed since Friedman’s influential piece, they put forward the idea of a “mindless economics,” arguing that what matters are the decisions people make, not the process by which they reach them [25]. We will develop this idea in the opposite direction and claim that the major challenges posited by neuroeconomics precisely relate to our understanding of the neurocognitive processes that underlie social behavior and, furthermore, open the possibility to embed economics in the biological processes taking place in the brain.

2.2 *Impact of Neurosciences on Modeling Individual and Social Behavior*

In the same way that behavioral economics has used insights from psychology to develop more “realistic” models of individual decision-making, in which people often did things that were not in their best interests, the evidence coming from neurobiology presents an additional challenge to the standard economic assumptions. Thus, evidence from neuroeconomics indicates that decision-making is far from being a unitary process (a simple matter of integrated and coherent utility maximization), suggesting instead that it is driven by the interaction of multiple systems or processes [26]. This range from the more basic dual-process approach that has influenced our general comprehension of human cognition and behavior beyond Descartes’ error (fast/hot module and the slow/cold, automatic vs. controlled processes [26–28]) to more complex multiple system approaches toward social behavior and social decision-making [29–31]. Steinbeis et al. [32], for instance, show that behavioral inhibition—modulated by the neuroanatomical development of the cognitive control systems—plays a crucial role in the implementation of fair behavior in bargaining games.

2.3 *Prediction Accuracy of Behavioral Models: Combining Psychological and Neurobiological Parameters*

A specific aspect of the relevance of the neuroeconomic program refers to its capacity to inform behavioral models in such a way that prediction accuracy can be improved. This point is very important, because if we do not build a bridge between neuroscience and algorithmic social decision theory, it will be very difficult for this program to reach the academic community of economists. To discuss the issues that could emerge from this challenge, we consider a distributional problem in the spirit of Andreoni and Miller [33], in which an agent i decides how to split an amount m between himself and another agent $-i$ for different budget constraints. For every monetary unit agent i sacrifices ($m - x_i$), his partner will receive $(m - x_i)/p$ monetary units. Thus, p can be interpreted as the price of altruism and agent i ’s choice can be represented as the consumer’s choice problem.

2.3.1 Neoclassical Model (*Homo Economicus*, Black Box)

$$\begin{aligned} & \max_{x_i} \{U_i(x_i)\} \text{ s.t. } x_i + px_{-i} = m, \\ & \text{which yields to } x_i(m, p) = x_i^* \text{ with } x_i^* = m \text{ and } x_{-i}^* = 0. \end{aligned}$$

In the case above, the only relevant argument of $U_i(\bullet)$ is his own material self-interest x_i . If, alternatively, we consider that agent i 's choice is also affected by the material welfare of his partner $-i$, we could represent his choice problem as follows.

2.3.2 Behavioral Model (Other-Regarding Preferences, Black Box)

$$\begin{aligned} & \max_{x_i, x_{-i}} \{U_i(x_i, x_{-i}, \theta_i)\} \text{ s.t } x_i + px_{-i} = m, \\ & \text{which yields to } x_i(m, p, \theta_i) = x_i^* \text{ with } x_i^* \left\langle m \text{ and } x_{-i}^* \right\rangle 0 \end{aligned}$$

where θ_i is a parameter that represents the intensity of the moral dispositions of the agent that counterbalances his self-interest.¹ Most models assume that θ_i is private. Now consider the possibility that θ_i can be estimated from the neurobiological activation n_i , $\hat{\theta}_i(n_i) = \theta_i + \varepsilon_i$. If this is the case, the lower the measurement error, the greater will be the predictive gains of opening the black box. The registered neurobiological activation n_i could give us information about θ_i through two channels: the individual's idiosyncratic characteristics and the dimensions of the stimuli not captured by the model. For the sake of simplicity, we will assume that n_i is simply a contextual modulator of θ_i . Thus, the structure of choice could be represented as follows.

2.3.3 Neurobiological Model (Other-Regarding Preferences, Neurobiological State)

$$\begin{aligned} & \max_{x_i, x_{-i}} \{U_i(x_i, x_{-i}, \theta_i | n_i) s_i\} \text{ s.t } x_i + px_{-i} = m, \\ & \text{which yields to } x_i(m, p, \theta_i | n_i) = x_i^{**}; x_i(m, p) = x_i^{**} > 0, x_{-i}^{**} > 0 \end{aligned}$$

The improvement in prediction accuracy of a model that incorporates n_i is an indicator of the incompleteness of the behavioral model. However, it is not only important to come up with a model that accurately predicts behavior in a particular context. Fehr and Camerer [34], among others, argue that prosocial behaviors occur in one-shot anonymous games as the result of a reflexive behavior that is highly adapted for repeated interactions where immediate prosocial behavior earns future benefits. Under this view, prosociality in one-shot games results from bounds on rationality in full response to changes in the economic structure. Alternatively, pro-

¹A simple functional specification of the agent's social preferences could be expressed as $U_i(x_i, x_{-i}, \theta_i) = (1 - \theta_i) \times u(x_i) + \theta_i \times u(x_{-i}) = (1 - \theta_i) \times \sqrt{x_i} + \theta_i \times \sqrt{x_{-i}}$ where θ_i represents the weight agent i attaches to his partner individual welfare. In some alternative functional specifications, both considerations to the efficiency and equity of the final distribution have been introduced (see [164]).

social behavior could reflect robust social preferences for treating others generously or reciprocally, and those preferences are similar to preferences for other kinds of primary and secondary rewards. Within this scheme, different arrangements of neurobiological activation $n_i^0 \uparrow n_i^1$ could lead to similar predictions in terms of cooperation that could indicate the motives underlying both cases. Such a case has been shown recently; see below [35]. Furthermore, these neural traits could provide crucial information to distinguish different types of individuals and, consequently, have more information about their behavior in the future or in different social contexts.

A crucial issue in this point is what are precisely these neurobiological traits and states and how these states weigh the parameters of self-interest and other-regarding preferences. Although neuroscientists are far from reaching consensus, there is accumulative evidence that can indicate some general structures of these traits and states. In the following section, we will review some critical evidence from developmental psychology and developmental neuroscience in order to give insight on how these neurobiological states mature and change during the ontogeny. Then, in the final section, we shall analyze how these neurobiological processes can be structured, with special focus on how the system weighs and values the regarding preferences of others.

3 Development of Social Preferences

One of the most relevant facts indicates that the neurobiological state has a decisive influence in the decision-making process is the human development. The maturity of different brain systems in different timelines generates several behavioral manifestations that are characteristic to a specific age [36, 37]. This is true not only during childhood and adolescence but also for older adults where pathological neuronal degeneration is expected [38, 39].

Regarding early human development research, one of the most intriguing human social phenomena is the ability to read the minds of others, known as “mentalization” or “theory of mind”. This ability has been described as one of the major landmarks in social development, because it enables children to handle more complex social interactions. Indeed, the ability to figure out and finally to attribute and understand the other person’s thoughts and feelings has been depicted as a distinctive human trait [40]. However, the mechanism by which this ability emerged has been the subject of drawn-out controversy [41–43]. The analysis of the development of human social functioning is a useful tool for understanding how social skills are structured. This analysis reveals that social ability development is not a unitary or an “all-or-nothing” type of outcome. Instead, it is an interactive specialization that entails both the association of an ability with a brain system and the specialization of this function in interaction with others [44]. In this context, one of the main drivers for this development is the necessity to anticipate and predict the behavior of others, which is crucial for both primate and human survival [45]. Certainly, the newborn ability to discriminate a relevant biological agent seems to be coordinated

to, first, a guarantee that the partner is actually a living being and, second, that this living being is actually human. As human babies are born premature [43], their extreme dependency puts them at higher risk; hence, they must draw the interlocutor's attention directly to them in order to modify the performance of others to get what he or she needs to survive. It seems possible that the later human ability to "read minds" arises from all those previous early stages of social development as a guarantee for survival since it constitutes a specialized expertise of social prediction. This section is organized in three overlapping stages of development, starting with the early capacity to identify biological/social agents and ending with the explicit manifestation of mentalization skills.

3.1 Identification of Social Agents in Newborns and Infants

The early stages of social development are the building blocks in which further social skills have grown. Certainly, the only way that a human infant can survive is if there is another being that can provide food, water, etc. Evidence in newborns showed that toddlers as young as only few days of age are able to discriminate different perceptual signs that indicate the existence of a social agent [46–48]. For example, they can identify points that emulate a coherent biological motion [48], face-like patterns [49], and direct versus averted gaze in faces [50, 51], and they can even imitate basic movements from another human being [52]. Indeed, from 2 months old, infants show a preference for looking at eyes rather than mouths or bodies [53]. This preference also describes a specialization process in 3-month-old toddlers, who prefer eyes only when they are accurately located in the upper part of the face configuration rather than placed in another location of the face [54]. All these findings are showing that there is an ontogenetic orientation toward the social agents, which seems to be in a growing process of behavioral and neural specialization. Indeed, comparative studies between preterm and full-term infants and among subjects of different ages [55–58] emphasize the role of the experience in the cerebral functions refinement [44]. From biological motion detection to imitation and face-like stimuli and direct gaze preference in newborns, human social development seems to be organized to detect, understand, and finally predict and manipulate the social agent [59].

EEG findings in infants and children are in accordance with this developmental perspective. The EEG technique is a noninvasive measurement of the brain activity through scalp electrodes widely used in neuroscience [60]. The evidence has shown that the electrical brain activity phase related to stimulus presentation, called event-related potentials (ERPs) [61], follows a developmental trajectory. An illustrative example is the N170, that is a negative deflection occurring at 170 ms after presentation of a human [60, 62–69], whose likely source is the ventral visual stream near the fusiform face area. In adults, the N170 evidenced a higher amplitude and latency for inverted human faces, while in infants it did not show any modulation by stimuli orientation. In 6-month-old infants, there is an "infant N170" (called P400 component)

characterized by higher amplitude in response to faces displaying direct gaze rather than an averted gaze [50], as well as to inverted faces only in the case of their mother [70], evidencing a specific selection process present in early life.

3.2 Being Able to Interact with Other Humans

It is important to note that these skills are present in a context of reciprocal interaction [71, 72]. While it is clear that infant behaviors like crying, screaming, gazing, and smiling are aimed to make the social partner answer their requirements, it is also clear that the partner cannot remain indifferent to those calls of attention. What actually happens when infants and their caregivers are coordinated or synchronized? It has been described that in mother-toddler relationships with infants from 3 to 6 months, the engagement periods came in a burst mode, with periods of asynchronous states [72]. Interestingly, these mismatch states are followed by repair sequences of the interactive errors by both the infant and the mother. These repair behaviors can have functionality in the interaction skills development. Indeed, the importance of stages as “reparation” contexts has been widely described in the attachment theory [71]. Precisely, these bonding-recovering stages emphasize the importance of the mutuality of the attachment between the caregiver and the infant which is crucial to underline [71, 73]. The higher social skills like mentalization abilities were the result of all these precursors or early stages of development, which are the building blocks in which further social skills are grown [59].

An important step in the development of the capacity to interact with other human beings is the joint attention (JA) skill. JA has been described as the capacity to share an interest with another person by alternating the gaze in order to coordinate the interest in an object with a social partner [74–80]. A key component of JA is the division and the alternation of the subject’s attention between the object and the partner [77, 81]. Several studies agree that JA emerges around the age of 9 months [74, 76, 77, 82], when children learn to use eye contact to derive information about another person’s goal-directed behavior [76]. Importantly, the ability to attend to an object jointly with another person has proved to be crucial for several capacities such as social synchronization, development of language [74, 76, 78, 79, 83, 84], and development of theory of mind [80]. The knowledge of the latter tends to be ambiguous to clarify if JA involves a level of “self-awareness” of the social agent [45]. Does the infant actually “know” the agent’s state of mind when is engaged in a JA interaction? There is a line of studies that defines JA as the situation in which two subjects are looking at the same object but without the awareness that the focus of attention is a common interest. The real capacity to realize that the focus of attention is a common element between the infant and the agent is what is called “shared attention” [45]. Accordingly, what is clearly a higher development of social knowledge is the mentalization ability, which is the capacity to understand and predict the behavior of other people and their knowledge, intentions, emotions, and beliefs [85, 86]. Furthermore, JA and shared attention would be intermediate

stages toward mentalization inasmuch as the theory of mind ability solely enables to notice and take into account the agent's mental state. Interestingly, the neuroimaging evidence revealed that JA and mentalization might be related. Specifically, fMRI is a method that measures changes in the hemodynamic brain response associated with neural activity—specifically, the blood-oxygen-level-dependent (BOLD) signal [87]. There is broad consensus about the brain network that is recruited when adult subjects participate in mentalizing tasks (see next section below). Interestingly, the same network is involved when participants show JA behavior in adulthood and later childhood. During early childhood, the EEG evidence shows that responses to JA are associated with the Nc component. This ERP refers to a negative deflection that occurs around 300–850 ms after stimulus onset [56, 66, 77, 82], and it is associated with attentional reorientation. In children during the age when they can achieve the false-belief mentalization, this component did not seem to present any differences. However, two neuronal measures seem to mark the mentalization achievement. One of these is the presence of a specific oscillatory activity in the temporoparietal areas of the mentalization network (see next section) and the maturation on neural fiber that connects the frontal and temporoparietal regions [88]. Thus, specific neuronal development seems to be a marker for more complex social skills achievement.

3.3 *Knowing the Others' Mental States*

What do infants know about the mental states of others? Do they actually try to modify the actions of others because they can infer what is in their minds? Premack and Woodruff [89] stated that the mentalization ability is a system of inference that enables us to attribute mental states both to oneself and to another—for instance, purposes, intentions, knowledge, belief, and thinking. Certainly, this system of inference is needed because such “mental states” are not directly observable, making it a “theory” of what are the others' mental states (i.e., theory of mind). The explicit skill to identify other people's false beliefs becomes evident not before 4 years of age [85, 90]. However, there is a line of research that describes how infants are able to do some kind of inferences about others' feelings and thoughts [91–94]. That line of studies appeared as alternative experimental paradigms to overcome the language-dependent bias which standard/classic false-belief tasks [86] have. Hence, the infants' difficulty both to inhibit their own knowledge about something that another person does not know and to think over different representations makes this task impossible to solve for children under 4 years old [95, 96]. Therefore, researchers use infants' longer looking time as a measure of children's anticipatory belief [94] or surprise as measure of a violation of the expectation paradigm [92, 93, 97] in nonverbal false-belief tasks. Thus, this line of research has shown that there is evidence of an “implicit” theory of mind [91]. However, there is another line of research that has been skeptical about this interpretation [41–43, 98, 99]. This evidence can be interpreted just as perceptual processes and competences

rather than high-level cognitive processes. Furthermore, high-level constructs that come from this experimental paradigm might be revealing the researcher's over-interpretation instead of the ability for which it was created [43]. Indeed, the increase in looking times that these studies have shown might be revealing a visual perception process related with a new arrangement of the stimuli rather than an interpretation of the agent's belief [99].

At this point of the controversy, it is important to consider that the implicit mentalization ability, the JA ability, the different levels of visual perspective taking (mentioned below), and the explicit theory of mind itself could be understood as stages of complexity inside the development process of the same capacity. The visual perspective taking (VPT) is the capacity to know that an object can be seen from a certain point of view and that someone else could not see it because there is a physical barrier [100]. Research of VPT should also be considered to understand the mentalization development as a dynamic building block process. These studies provide interesting evidence to consider the existence of an intermediate level of mentalization [59]. The first level of VPT [101, 102] can be understood as a previous step toward a well-consolidated theory of mind, because, around the age of 2, the child is only able to identify whether another person can see an object or not, but it says nothing about a genuine capacity to attribute the mental state of the agent. Nevertheless, this VPT level becomes more complex a couple of years after when it allows the child to identify the others' references and perspectives [90, 98, 101, 102]. This higher VPT level, known as Level 2 VPT, allows the child to understand that objects can be seen in different ways, depending on the form of presentation and point of view [98, 101, 102]. There is evidence that correlates Level 2 VPT with the development of mentalization ability [101]. Although the first theories point out that the visual perspective taking is the basic process from which more complex (social) perspectives arise, recent evidence indicates an opposite ontological development [102]. Early infants can track others' experiential backgrounds. In fact, several studies have found that infants take what others have witnessed into account when acting and responding toward them. In other words, the infants revert to the background constituted by past experiences and use it to understand an agent's desires, goals, and intentions. This ability becomes evident before infants can solve complex visual perspective-taking tasks (Level 2) and even before they can solve explicit mentalizing problems, like the false-belief task [91]. This evidence indicates that the developmental processes that lead to the explicit mentalizing ability are related to the integration of others' preferences into our behaviors. This skill, as an integrative process, becomes more complex through aging, incorporating more sources of information, such as memories, social knowledge, and visual skills, among others. Thus, the development of this skill serves as the basis for more complex explicit mentalizing or the theory of mind skill. Following the deconstruction of the mentalizing concept proposed elsewhere [103], the skill of valuing others can help us gather not well matching evidence, which has come from cognitive neuroscience and neuroeconomics. In the next section, we will review neuroscience evidence related to brain components of the system of other-regarding preferences.

4 Neurobiological System Related to Other-Regarding Preferences

Our brain has evolved to solve complex cognitive demands required for living in social groups of increasing size [104]. Experimental evidence has established that, unlike other social species, humans display a large amount of cooperative behaviors, including altruism, trust, and reciprocity [105, 106]. These behaviors are observed even when individuals interact with strangers and with individuals they will never meet again [107]. Trust, altruism, and reciprocity are crucial to establish and maintain cooperative links between different individuals. Recent work using neuroscience techniques has begun to reveal the brain states related to these prosocial dispositions [108]. In the following subsection, we will review evidence from neuroeconomic studies using two game theory experimental paradigms, namely, trust and dictator games. Then, we shall discuss evidence from the two putative systems related to other-regarding preferences or “valuing others” processes that can underlie human prosocial behaviors.

4.1 *Trust and Reciprocity*

The most widely used experimental setting to study trust and reciprocity is the trust game (TG) or invested game. In this game, two players, who do not know each other, engage in an anonymous interaction. The experimenter gives the “investor” (or trustor) some amount T of money. The trustor then decides how much of T send (or “invest”) in the other player, referred to as the trustee. The amount A_1 sent by the trustor is multiplied by an exchange factor r (typically 3). Thus, the trustee receives an amount of money three times the amount sent by the trustor (rA_1). Finally, the trustee decides how much of the money received (rA_1) is sent back to the trustor (A_2) [7]. The prediction from the self-interest hypothesis for TG is that the trustees will keep all the money. Assuming that the trustors have mentalizing capabilities (see above), they should anticipate this betrayal and send nothing. In the very first test of this game, 0.6% of the trustors sent nothing to the trustee, 66% sent half or more of their endowment, and about 50% ended the game with more money than their initial endowment (which implies, of course, that $A_2 > A_1$; in other words, trustees were trustworthy; [7]. These behaviors have been replicated in several studies. In a recent meta-analysis, Johnson and Mislim [109] collected the data from the 162 replications of the TG available at the time and found that, on average, trustors send 0.5 of his/her endowment to the trustee ($n = 23,900$; std = 0.12; min = 0.22; max = 0.89), while the trustee returns 0.37 of their total endowment ($n = 21,529$; std = 0.11; min = 0.11; max = 0.81 [109]. Repeated interactions of the TG show a similar pattern, indicating a high tendency toward trust and reciprocity by both players [107].

Trustee behavior is interesting. While, for trustors, there is an expected gain, this is not so clear for trustees. The trustee has the opportunity to break the trust, which

is, as stated above, the classical self-interest prediction. This is particularly true for one-shot, anonymous interactions since there are no incentives to build reputation and create a greater amount of trust for future interactions. Classically, trustee's behavior has been considered just reciprocity, but this is only true if allocations made by trustees are different from allocations made by a subject in a context where his/her behavior is unrelated with the perceived intentions of cooperation from the other player [110]. There is a difference between intention-based behaviors, such as the behavior in the TG, where trustee's behavior depends on ascribing cooperative intentions to the trustor, and outcome-based behaviors, such as the behavior in the dictator game (DG, described below), where subject behavior depends only on the final share of the game and not on the others' intentions.

4.2 *The Neural Dynamics of the TG*

In a TG, the very first decision by the trustor involves deciding whether to trust the other player or not. From the trustors' perspective, this involves (1) knowing whether they are playing with another human or a non-intentional entity (generally a computer which makes random allocations) and (2) then deciding to send or not to send some amount of money to the trustee. Several reports have shown increased activity in the medial prefrontal cortex (mPFC, a structure involved in metallization processes; [111] when trustors decide to trust another human partner [112–115]. In addition, during the first stage, the trustor has not received any feedback on the trustworthiness of his/her partner; therefore, the reinforcement learning system must be engaged to adjust trustor behavior based on feedback reward. Delgado et al. [114] read the descriptions of the life events of different trustees to trustors, indicating praiseworthy, neutral, or suspicious moral characters for each of them. Not surprisingly, rates of cooperation were higher when playing with the praiseworthy partner. Interestingly, trustors showed different activation in the ventral striatum (VS) for positive and negative feedback but only when they were playing with the neutral trustee. The VS has been involved in processing feedback and prediction error [114, 116], suggesting that, in the neutral condition, trustors activate the reinforcement system to learn about the trustworthiness of their partners, while praiseworthy and suspicious moral characters bias the behavior of trustors [114]. Interestingly, the neuropeptide oxytocin (OXT) has been associated with trust behaviors in humans [117, 118]. Kosfeld et al. [119] used a TG experiment to show that intranasal infusions of OXT increase trust in humans (but not in other nonsocial interactions), do not increase risk-taking behavior, and did not change trustees' behavior. Although the mechanism of action of OXT is not clear, evidence suggests that OXT decreases stress responses and anxiety in social interactions, likely modulating the amygdala and anterior cingulate cortex (ACC) activity [117, 120].

Considering now the situation of the trustees, reports show that the mentalization system becomes active when they receive an allocation from trustors. Van den Bos et al. [121] has shown that the mPFC increases its activation when trustees defect.

On the other hand, when trustees reciprocate a high-risk allocation (i.e., the trustor could lose a large amount of money if the trustee chose to defect), there is greater activation of the temporoparietal junction, which is also a part of the mentalization system [122–124]. Moreover, trustees' reciprocity in low-risk allocations correlated with the activity in the anterior insula cortex (AIC), a structure involved in emotional and salience processing [113, 125]. Furthermore, trustees reciprocating low benefit allocations (i.e., when the monetary incentives to reciprocate are low) were associated with an increased activity in the ACC and the dorsolateral prefrontal cortex (dlPFC), which are structures involved in cognitive control and the inhibition of selfish impulses [126–129].

Another interesting finding is the effect of individual traits in reciprocal interaction [121, 130]. For example, people with more traits characterized by positive emotionality trust more in others, while people with less tendency to psychopathic traits show more reciprocate behaviors [130]. Other study shows that when a prosocial subject reciprocated, they showed an increased activation in VS, while defection increased the activity in ACC, AIC, and right TPJ. In contrast, pro-self individuals showed the opposite pattern, showing increasing ACC, AIC, and right TPJ activity after they reciprocated. This shows that these structures were more active when participants chose their less frequent behavior, considering their personal trait or past history [121].

Trustees' reciprocal behavior is also influenced by expectations [131]. Chang et al. [131] asked trustees about their second-order beliefs (i.e., how much money they think the trustor expects) and compared these second-order beliefs with the amount that trustees actually send. With this information, they could categorize the allocations made by trustees as “minimizing guilt” (when the amount sent was close to the trustees' second-order beliefs) or “maximizing outcome” (when trustees sent an amount significantly smaller than what they expected based on their second-order beliefs). When trustees minimized guilt, they exhibited higher activation in dlPFC, AIC, and dorsal ACC, which are structures reported to be activated by negative affective states [132–134]. On the other hand, when trustees maximized outcome, higher activation occurred in ventral mPFC, VS, and dorsal mPFC. The authors proposed a model where minimizing guilt increased AIC activation, which increased activation in dorsal mPFC, while maximizing outcome decreased AIC activation, which increased activation in the VS [131].

4.3 *Altruism*

Historically, altruism has been studied by means of the dictator game (DG). In this game, there are also two players involved in an anonymous one-shot interaction. The first player, called “dictator,” receives an amount T of money and donates some a part of it ($A1 \in [0, T]$) to the second player, called the “recipient.” This decision ends the game and the recipient has no participation in deciding about this distribution. Crucially, the recipient has no chances of punishing the dictator if the amount

is not acceptable to him. Thus, there are not direct incentives for a strictly self-interested dictator to share any portion of the received money, and any donation is defined as an altruistic act [108, 135]. Behavioral evidence shows that even when participants play this game with unknown others, dictators tend to donate around 25% of their money to the recipient [136]. Interesting variants have been introduced to the game. Cherry et al. [137], for instance, made the dictators earn their own money, thereby giving subjects a sense of ownership. In this case, about 91% of the dictators don't send anything to the recipient. In addition, there have been recent efforts to include social knowledge about the recipient in the DG [138, 139]. Such experimental settings have shown that there are important variables which explain allocations, such as the knowledge about who the recipient is and how the game is explicitly described to the players [136]. Likewise, social distance is an important modulator of behavior in the DG. Hoffman et al. [140] showed that 64% of dictators kept all the money when social distance was maximized. In addition, some authors have shown that donations tend to be higher when people are informed that the recipient is a real charitable organization [138, 139].

4.4 Neuronal Dynamic of the DG

Despite its simplicity, and the fact that it has been used widely in behavioral economics, few neuroeconomics experiments have used the DG to assess the neural basis of altruism. In a recent article, Hutcherson et al. [141] made subjects participate in a DG where subjects had to choose between two options of allocation. By using this protocol, they induced choices between the default 50–50% split, generous (benefiting the other at a cost to oneself) or selfish behavior (benefiting oneself at a cost to another). The authors fitted a drift-diffusion model which assumes that choices are the output of a noisy process that weighs the linear sum of monetary outcomes for self and others. In this model, the choice is made when sufficient neural evidence has accumulated in favor of one of the options, and it assumes that the valuing of self and other outcomes is computed independently and then integrated in an overall value signal. At the neural level, the authors found that ventromedial prefrontal cortex (vmPFC) activity correlated positively with the value that subjects assigned to proposals, as measured by the Likert response scale. vmPFC has been reported to encode stimulus values at the time of decision in a wide range of tasks [142, 143]. Moreover, fitting general linear models (GLM), they found that valuations toward self-outcomes correlate with the activity in both vmPFC and VS, while valuations toward other outcomes correlate with the activity in the right TPJ, precuneus, and vmPFC. These results, further discussed below, show that the right TPJ is an area that becomes activated specifically when focusing on others, while vmPFC combines information about self and others.

In another experiment, Hein et al. [35] studied the role of empathy and reciprocity motives in human altruism. Using a DG, they investigated differences in altruistic behavior from experimental subjects when they observed recipients (1) receiving

painful shocks (empathy partner) or (2) giving an amount of money to save some of those empathy partners from painful shocks (reciprocity partner), an action perceived as kind and, thus, one that should elicit reciprocity motives. A baseline partner neither received painful shock nor was instructed to give money for saving subjects from shock. Authors observed that subjects behave more altruistically toward the empathy and the reciprocity partners, noteworthy, without significant differences in allocations between the two motive inductions. At the neural level, a network consisting of AIC, VS, and ACC was activated in both motive-induction conditions. Moreover, individual pattern of brain connectivity in this network predicts subjects' altruist behavior. Interestingly, this prediction was particular for each treatment. Thus, a positive connectivity between ACC and AIC and a slightly negative connectivity between AIC and VS predict empathy-driven altruism, while a strong bidirectional projection between AIC and ACC and a positive connectivity between AIC and VS predict reciprocity-driven altruism. Additionally, the ACC connectivity to AIC correlates positively with baseline levels of altruism. Notice that, at the behavioral level, both motives were indistinguishable, because motives are a mental construct hidden to revealed preferences. A neuroeconomic approach is able to unravel both motives and their influence on altruistic behavior.

4.5 Two Putative Systems for Valuing Others' Outcomes

4.5.1 Anterior Cingulate Cortex and Vicarious Performance Monitoring

As seen above, a set of cognitive and affective functions determining the need for adaptive control prove central to economic decision-making [144]. A key neural structure that participates in these functions is the ACC, which is involved in interactions such as reciprocity, choosing the less common behavior [128, 145, 146], empathy and reciprocity-driven motives in human altruism [35], violations of social norms [147, 148], and mediating the effects of OXT in trust behavior [117].

The ACC is the frontal part of the cingulate cortex. Anatomically, the ACC has classically been subdivided in a rostral (rACC) and a dorsal part (dACC) [149]. The inputs to dACC include the amygdala, AI, orbitofrontal cortex, vmPFC and mid-brain, and prominent ventral tegmental area. Its outputs target the lateral PFC, the motor cortex, striatum, subthalamic nucleus, and locus coeruleus [150]. The activity of the dACC has been correlated with almost the whole set of known psychological variables. Broadly speaking, dACC has been considered a key hub in a network of brain regions implicated in domain-general executive functions in humans [127], being important for cognitive control (i.e., our ability to flexibly adjust behavior according to internally maintained goals and away from behaviors that are more automatic but distracted from those goals [149]. Consequentially, there exists some agreement relating the involvement of the dACC in motivation and reward-based decision-making [127, 151].

However, there is no clear consensus on the function of dACC. Currently, two main proposals interpret its functioning: the expected value of control (EVC) theory and the foraging value theory (FVT). EVC [150] proposes that dACC plays a central role in decisions about the allocation of cognitive control based on a cost (for instance, the effort needed) and benefit (for instance, improved performance) analysis that identifies the highest EVC. The FVT theory, on the other hand, argues that difficulty or control allocation is insufficient to account for all dACC activity [152]. Instead, the dACC plays a key role in behavioral flexibility. Its activity reflects the history, weighted by time of occurrence, of previously chosen rewards, computing the value of persisting in the current environment versus the value of switching away from it [153].

Following the evidence review above and other experiments using economic social exchanges [115, 147, 154], some researchers argue that particular areas of ACC track, specifically, behavioral motivation and prediction errors not of self but specifically of others [149]. In this line, studies suggest that the gyrus region of the ACC (ACCg) computes “other-oriented” information (i.e., information about other agents that might be animals or people, rather than ourselves). Apps et al. [155], for instance, examined the brain activity of human subjects when they received cues about the level of an economic reward and the cost incurred for receiving this reward, under conditions in which the costs and rewards pertained to the same experimental subjects or to a third person. In this experiment, ACCg activity correlated with the net value of rewards to be received by the third person when the third person incurred the cost of the effort. By contrast, the ACC sulcus signaled the effort level regardless of whether the effort was exerted by the subject or by a third person [149]. Authors found, “with a striking consistency,” that the ACCg responds exclusively to other-oriented information.

4.5.2 Temporoparietal Regions and Valuing Others’ Processes

As reviewed above, mentalization is our ability to represent and attribute others’ mental or internal states, such as ideas, beliefs, desires, emotions, and motivations [31, 156]. Similarly, perspective taking (PT) is the ability to comprehend that the same event or object can be seen or constructed in multiple ways, depending on each subject’s point of view. Both processes enable humans to weight others’ behaviors and preferences into the subjective valuations that underlie decision-making, a process that can be called “valuing others” [38, 39, 124]. At the neurobiological level, meta-analysis studies have shown that this area becomes active in all the tasks involving PT or mentalization [157]. Furthermore, some scholars have proposed that TPJ is a key neural structure underlying the distinction between self and others’ perspectives [156, 158–160].

The involvement of the TPJ in general mentalizing functions can be linked to its anatomical characteristics. TPJ is constituted by the posterior part of the temporal lobe, the inferior part of parietal lobe, and the lateral part of occipital lobe [161]. This area is a heteromodal association cortex integrating multiple sources of sensory

(and non-sensory) information. In addition, this region is located at a maximum synaptic/geodesic distance from sensory and motor areas. This seems to be useful for generating integrative computations addressing inner (abstract) and social processes [162].

There is plenty of evidence highlighting a consistent role of the TPJ in other preferences and how much these preferences affect personal decisions. TPJ is engaged, for instance, when subjects must anticipate others' decisions and behaviors [38, 39, 123, 141, 160, 163], when trustees reciprocate a high-risk allocation when pro-self individuals reciprocate [121], or when dictators evaluate the outcomes of others [141]. All these findings point to the existence of neuronal processes that compute others' preferences and behaviors, where TPJ is a key structure underlying the mechanism that allows us to integrate the others' preferences during a social interaction.

5 Conclusions

Currently there is a broad interest to combine evidence from different fields to better understand our complex social behavior. Our review suggests that, while the integration between social and natural sciences is still elusive, the evidence warrants five conclusions that may guide interdisciplinary discussion among behavioral economics, developmental psychology, and neuroscience. In particular, we believe that it is necessary to take care of the following observations:

1. The process of social decision-making can be understood as an algorithmic process that necessarily needs to be in contrast with real decision-making data.
2. In this algorithmic process, humans take into account multiple motivators (parameters), where self-interest (wellbeing/survival) and other-regarding preferences (valuing others' processing) are the most relevant.
3. The ways by which these motivators are finally integrated strongly depend on the neurobiological organization of multiple (not unitary) systems.
4. The neurobiological system (understood as neurophysiological states and traits) implicates both a general and a variable organization.
5. The variations of these neurobiological systems (not only one black box) depend at least on ontogenic (developmental) states, contextual constraints, and individual predispositions.

The social skills analyzed here are only an example of the areas where multiple disciplines have focused their efforts. Currently, it is extremely necessary to work on establishing common concepts in order to gather disperse perspectives. Through this chapter, we intend to generate a conceptual bridge among the knowledge input from psychology, neuroscience, and economics. This is certainly not a global theoretical framework but rather a starting point for building common conceptual framings in order to increase an interdisciplinary dialogue. In this way, we expect to be able to address difficult and unanswered questions about our amazing and, at the same time, conflictive social behavior.

Acknowledgments This work was supported by Comisión Nacional de Investigación Científica y Tecnológica CONICYT (Grant FONDECYT 11405268 to CR-S, Grant FONDECYT Inicio 11140535 to PB and Grant PCHA/DoctoradoNacional/2014-21140043 to PS-I).

References

1. Friedman M. *Essays in positive economics*. Chicago: University of Chicago Press; 1953.
2. Stigler G, Becker G. De Gustibus Non Est Disputandum. *Am Econ Rev*. 1977;67(2):76–90.
3. Ashraf N, Camerer CF, Loewenstein G. Adam Smith, behavioral economist. *J Econ Perspect*. 2005;19(3):131–45.
4. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica*. 1979;47(2):263–91.
5. O'Donoghue T, Rabin M. Doing it now or later. *Am Econ Rev*. 1999;89(1):103–24.
6. Laibson D. golden eggs and hyperbolic discounting. *Q J Econ*. 1997;112(2):443–78.
7. Berg J, Dickhaut J, McCabe K. Trust, reciprocity, and social history. *Games Econ Behav*. 1995;10(1):122–42.
8. Cox JC. How to identify trust and reciprocity. *Games Econ Behav*. 2004;46(2):260–81.
9. Falk A, Kosfeld M. The hidden costs of control. *Am Econ Rev*. 2006;96(5):1611–30.
10. Sheremeta RM, Zhang J. Three-player trust game with insider communication. *Econ Inq*. 2014;52(2):576–91.
11. Heyes A, List JA. Supply and demand for discrimination: strategic revelation of own characteristics in a trust game. *Am Econ Rev*. 2016;106(5):319–23.
12. Fehr E, Gächter S. Cooperation and punishment in public goods experiments. *Am Econ Rev*. 2000;90(4):980–94.
13. Reuben E, Riedl A. Enforcement of contribution norms in public good games with heterogeneous populations. *Games Econ Behav*. 2013;77(1):122–37.
14. Oprea R, Charness G, Friedman D. Continuous time and communication in a public-goods experiment. *J Econ Behav Organ*. 2014;108:212–23.
15. Brañas-Garza P, Espín AM, Exadaktylos F, Herrmann B. Fair and unfair punishers coexist in the ultimatum game. *Sci Rep*. 2015;4(1):6025.
16. Nowak MA. Fairness versus reason in the ultimatum game. *Science*. 2000;289(5485):1773–5.
17. Güth W, Kocher MG. More than thirty years of ultimatum bargaining experiments: motives, variations, and a survey of the recent literature. *J Econ Behav Organ*. 2014;108:396–409.
18. Fehr E, Gächter S. Altruistic punishment in humans. *Nature*. 2002;415(6868):137–40.
19. Bolton GE, Ockenfels A. ERC: a theory of equity, reciprocity, and competition. *Am Econ Rev*. 2000;90(1):166–93.
20. Fehr E, Schmidt K. A Theory of fairness, competition and cooperation. *Q J Econ*. 1999;114(August):817–68.
21. Rabin M. Incorporating fairness into game theory and economics. *Am Econ Rev*. 1993;83:1281–302.
22. Falk A, Fischbacher U. A theory of reciprocity. *Games Econ Behav*. 2006;54(2):293–315.
23. Dufwenberg M, Kirchsteiger G. A theory of sequential reciprocity. *Games Econ Behav*. 2004;47(2):268–98.
24. Falk A, Fehr E, Fischbacher U. On the nature of fair behavior. *Econ Inq*. 2003;41(1):20–6.
25. Andreoni J, Barton B, Bernheim BD, Aydin D, Naecker J. When fair isn't fair: sophisticated time inconsistency in social preferences. *Work Pap*. 2016;1996:58.
26. Loewenstein G, Rick S, Cohen JD. Neuroeconomics. *Annu Rev Psychol*. 2008;59:647–72.
27. Damasio A. Feelings of emotion and the self. *Ann N Y Acad Sci*. 2003 Oct;1001:253–61.
28. Kahneman D. *Thinking, fast and slow*. New York: Macmillan and Company; 2011.

29. King-Casas B, Chiu PH. Understanding interpersonal function in psychiatric illness through multiplayer economic games. *Biol Psychiatry*. 2012;72(2):119–25.
30. Declerck CH, Boone C, Emonds G. When do people cooperate? The neuroeconomics of prosocial decision making. *Brain Cogn*. 2013;81(1):95–117.
31. Billeke P, Aboitiz F. Social cognition in schizophrenia: from social stimuli processing to social engagement. *Front Psychiatry*. 2013;4(February):1–12.
32. Steinbeis N, Bernhardt BC, Singer T. Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron*. 2012;73(5):1040–51.
33. Andreoni J, Miller J. Giving according to GARP: an experimental test of the consistency of preferences for altruism. *Econometrica*. 2002;70(2):737–53.
34. Fehr E, Camerer CF. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn Sci*. 2007;11:419–27.
35. Hein G, Morishima Y, Leiberg S, Sul S, Fehr E. The brains functional network architecture reveals human motives. *Science*. 2016;351(6277):1074–8.
36. McAuliffe K, Blake PR, Steinbeis N, Warneken F. The developmental foundations of human fairness. *Nat Hum Behav*. 2017;1(2):42.
37. Dalgleish T, Walsh ND, Mobbs D, Schweizer S, van Harmelen A-L, Dunn B, et al. Social pain and social gain in the adolescent brain: a common neural circuitry underlying both positive and negative social evaluation. *Sci Rep*. 2017;7(February 2016):42010.
38. Ibáñez A, Billeke P, de la Fuente L, Salamone P, García AM, Melloni M. Reply: Towards a neurocomputational account of social dysfunction in neurodegenerative disease. *Brain*. 2017;140(3):e15.
39. Melloni M, Billeke P, Baez S, Hesse E, de la Fuente L, Forno G, et al. Your perspective and my benefit: multiple lesion models of self-other integration strategies during social bargaining. *Brain*. 2016;139(11):3022–40.
40. Baars B, Gage N. Social cognition: perceiving the mental states of others. In: *Cognition, brain and consciousness: introduction to cognitive neuroscience*. 2nd ed. San Diego, CA: Elsevier; 2010.
41. Penn DC, Povinelli DJ. On the lack of evidence that non-human animals possess anything remotely resembling a “theory of mind”. *Philos Trans R Soc Lond Ser B Biol Sci*. 2007;362(January):731–44.
42. Povinelli DJ, Vonk J. Chimpanzee minds: suspiciously human? *Trends Cogn Sci*. 2003;7(4):157–60.
43. Aboitiz FA. *Brain for speech. A view from evolutionary neuroanatomy*. London: Palgrave Macmillan; 2017.
44. Johnson MH. Interactive specialization: a domain-general framework for human functional brain development? *Dev Cogn Neurosci*. 2011;1(1):7–21.
45. Emery NJ. The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci Biobehav Rev*. 2000;24(6):581–604.
46. Bertenthal BI, Proffitt DR, Cutting JE. Infant sensitivity to figural coherence in biomechanical motions. *J Exp Child Psychol*. 1984;37(2):213–30.
47. Pavlova M, Sokolov A. Orientation specificity in biological motion perception. *Percept Psychophys*. 2000;62(5):889–99.
48. Simion F, Regolin L, Bulf H. A predisposition for biological motion in the newborn baby. *Proc Natl Acad Sci U S A*. 2008;105(2):809–13.
49. Macchi Cassia V, Simion F, Umiltà C. Face preference at birth: the role of an orienting mechanism. *Dev Sci*. 2001;4(1):101–8.
50. Farroni T, Csibra G, Simion F, Johnson MH. Eye contact detection in humans from birth. *Proc Natl Acad Sci*. 2002;99(14):9602–5.
51. Farroni T, Mansfield EM, Lai C, Johnson MH. Infants perceiving and acting on the eyes: tests of an evolutionary hypothesis. *J Exp Child Psychol*. 2003;85(3):199–212.

52. Meltzoff AN, Moore MK. Imitation of facial and manual gestures by human neonates. Published by: American Association for the Advancement of Science Stable. URL: <http://www.jstor.org/stable/1744187>. 1977;198(4312):75–8.
53. Jones W, Klin A. Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*. 2013;504(7480):427–31.
54. Turati C, Valenza E, Leo I, Simion F. Three-month-olds' visual preference for faces and its underlying visual processing mechanisms. *J Exp Child Psychol*. 2005;90(3):255–73.
55. Macchi Cassia V, Bulf H, Quadrelli E, Proietti V. Age-related face processing bias in infancy: evidence of perceptual narrowing for adult faces. *Dev Psychobiol*. 2014;56(2):238–48.
56. Luyster RJ, Powell C, Tager-Flusberg H, C a N. Neural measures of social attention across the first years of life: characterizing typical development and markers of autism risk. *Dev Cogn Neurosci*. 2014;8:131–43.
57. De Haan M, Johnson MH, Halit H. Development of face-sensitive event-related potentials during infancy. In: De Haan M, editor. *Infant EEG and event-related potentials*. 1st ed. New York: Psychology Press; 2007.
58. Pena M, Arias D, Dehaene-Lambertz G. Gaze following is accelerated in healthy preterm infants. *Psychol Sci*. 2014;25(10):1884–92.
59. Soto-Icaza P, Aboitiz F, Billeke P. Development of social skills in children: neural and behavioral evidence for the elaboration of cognitive models. *Front Neurosci*. 2015;9(September):1–16.
60. Haan M De. Introduction to infant EEG and event-related potentials. In: Haan M, editor. *Infant EEG and event-related potentials*. New York, USA: Psychology Press Ltd New York; 2002. p. 39–76.
61. Luck SJ. Ten simple rules for designing and interpreting ERP experiments University of Iowa. In: Handy TC, editor. *Event related potentials: a methods handbook*. Cambridge, MA: MIT Press; 2004.
62. Csibra G, Kushnerenko E, Grossmann T. Electrophysiological methods in studying infant cognitive development. In: Nelson CA, Luciana M, editors. *Handbook of developmental cognitive neuroscience*. Cambridge, MA: MIT Press; 2008. p. 1–50.
63. Hileman CM, Henderson H, Mundy P, Newell L, Jaime M. Developmental and individual differences on the P1 and N170 ERP components in children with and without autism. *Dev Neuropsychol*. 2013;36(2):214–36.
64. Itier RJ. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cereb Cortex*. 2004;14(2):132–42.
65. Courchesne E, Ganz L, Norcia A M. Event-related brain potentials to human faces in infants. *Child Dev*. 1981;52(3):804–11.
66. Dawson G, Webb SJ, McPartland J. Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. *Dev Neuropsychol*. 2005;27(3):403–24.
67. de Haan M, CA N. Brain activity differentiates face and object processing in 6-month-old infants. *Dev Psychol*. 1999;35(4):1113–21.
68. Elsabbagh M, Volein A, Csibra G, Holmboe K, Garwood H, Tucker L, et al. Neural correlates of eye gaze processing in the infant broader autism phenotype. *Biol Psychiatry*. 2009;65(1):31–8.
69. Johnson MH, Griffin R, Csibra G, Halit H, Farroni T, de Haan M, et al. The emergence of the social brain network: evidence from typical and atypical development. *Dev Psychopathol*. 2005;17(3):599–619.
70. Balas BJ, Nelson CA, Westerlund A, Vogel-Farley V, Riggins T, Kuefner D. Personal familiarity influences the processing of upright and inverted faces in infants. *Front Hum Neurosci*. 2010;4(February):1.
71. Bretherton I. The origins of attachment theory: John Bowlby and Mary Ainsworth. *Dev Psychol*. 1992;28(5):759–75.
72. Tronick EZ, Cohn JF. Infant-mother face-to-face interaction: age and gender differences in coordination and the occurrence of miscoordination. *Child Dev*. 1989;60(1):85.

73. Harlow HF, Zimmermann RR. Affectional response in the infant monkey: orphaned baby monkeys develop a strong and persistent attachment to inanimate surrogate mothers. *Science*. 1959;130(3373):421–32.
74. Mundy P, Card J, Fox N. EEG correlates of the development of infant joint attention skills. *Dev Psychobiol*. 2000;36:325–38.
75. Charman T. Why is joint attention a pivotal skill in autism? *Philos Trans R Soc Lond Ser B Biol Sci*. 2003;358(January):315–24.
76. Morgan B, Maybery M, Durkin K. Weak central coherence, poor joint attention, and low verbal ability: independent deficits in early autism. *Dev Psychol*. 2003;39(4):646–56.
77. Striano T, Reid VM, Hoehl S. Neural mechanisms of joint attention in infancy. *Eur J Neurosci*. 2006;23(10):2819–23.
78. Lachat F, Hugueville L, Lemaréchal J-D, Conty L, George N. Oscillatory brain correlates of live joint attention: a dual-EEG study. *Front Hum Neurosci*. 2012;6(June):156.
79. Hopkins WD, Taglialatela JP. Initiation of joint attention is associated with morphometric variation in the anterior cingulate cortex of chimpanzees (*Pan troglodytes*). *Am J Primatol*. 2013;75(5):441–9.
80. Charman T, Baron-Cohen S, Swettenham J, Baird G, Cox A, Drew A. Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cogn Dev*. 2000;15(4):481–98.
81. Bakeman R, Adamson LB. Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Dev*. 1984;55(4):1278–89.
82. Kopp F, Lindenberger U. Effects of joint attention on long-term memory in 9-month-old infants: an event-related potentials study. *Dev Sci*. 2011;14(4):660–72.
83. Striano T, Reid VM. Social cognition in the first year. *Trends Cogn Sci*. 2006;10(10):471–6.
84. Hirotani M, Stets M, Striano T, Friederici AD. Joint attention helps infants learn new words: event-related potential evidence. *Neuroreport*. 2009;20(6):600–5.
85. Wimmer H, Perner J. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*. 1983;13(1):103–28.
86. Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a “theory of mind”? *Cognition*. 1985;21(1):37–46.
87. Auer DP. Spontaneous low-frequency blood oxygenation level-dependent fluctuations and functional connectivity analysis of the “resting” brain. *Magn Reson Imaging*. 2008;26(7):1055–64.
88. Grosse Wiesmann C, Schreiber J, Singer T, Steinbeis N, Friederici AD. White matter maturation is associated with the emergence of theory of mind in early childhood. *Nat Commun*. 2017;8:14692.
89. Premack D, Woodruff G. Does the chimpanzee have a theory of mind. *Behav Brain Sci*. 1978;1:515–26.
90. Perner J, Roessler J. From infants' to children's appreciation of belief. *Trends Cogn Sci*. 2012;16:519–25.
91. Baillargeon R, Scott RM, He Z. False-belief understanding in infants. *Trends Cogn Sci*. 2010;14(3):110–8.
92. Choi YJ, Luo Y. 13-Month-olds' understanding of social interactions. *Psychol Sci*. 2015;26(3):274–83.
93. Kovács ÁM, Téglás E, Endress AD. The social sense: susceptibility to others' beliefs in human infants and adults. *Science*. 2010;330(6012):1830–4.
94. Southgate V, Senju A CG. Action anticipation through attribution of false belief by 2-year-olds. *Psychol Sci*. 2007;18(7):587–92.
95. SAJ B, Bernstein DM. What can children tell us about hindsight bias: a fundamental constraint on perspective-taking? *Soc Cogn*. 2007;25(1):98–113.
96. Bloom P, German TP. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*. 2000;77:25–31.

97. Surian L, Caldi S, Sperber D. Attribution of beliefs by 13-month-old infants. *Psychol Sci*. 2007;18(7):580–6.
98. Moll H, Meltzoff AN. How does it look? Level 2 perspective-taking at 36 months of age. *Child Dev*. 2011;82(2):661–73.
99. Aichhorn M, Perner J, Kronbichler M, Staffen W, Ladurner G. Do visual perspective tasks need theory of mind? *NeuroImage*. 2006;30(3):1059–68.
100. Moll H, Tomasello M. Level 1 perspective-taking at 24 months of age. *Br J Dev Psychol*. 2006;24(3):603–13.
101. Hamilton AF de C, Brindley R, Frith U. Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*. 2009;113(1):37–44.
102. Moll H, Kadipasaoglu D. The primacy of social over visual perspective-taking. *Front Hum Neurosci*. 2013;7(September):558.
103. Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R. Deconstructing and reconstructing theory of mind. *Trends Cogn Sci*. 2015;19(2):65–72.
104. Dunbar RIM, Shultz S. Evolution in the social brain. *Science*. 2007;317(5843):1344–7.
105. Fehr E, Fischbacher U. Third-party punishment and social norms. *Evol Hum Behav*. 2004;25(2):63–87.
106. Camerer CF, Fehr E. When does “economic man” dominate social behavior? *Science*. 2006;311(5757):47–52.
107. Krueger F, Grafman J, McCabe K. Neural correlates of economic game playing. *Philos Trans R Soc Lond Ser B Biol Sci*. 2008;363(1511):3859–74.
108. Lee D. Game theory and neural basis of social decision making. *Nat Neurosci*. 2008;11(4):404–9.
109. Johnson ND, Mislin AA. Trust games: a meta-analysis. *J Econ Psychol*. 2011;32(5):865–89.
110. Camerer CF, Loewenstein G, Prelec D. Neuroeconomics: How neuroscience can inform economics. *J Econ Lit*. 2005;43(1):9–64.
111. Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci*. 2006;7(4):268–77.
112. McCabe K, Houser D, Ryan L, Smith V, Trouard T. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc Natl Acad Sci U S A*. 2001;98:11832–5.
113. Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD. The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*. 2004;22(4):1694–703.
114. Delgado MR, Frank RH, Phelps EA. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat Neurosci*. 2005;8:1611–8.
115. King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR. Getting to know you: reputation and trust in a two-person economic exchange. *Science*. 2005;308:78–83.
116. Delgado MR, Li J, Schiller D, E a P. The role of the striatum in aversive learning and aversive prediction errors. *Philos Trans R Soc Lond Ser B Biol Sci*. 2008;363(1511):3787–800.
117. Baumgartner T, Heinrichs M, Vonlanthen A, Fischbacher U, Fehr E. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*. 2008;58(4):639–50.
118. Zak PJ, Kurzban R, Ahmadi S, Swerdloff RS, Park J, Efremidze L, et al. Testosterone administration decreases generosity in the ultimatum game. *PLoS One*. 2009;4(12):e8330.
119. Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E. Oxytocin increases trust in humans. *Nature*. 2005;435(June):673–6.
120. Aspé-sánchez M, Moreno M, Rivera MI, Rossi A. Oxytocin and vasopressin receptor gene polymorphisms: role in social and psychiatric traits. *Front Neurosci*. 2016;9(January):510.
121. van den Bos W, Güroğlu B, van den Bulk BG, Rombouts SA, Crone E. Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Front Hum Neurosci*. 2009;3(December):52.
122. Mitchell JP. Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb Cortex*. 2008;18(2):262–71.
123. Billeke P, Boardman S, Doraiswamy PM. Social cognition in major depressive disorder: a new paradigm? *Transl Neurosci*. 2013;4(4):437–47.

124. Billeke P. The more I get to know you, the more I distrust you? Non-linear relationship between social skills and social behavior. *Front Psychiatry*. 2016;7:49.
125. de Vignemont F, Singer T. The empathic brain: how, when and why? *Trends Cogn Sci*. 2006;10:435–41.
126. Shenhav A, Botvinick MM, Cohen JD. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*. 2013;79(2):217–40.
127. Ebitz RB, Platt ML, Ebitz RB, Platt ML. Neuronal activity in primate dorsal anterior cingulate cortex signals task conflict and predicts adjustments in pupil-linked arousal Article Neuronal Activity in Primate Dorsal Anterior Cingulate Cortex Signals Task Conflict and Predicts Adjustments in Pu. *Neuron*. 2015;85(3):628–40.
128. Billeke P, Zamorano F, López T, Rodríguez C, Cosmelli D, Aboitiz F. Someone has to give in: theta oscillations correlate with adaptive behavior in social bargaining. *Soc Cogn Affect Neurosci*. 2014;9(12):2041–8.
129. Billeke P, Zamorano F, Cosmelli D, Aboitiz F. Oscillatory brain activity correlates with risk perception and predicts social decisions. *Cereb Cortex*. 2013;23(12):2872–83.
130. Ibáñez MI, Sabater-Grande G, Barreda-Tarrazona I, Mezquita L, López-Ovejero S, Villa H, et al. Take the money and run: psychopathic behavior in the trust game. *Front Psychol*. 2016;7(November):1–15.
131. Chang LJ, Smith A, Dufwenberg M, Sanfey AG. Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*. 2011;70(3):560–72.
132. Yoshimura S, Okamoto Y, Onoda K, Matsunaga M, Ueda K, Suzuki S, et al. Rostral anterior cingulate cortex activity mediates the relationship between the depressive symptoms and the medial prefrontal cortex activity. *J Affect Disord*. 2010;122(1–2):76–85.
133. Damasio AR, Grabowski TJ, Bechara A, Damasio H, Ponto LL, Parvizi J, et al. Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat Neurosci*. 2000;3:1049–56.
134. Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD. Empathy for pain involves the affective but not sensory components of pain. *Science*. 2004;303:1157–62.
135. Rilling JK, Sanfey AG. The neuroscience of social decision-making. *Annu Rev Psychol*. 2011;62:23–48.
136. Camerer CF. Behavioural studies of strategic thinking in games. *Trends Cogn Sci*. 2003;7:225–31.
137. Cherry T, Frykblom P, Shogren J. Hardnose the Dictator. *Am Econ Rev*. 2002;92(4):1218–22.
138. Moll J, Krueger F, Zahn R, Pardini M, de Oliveira-Souza R, Grafman J. Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc Natl Acad Sci U S A*. 2006;103(42):15623–8.
139. Wu S-W, Delgado MR, Maloney LT. The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *J Neurosci*. 2011;31(24):8822–31.
140. Hoffman E, McCabe K, Shachat K, Smith V. Preferences, property rights, and anonymity in bargaining games. *Games Econ Behav*. 1994;7:346–80.
141. Hutcherson CA, Bushong B, Rangel A. A neurocomputational model of altruistic choice and its implications. *Neuron*. 2015;87(2):451–62.
142. Raposo A, Vicens L, Clithero JA, Dobbins IG, Huettel SA. Contributions of frontopolar cortex to judgments about self, others and relations. *Soc Cogn Affect Neurosci*. 2011;6(3):260–9.
143. Kable JW, Glimcher PW. The neurobiology of decision: consensus and controversy. *Neuron*. 2009;63(6):733–45.
144. Ullsperger M, Fischer AG, Nigbur R, Endrass T. Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci*. 2014;18(5):259–67.
145. Rilling J, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C. A neural basis for social cooperation. *Neuron*. 2002;35:395–405.
146. McClure EB, Parrish JM, Nelson EE, Easter J, Thorne JF, Rilling JK, et al. Responses to conflict and cooperation in adolescents with anxiety and mood disorders. *J Abnorm Child Psychol*. 2007;35(4):567–77.

147. 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.
148. Spitzer M, Fischbacher U, Herrnberger B, Grön G, Fehr E. The neural signature of social norm compliance. *Neuron*. 2007;56(1):185–96.
149. Apps MAJ, Rushworth MFS, Chang SWC. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron*. 2016;90(4):692–707.
150. Shenhav A, Straccia MA, Botvinick MM, Cohen JD. Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cogn Affect Behav Neurosci*. 2016;16(6):1127–39.
151. Wittmann MK, Kolling N, Akaishi R, Chau BKH, Brown JW, Nelissen N, et al. Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat Commun*. 2016;7:12327.
152. Kolling N, Wittmann MK, Behrens TEJ, Boorman ED, Mars RB, Rushworth MFS. Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci*. 2016;19(10):1280–5.
153. Wittmann MK, Kolling N, Faber NS, Scholl J, Nelissen N, MFS R. Self-other mergence in the frontal cortex during cooperation and competition. *Neuron*. 2016;91(2):482–93.
154. Ruff CC, Fehr E. The neurobiology of rewards and values in social decision making. *Nat Rev Neurosci*. 2014;15(8):549–62.
155. Apps MAJ, Lesage E, Ramnani N. Vicarious reinforcement learning signals when instructing others. *J Neurosci*. 2015;35(7):2904–13.
156. Abu-Akel A, Shamay-Tsoory S. Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia*. 2011;49(11):2971–84.
157. Schurz M, Radua J, Aichhorn M, Richlan F, Perner J. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci Biobehav Rev*. 2014;42:9–34.
158. Saxe R, Xiao D-K, Kovacs G, Perrett DI, Kanwisher N. A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*. 2004;42(11):1435–46.
159. Billeke P, Zamorano F, Chavez M, Cosmelli D, Aboitiz F. Functional network dynamics in alpha band correlate with social bargaining. *PLoS One*. 2014;9(10):e109829.
160. Billeke P, Armijo A, Castillo D, López T, Zamorano F, Cosmelli D, et al. Paradoxical expectation: oscillatory brain activity reveals social interaction impairment in schizophrenia. *Biol Psychiatry*. 2015;78(6):421–31.
161. Corbetta M, Patel G, Shulman GL. The reorienting system of the human brain: from environment to theory of mind. *Neuron*. 2008;58(3):306–24.
162. 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;113:12574–9.
163. Carter RM, Bowling DL, Reeck C, S a H. A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science*. 2012;337(6090):109–11.
164. Charness G, Rabin M. Understanding social preferences with simple tests. *Q J Econ*. 2002;117(3):817–69.