

ADVANCED REVIEW

Neurocomputations of strategic behavior: From iterated to novel interactions

Yaomin Jiang  | Hai-Tao Wu  | Qingtian Mi  | Lusha Zhu 

School of Psychological and Cognitive Sciences, Beijing Key Laboratory of Behavior and Mental Health, IDG/McGovern Institute for Brain Research, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China

Correspondence

Lusha Zhu, School of Psychological and Cognitive Sciences, Beijing Key Laboratory of Behavior and Mental Health, IDG/McGovern Institute for Brain Research, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.
Email: lushazhu@pku.edu.cn

Funding information

NSFC, Grant/Award Number: 32071095; Peking-Tsinghua Center for Life Sciences

Edited by: Joe Kable, Editor

Abstract

Strategic interactions, where an individual's payoff depends on the decisions of multiple intelligent agents, are ubiquitous among social animals. They span a variety of important social behaviors such as competition, cooperation, coordination, and communication, and often involve complex, intertwining cognitive operations ranging from basic reward processing to higher-order mentalization. Here, we review the progress and challenges in probing the neural and cognitive mechanisms of strategic behavior of interacting individuals, drawing an analogy to recent developments in studies of reward-seeking behavior, in particular, how research focuses in the field of strategic behavior have been expanded from adaptive behavior based on trial-and-error to flexible decisions based on limited prior experience. We highlight two important research questions in the field of strategic behavior: (i) How does the brain exploit past experience for learning to behave strategically? and (ii) How does the brain decide what to do in novel strategic situations in the absence of direct experience? For the former, we discuss the utility of learning models that have effectively connected various types of neural data with strategic learning behavior and helped elucidate the interplay among multiple learning processes. For the latter, we review the recent evidence and propose a neural generative mechanism by which the brain makes novel strategic choices through simulating others' goal-directed actions according to rational or bounded-rational principles obtained through indirect social knowledge.

This article is categorized under:

Economics > Interactive Decision-Making

Psychology > Reasoning and Decision Making

Neuroscience > Cognition

KEYWORDS

decision neuroscience, game theory, social cognition, strategic behavior

Yaomin Jiang and Hai-Tao Wu contributed equally to this study.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *WIREs Cognitive Science* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Strategic interactions are ubiquitous among social animals, including humans (Camerer, 2011; Clutton-Brock, 2009; Darwin, 2004; Tomasello, 2019). We play poker, invest in others' risky businesses, and collaborate on scientific projects. One key feature shared by this broad range of interactions is that an individual's payoff depends not only on their own decision but also on the choices of other intelligent agents motivated by their respective personal goals (Osborne & Rubinstein, 1994). Therefore, to behave strategically, a goal-directed individual needs to not only account for potential reward and punishment in the environment, but also anticipate others' actions and/or intentions in a given context—such as predicting the card that the other poker player will select, or when and how much the trustee will repay the investment. Strategic behavior has been the subject of intense study in game theory (Camerer, 2011; Osborne & Rubinstein, 1994), evolutionary biology (Nowak, 2006; Weibull, 1997), sociology (Easley & Kleinberg, 2010; Moulin, 1986), political science (Farquharson, 1969; Stewart et al., 2019), cognitive and decision neuroscience (Glimcher & Fehr, 2014; Rilling & Sanfey, 2011), multi-agent artificial intelligence (Parkes & Wellman, 2015; Schrittwieser et al., 2020), and other related fields.

In cognitive and decision neuroscience, research has been focusing on establishing quantitative, neurobiologically plausible frameworks to decompose the intertwining cognitive operations underlying strategic choices observed in social animals and quantify their related neural processes (Camerer & Hare, 2014; Konovalov & Ruff, 2022; Lee & Seo, 2015; Montague & Lohrenz, 2007). Over the past decade, notable advances have been made in connecting brain circuits implicated in reward processing (e.g., the orbitofrontal cortex [OFC], striatum, etc.) and theory of mind (ToM; e.g., the temporoparietal junction [TPJ], dorsomedial prefrontal cortex [dmPFC], etc.) with behaviors in a variety of stylized, multiplayer games adapted from game theory and evolutionary biology (Bhatt et al., 2012; Hampton et al., 2008; Haroush & Williams, 2015; Ong et al., 2020; Seo et al., 2014). Central to this interdisciplinary approach is the role of computational models of learning, particularly variants of reinforcement learning (RL) models (Schultz et al., 1997; Sutton & Barto, 1998), which formalize how prior experience dynamically shapes other-predictive signals during repeated interactions at behavioral and neural levels. For example, trust building has been characterized as a neuro-behavioral process of social reward reinforcement in analogous to reinforcement implicated in simple conditioning tasks (King-Casas et al., 2005; Montague et al., 2015). This line of research has been greatly inspired by and developed in parallel with the progress in applying the RL framework to associate trial-and-error learning and goal-directed behavior with the underlying neurobiology, such as dopaminergic functions (Lee & Seo, 2015; see also Dayan & Abbott, 2001; Glimcher, 2011; and Niv, 2009 for reviews on RL-related research). Importantly, studies applying formal models of strategic learning have begun to elucidate the interplay between multiple learning systems (Hill et al., 2017; Kikumoto & Mayr, 2019; Tervo et al., 2014; Zhu et al., 2012, 2019), echoing the development in more basic decisions involving reward and punishment using model-free and model-based learning (Daw et al., 2011; Gläscher et al., 2010).

Despite its successes, this learning approach provides an incomplete account of strategic interactions, especially those in novel, one-shot situations. Humans excel at interacting with strangers in new environments by merely imagining or anticipating the actions or intentions of others (Baker et al., 2017; Camerer, 2011; Koster-Hale & Saxe, 2013). Such anticipation requires the ability to go beyond forming expectations of social partners by repeatedly interacting with the same partner within the same decision context. Consider a simple example of social signaling. Two individuals, Alice and Bob, meet in a room, and Alice points in the direction of a window for Bob. Bob, who has never been in this room or met Alice before, would immediately realize that Alice intends for him to look outside the window. This simple pointing-and-understanding interaction is strategic in the sense that the successful delivery of Alice's intention depends both on how Alice would choose a gesture to signal her intention and on what information Bob would recover from the observed gesture.

Although seemingly effortless, this simple form of strategic behavior has been considered uniquely human and involves complex, intertwined cognitive operations (Tomasello, 2010, 2019). First, it has been proposed that, as in many repeated strategic interactions, the pointing-and-understanding interaction may require recursive inferences of one's beliefs about others (Clark, 1996; Goodman & Frank, 2016; Noveck & Reboul, 2008; Tomasello, 2010). For example, one possibility is that accessing Alice's intention would require the understanding from Bob that Alice is pointing to inform him (Alice is cooperative), and that Alice believes that Bob will try hard to infer why Alice thinks the gesture would be informative for him (Alice believes that Bob is also cooperative) (Grice, 1975; Tomasello, 2010). More importantly, this inferential process is context-dependent: When the room is hot, for instance, Bob would correctly interpret that Alice wants him to open the window, rather than looking outside. Such flexible inferences potentially involve building an internal generative model of the social partner that maps from the potential cause (e.g., the intention of

asking Bob to look outside or open the window) to the observed consequence (e.g., pointing to the window), conditional on the specific decision context that one may have never directly encountered before. How, then, do individuals extract relevant information from social contexts in service of strategic reasoning, and what neural systems are involved in representing and organizing the extracted information? What computations underlie the internal simulation of others, and to what extent is such mental calculation specific to social and strategic behavior? Is the reasoning process based on a set of rules and heuristics that vary across contexts or some general assumptions regarding social behavior?

The ability to interpret and predict others' actions and/or cognitive states has long been associated with ToM (or mentalization) and investigated using experimental paradigms such as the false belief task (Wimmer & Perner, 1983). These paradigms make important contributions in revealing where (Jamali et al., 2021; Saxe et al., 2004) and how (Apps et al., 2013) the brain predicts others' false beliefs and how such process may go awry in clinical populations (Balsters et al., 2017). Despite their popularity, however, the classic false belief tasks are insufficient to capture mechanisms central to strategic behavior for various reasons. First, these tasks are usually *noninteractive*, leaving out a key component in strategic interactions which requires the prospective evaluation of how one's own decision would influence the mental states and choices of a *reactive* partner (Bhatt et al., 2012; Camerer et al., 2004; Hampton et al., 2008). Moreover, participants in the standard false belief tasks are typically instructed by the experimenters to infer other's mental states, raising questions on whether and how such instructed, explicit processes can support the voluntary, implicit mentalization central to strategic decisions (Bloom & German, 2000; Schaafsma et al., 2015). Indeed, prior research has shown that, in some strategic scenarios, self-reported beliefs deviate drastically from latent beliefs recovered from the choice behavior (Nyarko & Schotter, 2002), probably due to the differential levels of sophistication in instructed versus voluntary strategic reasoning.

Here, we propose that a constellation of methods from decision neuroscience and game theory, previously shedding light on strategic learning, may also have the potential to characterize some key aspects of strategic behavior in novel situations. First, recent developments in decision neuroscience have begun to elucidate neural and cognitive processes related to knowledge transfer, an important mechanism supporting flexible behavior including novel reward-seeking behavior (Baram et al., 2021; Bongioanni et al., 2021; Mark et al., 2020). Moreover, such new insights can be combined with or evaluated within the framework of game theory, which provides rich theoretical descriptions and empirical methods for examining novel, one-shot interactions (Camerer, 2011; Montague & Lohrenz, 2007; Osborne & Rubinstein, 1994). Thus, in addition to shedding light on strategic learning, these methods also have the potential to open a new avenue for studying whether and how the brain constructs generative models concerning others with value-based goals for choices in novel interpersonal settings.

In the remainder of this paper, we will review work on strategic interactions from the perspective of decision neuroscience. Rather than providing a comprehensive survey, we aim at offering an illustrative outline of key topics, models, evidence, and open questions in the field. We will first review recent progress in understanding adaptive strategic behavior, focusing on the multiplicity of learning systems and their interplay. We will then discuss strategic inferences and choices beyond learning paradigms, highlighting potential challenges and recent developments in elucidating the neurocomputations within novel strategic environments.

2 | STRATEGIC INTERACTIONS IN ITERATED CONTEXTS

Much of the previous research on strategic interactions has been focusing on learning paradigms (Fudenberg & Levine, 1998), where agents interact with one another repeatedly and adjust their behavior according to the outcomes and feedbacks arising from past interactions (Abe & Lee, 2011; Ferreira et al., 2021; Hampton et al., 2008; King-Casas et al., 2005; Ong et al., 2020; Park et al., 2019; Ray et al., 2008; Yoshida et al., 2008; Yoshida, Seymour, et al., 2010; Zhu et al., 2012). Learning paradigms provide a suitable starting point for probing neurocognitive operations of strategic behaviors for a number of reasons. The first involves the widespread success of the RL framework in explaining neural activity in value-based decisions (Dayan & Abbott, 2001; Glimcher, 2011; Montague et al., 2006; Niv, 2009; Schultz et al., 1997). In social and strategic settings, accumulating evidence suggests the involvement of simple RL-like behavior, even in situations where RL is suboptimal (Camerer & Ho, 1999; Erev & Roth, 1998; Seo et al., 2014; Zhu et al., 2012). Moreover, while strategic interactions often require agents to form beliefs and mental models of other individuals, extant data suggest that at least some forms of such inferential and updating processes can be characterized at the algorithmic level in a manner similar to more sophisticated RL algorithms, such as those involving model-based components (Behrens et al., 2008; Burke et al., 2010; Dunne & O'Doherty, 2013; Hampton et al., 2008; Suzuki et al., 2012; Zhu et al., 2012). Significant progress has been made over the past decade in identifying multiple learning

systems involved in repeated strategic decisions. Important research questions have therefore been proposed and actively investigated, including the necessity of these learning systems for strategic behavior, and how the brain arbitrates between different systems within a decision and across a sequence of decisions.

2.1 | Learning systems in repeated strategic behavior

First, we will review algorithms widely implicated in strategic learning, including reward learning, belief learning, and mental-state learning (Figure 1), drawing an analogy at the algorithmic level with multiple RL systems in more basic reward-seeking situations (Daw, 2018; Gläscher et al., 2010; Kool et al., 2018; Montague et al., 2006; Niv, 2009). Although equally important, social learning without strategic considerations (e.g., observational learning; Burke et al., 2010; Charpentier et al., 2020; Suzuki et al., 2012) is excluded from this review, given the potential differences in the underlying processes when one's payoff is determined jointly by the actions of multiple agents or uniquely by one's own action.

2.1.1 | Reward learning

There is much evidence that strategic agents learn about reward and punishment arising from past interactions in a manner mimicking model-free RL in nonsocial settings (i.e., reward learning, Figure 1). Computationally, strategic agents may treat candidate actions as different slot machines, like in the “multi-armed bandit” task widely used in reward learning studies in nonsocial settings (Daw et al., 2006; Sutton & Barto, 1998), and update the estimated value of each action in a temporal difference form (Montague et al., 2006; Schultz et al., 1997; Sutton & Barto, 1998), without considering what others have or have not selected (Camerer, 2011; Erev & Roth, 1998; Fudenberg & Levine, 1998).

At the neural level, neuroimaging studies on various strategic interactions indicate that the reward prediction error (RPE)—a key learning signal reflecting the discrepancy between expected and received rewards—is tracked by activity in the reward circuits (van den Bos et al., 2013; Zhu et al., 2012) previously implicated in representing RPE in more basic non-strategic environments (O'Doherty et al., 2003, 2004; Pagnoni et al., 2002; Pessiglione et al., 2006). Interestingly, the neural encoding of RPE during strategic learning not only reflects the update of predictive value in monetary gains and losses, but is also associated with social values, such as the prediction error related to defeating or being defeated by an opponent in competitive scenarios (Hertz et al., 2017; Ligneul et al., 2016; see also Lockwood & Klein-Flügge, 2021; Zhang et al., 2020 for helpful introductions and tutorials on how to study social learning using computational models of RL and functional magnetic resonance imaging [fMRI]).

2.1.2 | Belief learning

Notably, the reward learning algorithm provides an incomplete account of strategic learning, as reward-learning agents have no notion of the game structure or others' actions and inferences, and can consequently be exploited by more

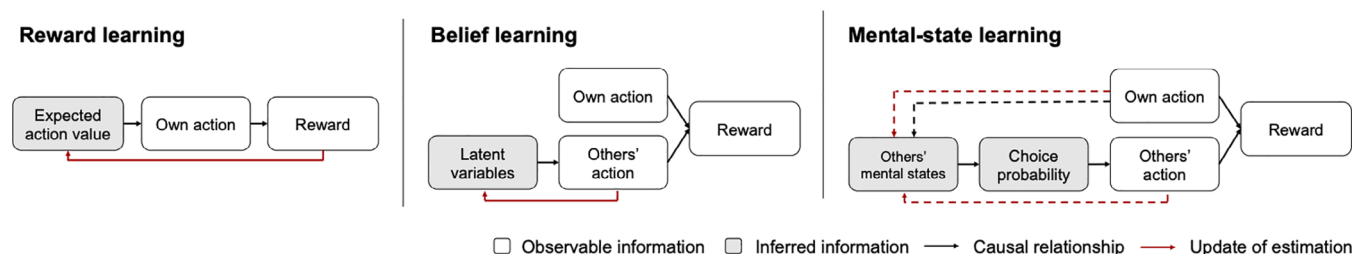


FIGURE 1 Illustration of strategic learning algorithms. To provide a better comparison against other learning rules, only one of the belief learning strategies (i.e., action-based learning) is illustrated here, although belief learning can be achieved by differential algorithmic processes. Mental-state learning involves learning about various mental states, including but not limited to the “states” of others' preferences (Ferreira et al., 2021), pro-sociality tendency (Ray et al., 2008), and the level of sophistication in strategic reasoning (Yoshida et al., 2008). The dashed lines reflect potential relationships that may or may not exist depending on the specific interactive scenario

sophisticated opponents (Camerer, 2011; Fudenberg & Levine, 1998; Hampton et al., 2008; Seo et al., 2014; Zhu et al., 2012). Researchers have proposed an array of higher-level learning models that characterize the underlying processes by which individuals build and update the causal structure that gives rise to others' choice behavior. Based on the sophistication of algorithms, these models can be roughly classified as belief learning and mental-state learning (Figure 1).

Belief learning requires agents to form and update first-order beliefs regarding the likelihood of future actions of other individuals through experience. For example, under the belief learning assumption, a goalkeeper facing penalty kicks assumes that the direction of the next kick taken by a certain soccer player should follow a static and unknown probabilistic distribution, which can be inferred from the past kicks of the same player. Computationally, this can be achieved by learning either directly about the frequency of others' actions, or from one's received and could-have-received reward (i.e., payoffs associated with the chosen and unchosen actions), conditional on the action selected by the strategic partner (Camerer & Ho, 1999). The belief learning algorithm is mathematically related to model-based RL (Daw et al., 2005) under mild mathematical constraints, as it enables an individual to learn the predictive value of a specific action at a certain state (Hunter et al., 2021).

At the neural level, belief learning in strategic settings recruits a partially overlapping but distinct set of brain regions in comparison to model-based RL in nonsocial setups. Key belief learning signals such as the first-order belief about others' actions (expectation) or its discrepancy from observations (prediction error) have been identified in the ventral striatum (Zhu et al., 2012), rostral and middle parts of the anterior cingulate cortex (ACC) (Park et al., 2019; Zhu et al., 2012), and TPJ (Park et al., 2019) in human neuroimaging studies. In nonhuman primates, neurons within a number of prefrontal regions have been implicated in encoding signals with a belief-learning flavor in both competitive and cooperative environments. For example, in a competitive game of rock-paper-scissors, rhesus monkeys display behavioral signatures of belief learning when playing against a computer-simulated opponent (Abe & Lee, 2011). Neurons in the OFC and dorsolateral prefrontal cortex (dlPFC) are found to encode reward associated with both the chosen and unchosen actions (conditional on the opponent's behavior), consistent with the prediction of belief learning algorithms (Abe & Lee, 2011). In a cooperative game of the prisoner dilemma, neurophysiological recordings also demonstrate that neurons in monkeys' dorsal anterior cingulate carry signals predictive of an opponent's yet unknown decision to cooperate or defect during repeated interactions (Haroush & Williams, 2015). Disrupting cingulate activity selectively reduces monkeys' tendency to cooperate following a cooperative choice of the opponent (Haroush & Williams, 2015).

2.1.3 | Mental-state learning

At the core of belief learning is the assumption that others' choices follow a fixed but unknown distribution, yet these models are agnostic about why such distribution takes a particular form. In contrast, algorithms of mental-state learning explicitly model the process that generates the distribution of others' choices, by building an internal model based on the learned statistical relationship on how others' unobservable mental states would contribute to their observable choice behavior. Mental-state learning can be considered as a counterpart of the more sophisticated model-based RL algorithms, with the "model" containing hierarchical causal relationships regarding the action selection process of strategic partners (Figure 1). These mental-state learning models sometimes account for the prospective influence of one's decision on others' mental states based on the learned relationships. For example, in a repeated rock-paper-scissors game, a player may engage in prospective reasoning by mentalizing how the other player would predict her future behavior following her taking a particular action (second-order belief).

At the neural level, regions related to the TPJ have been implicated in representing signals important for mental-state learning. In a simple two-player competitive game, Hampton et al. (2008) formalize the process of mental-state learning in the context of second-order beliefs. Their model captures not only how an individual responds to the opponent's action (first-order belief), but also how an individual predicts, recursively, the influence of her own action on the opponent's behavior (second-order belief). The fMRI signals related to the update of the second-order belief, a key mental-state learning signal, are identified in the TPJ/superior temporal sulcus (STS) in a temporal difference form at the time of feedback. Consistent with this finding, a recent study of a repeated coordination game demonstrates that monkeys' choices are best explained by a learning algorithm that entails modeling the opponent's intention to take certain actions (Ong et al., 2020). Specifically, the study associates signals related to mental-state updates with signals in neurophysiological recordings in monkeys' middle STS. This region has been widely implicated in social perception in

macaques (Sliwa & Freiwald, 2017; Wittmann et al., 2018) and, more recently, proposed as a potential homolog of the human TPJ based on profiles of resting-state connectivity in macaques and humans (Mars et al., 2013). The finding of the middle STS signals, therefore, supports the potential functional overlap between the human TPJ and the primate middle STS, raising intriguing questions regarding the extent to which high-level strategic learning is shared by other social species (see Box 1 for discussions on the specificity of social and strategic learning processes).

Of note, while we draw the analogy based on the algorithmic similarity between RL and strategic learning models, we do not mean that these models are completely identical or that they share the entire internal representations of the inputs, outputs, or intermediate learning signals. Rather, we hope the analogy would help connect the high-level logic and relational similarity embedded in these models and promote integration of ideas and methods across domains. For example, a recent study (Hunter et al., 2021) leverages the computational psychiatric methods developed within the model-free and model-based RL framework to examine the relative use of reward and belief learning by individuals

BOX 1 The specificity of social and strategic learning processes

There has been much discussion on whether and how neural and cognitive substrates identified in strategic learning are unique to social behavior. This discussion has its basis in a long-standing social brain hypothesis (Dunbar, 1998), postulating that certain brain structures have been specifically shaped by evolutionary adaptation to social interactions. For example, a recent study examines social specificity of the TPJ and dmPFC in a mental-state learning task with matching nonsocial conditions and suggests that the TPJ is related to the context-dependent processing of learning outcomes, whereas the dmPFC is involved in implementing learning strategies in response to reactive opponents, regardless of whether the opponent is a human or computer (Konovalov et al., 2021). This study supports the idea that different “social brain” regions may be related to disparate operations in service of strategic interactions, and some of these operations may be domain general.

More broadly, it has been proposed that social specificity in strategic learning (and social behavior in general) should be examined at multiple levels of analyses, rather than being treated as a single-dimension comparison (Lockwood et al., 2020). Belief learning in strategic behavior, for example, is algorithmically similar to model-based learning in nonsocial contexts, yet it is associated with differential neural substrates from model-based learning (Apps et al., 2016; Dunne & O’Doherty, 2013; Hunter et al., 2021; Lockwood et al., 2020; Wittmann et al., 2018; Zhu et al., 2012), suggesting dissociable levels of specificity in cognitive operations and neural implementations (see also Lockwood et al., 2020 for more examples). Future studies are needed to address the fine-grained nature of social specificity possibly by combining causal approaches like brain lesions and noninvasive brain stimulations with experiments carefully designed to contrast the need for separate cognitive operations in strategic versus nonstrategic interactions.

BOX 2 Confidence and strategic decisions

In social neuroscience such as studies related to observational learning and other nonstrategic social decisions, there has been growing interest in how social behavior interacts with confidence (Campbell-Meiklejohn et al., 2017; De Martino et al., 2017; Fisher & Oppenheimer, 2021; Pescetelli & Yeung, 2021; Soldà et al., 2021). Although there has not been much research on this topic in strategic contexts, strategic interactions may offer a unique test bed for examining the role of confidence in social behavior. Behavioral evidence suggests that confidence can be used as a signal to manipulate other’s choices or mental states (Charness et al., 2017; Hertz et al., 2017) and therefore may be exaggerated during strategic interactions, especially the competitive ones. This phenomenon leads to exciting questions related to the actual, expressed, and perceived confidence of self and other, such as how differential metacognitive signals are computed, represented, used, and updated in the brain. Future studies may explore whether these confidence signals are produced in the brain in a context-dependent manner varying across competitive versus cooperative scenarios, and how confidence signals would contribute to aggressive versus conservative choices in strategic contexts.

with social anxiety disorder. This combination of methods allows for showing that social anxiety selectively increases the tendency of deliberation during repeated strategic interactions. More broadly, the proposed integration of ideas and methods may provide insights to some important open questions that have been intensively investigated in the fields of social and decision neurosciences but remain relatively unexplored in strategic contexts. These include how confidence arises from and interacts with the processes of making strategic decisions (see Box 2) and how complex social contexts such as norms and culture affect the internal computation underlying strategic interactions (see Box 3).

2.2 | Multiplicity of learning systems in repeated strategic interactions

Previous research indicates that repeated strategic interactions may simultaneously involve a multiplicity of learning processes subserved by potentially dissociable neural systems. For example, neurobehavioral signatures of both belief and mental-state learning are identified within fixed pairs of human subjects during repeated competition that requires higher-order mentalization (Hampton et al., 2008). Similar hybrid learning behavior has also been observed in non-human primates in an iterated coordination game with a fixed primate partner (Ong et al., 2020). In comparison, when higher-order beliefs are experimentally controlled (e.g., instead of fixed pairing, repairing participants randomly on a trial-by-trial basis in learning experiments), strategic learning in competitive scenarios sometimes involves a hybrid of belief and reward learning, rather than mental-state learning (Zhu et al., 2012). Results of these studies are consistent with the general possibility that strategic learning behavior may be guided by inputs from multiple systems that learn from differential aspects of the feedback information.

How does the brain implement a hybrid of algorithms during strategic learning? One possibility is that the brain implements these algorithms in parallel, and integrates their outputs linearly or nonlinearly to guide subsequent decisions. In line with this hypothesis, neural correlates of key learning signals (e.g., prediction error) derived from different algorithms have been identified in distinct but partially overlapping brain areas: For example, striatal activity has been implicated in both reward and belief learning (van den Bos et al., 2013; Zhu et al., 2012), whereas TPJ/STS activity has been implicated in both belief and mental-state learning (Hampton et al., 2008; Ong et al., 2020; Park et al., 2019). Moreover, there is evidence that the trial-by-trial prediction error signals separately derived from the reward learning and belief learning models were integrated in the brain in a manner consistent with the experience-weighted attraction model (Zhu et al., 2012), a well-established strategic learning algorithm that assumes nonlinear integration over multiple learning rules (Camerer & Ho, 1999).

BOX 3 Strategic behavior in complex social contexts

Social interactions are embedded in complex hierarchy of contexts—from social norms and group culture to intricate structures in reward and social percepts (Baker et al., 2017; FeldmanHall & Nassar, 2021). The unparalleled scale and complexity in strategic behavior in humans raise the intriguing question regarding the extent to which the abstract strategic games discussed in the current review can capture features relevant to real-world interactions. This is particularly important given the rapid developments in computational psychiatry that has begun to probe social dysfunctions as well as their biomarkers using strategic games (King-Casas & Chiu, 2012; Kishida et al., 2010; Robson et al., 2020). The diagnostic utility of a neural measure of a certain psychiatric condition relies critically on how sensitive the measure is to latent functions relevant to everyday life.

A major challenge, however, lies in how to construct neurobiologically plausible models capable of capturing the complexity in multilayer reasoning that is central to strategic interactions. Even in individual learning that only involves states, outcomes, and policies of a single person, it is well known that the consideration of increasingly high-dimensional states will quickly lead to a combinatorial explosion and inefficient learning performances (Bellman, 1957). Such “curse of dimensionality” may be even severer in real-life strategic interactions that entail recursive interpersonal reasoning in a high-dimensional state space. Future research is needed to explore how to constrain vast state spaces resulted from complex hierarchy of contexts in an ecologically valid manner. Promising approaches may involve considering the limited cognitive hierarchies in strategic reasoning (Camerer et al., 2004) or using norms, conventions, and moral principles as the potential mechanisms for reducing dimensions or reorganizing the underlying state space (Hawkins et al., 2019; Niv, 2019; Radulescu et al., 2021).

The overlapping neural activation further raises questions regarding the functional relevance of these regions in their respective learning systems and the specific computational role of these regions in guiding strategic behavior. Recent studies have addressed these questions using causal experiments of either noninvasive brain stimulation or patients with focal brain lesions (Hill et al., 2017; Zhu et al., 2019). By disrupting a specific brain region while leaving others uninfluenced, these causal studies help establish the necessity and specificity of a brain area in the putative hybrid process, thereby providing key insights into the neurocomputational architecture of strategic learning.

For example, a recent study examined the choice behavior in patients with focal lesions in the basal ganglia (BG), a group of subcortical nuclei previously implicated in both reward and belief learning (Zhu et al., 2019). In a competitive game known to engage a hybrid of reward and belief learning, the study provides causal evidence that belief learning is dissociable from reward learning and does not rely on the integrity of the BG. Specifically, patients with focal lesions in the BG display intact learning in a strategic context where both learning models can be adopted, despite impairment in a nonstrategic context where reward learning can be adopted, but the social context for belief learning is absent. Preserved learning of patients with BG lesions in the strategic context may thus reflect compensation of the belief learning system that is possibly supported by intact prefrontal regions.

Using transcranial magnetic stimulation (TMS), other researchers provide evidence for the necessity of the right TPJ in mental-state learning in a hybrid process that involves both mental-state and belief learning (Hill et al., 2017). Widely implicated in ToM, the TPJ/STS is associated with both belief and mental-state learning (Hampton et al., 2008; Ong et al., 2020; Park et al., 2019). In particular, temporally disrupting the neural excitability in the right TPJ with continuous theta-burst TMS selectively reduces the extent to which participants rely on higher-order mental-state learning, yet it does not influence belief learning (Hill et al., 2017). Interestingly, more disrupted functional connectivity between the right TPJ and ventromedial prefrontal cortex (vmPFC) is associated with more decreased behavioral reliance on computations related to higher-order mentalization (Hill et al., 2017). These results suggest dissociable neural implementation for belief and mental-state learning: The latter relies on functions of the right TPJ and probably also its communications with the vmPFC, while the former does not.

2.3 | Arbitration of learning systems in repeated strategic interactions

Although hybrid learning models typically assume parallel implementation of multiple learning systems, it is also possible that the brain arbitrates among learning strategies according to how well a particular strategy performs or how one's internal state or external environment changes during learning. Similar strategy switches have been observed between model-free and model-based RL in the nonsocial domain based on the internal trade-off between the cost (e.g., cognitive demand) and benefit (e.g., the expected payoff of a recommended action) associated with the usage of model-free or model-based strategies (Kool et al., 2018; Otto et al., 2013).

Recent studies of strategic learning in humans, nonhuman primates, and rodents lend credence to this possibility, suggesting that the medial prefrontal cortex, including the ACC, may play a role in the arbitration process. For example, in a modified matching-pennies game, human participants switch between model-based and stochastic choices based on the outcome associated with the previous choice (Kikumoto & Mayr, 2019). Here, "model-based" choices are similar to those generated by belief learning that are based on the first-order prediction of opponents' actions. Such choices exploit detectable patterns in opponent's behavior but may also lead to potential losses when the prediction is not accurate enough or is counter-predicted by the opponent. According to this study, participants switch to more stochastic choices when losses occur, creating opportunities to abscond from suboptimal strategies and making their own choices less predictable for opponents. Mid-frontal EEG activity encoding information about the opponent's strategy at feedback also attenuates after losses compared with wins, consistent with a suppressed internal model of opponents' choices, which supports memory-free stochastic choices.

Primates and rodents also demonstrate the ability to switch between strategies as if they are performing cost-benefit analyses of different learning systems during strategic learning. In a computer-based competitive games, monkeys are found to show greater deviation from RL when such deviation is associated with greater gains (Seo et al., 2014). Neurons in monkeys' dmPFC show switch signals, whose strength is associated with the level of deviation from model-free RL. Yet it is not clear in this study what kind of learning model the animals switch to. Tervo et al. (2014) find that rats shift toward more stochastic choices when the opponent (computer) becomes more sophisticated (i.e., when the opponent choices are more difficult to predict and their own choices are more likely to be exploited). This behavioral pattern is consistent with the hypothesis of cost-mediated arbitration. Using circuit perturbations in transgenic rats, this study

demonstrates that the behavioral switches between strategies are casually related to inputs from the locus coeruleus to the ACC, such that enhancing (suppressing) the action of the noradrenergic system in the ACC leads to increased (decreased) stochastic behavior (Tervo et al., 2014).

3 | STRATEGIC INTERACTIONS IN NOVEL CONTEXTS

In addition to repeated interactions, humans often engage in strategic behavior in which they have no direct experience (Camerer, 2011; Tomasello, 2019). We ask strangers to watch our laptop in a café and coordinate with other commuters by walking on opposite sides of a road, even though we have never met those people, been to those places, or encountered the same set of choice options. Although these scenarios are widespread in the daily life, they are “novel” in the sense that such interactions often rely on transferred knowledge or general principles established from indirect experiences. Similar cases have been established in nonsocial domains such as how an iPhone user first switches to an Android system by leveraging her prior knowledge on smart phones (Griffiths et al., 2019), and how human can correctly predict the trajectories of moving objects in novel settings based on the intuition of physical laws (Battaglia et al., 2013; Fischer et al., 2016). In the social world, the ability to make decisions in novel situations may serve as a starting point for building long-term relationships and adapting to new norms and cultures. Despite being a powerful determinant of social life, its neural and cognitive mechanisms remain to be explored.

Recent development in value-based nonsocial decisions has begun to shed light on neurocognitive underpinnings of novel reward-seeking behavior in both human and nonhuman primates, highlighting potential mechanisms such as memory integration (Backus et al., 2016; Schlichting & Preston, 2015; Spalding et al., 2018; Zeithamova et al., 2012) and cognitive maps (Bongioanni et al., 2021). In the social domain, a handful of fMRI studies have provided evidence for map-like representations underlying social perception, which are thought to support the generalization of social judgments to novel situations (Behrens et al., 2018; Park et al., 2020; Park et al., 2021; Tavares et al., 2015). Much less is known, however, about whether similar neurocognitive substrates support novel strategic choices, even though social decisions are believed to share some of the same neural circuitry with individual value-based behavior (Behrens et al., 2009; Ruff & Fehr, 2014) and incorporate social perception capturing how people see others (Jenkins et al., 2018).

A fundamental question for novel strategic interactions is how individuals form accurate (or sometimes inaccurate) predictions of others in the absence of direct interacting experience. One possible mechanism involves constructing an internal generative model reasoning about actions, consequences, and/or relevant mental states of others, in substitute for models (or expectations) learned from repeated interactions. The notion of generative models has deep roots in artificial intelligence (Dayan et al., 1995; Hinton, 2007) and multiple domains in cognitive neuroscience, including visual perception (Knill & Richards, 1996) and motor control (Kording & Wolpert, 2006), and has been recently proposed to account for inferences related to nonstrategic social cognition (Jara-Ettinger et al., 2016; Koster-Hale & Saxe, 2013). Computationally, a generative model is defined as a probabilistic mapping from covert causes (e.g., an object in sight) to overt consequences (e.g., a retinal image) (Figure 2a) (Rao, 1999), which, when combined with hypotheses like active inference (Friston et al., 2017), provides a mechanism by which the brain constructs higher-level expectations about the external environment in service of interpreting incoming information and guiding behavioral responses.

In the context of strategic interaction, the putative generative process may capture how intertwining causes—intention, context, knowledge, and other mental states—would give rise to others' actions. One intriguing possibility is that our brain models others' choice processes as goal-directed, strategic decisions, based on some general assumptions of human behavior in social environments (Camerer, 2011; Goodman & Frank, 2016). For example, rational predictions may be generated based on the assumption that social partners would maximize their personal payoffs in a given context and the assumption that partners would also engage in similar generative reasoning about others. Such rational principles, which have deep roots in economics and game theory (Camerer, 2011; Glimcher & Fehr, 2014; Osborne & Rubinstein, 1994), may come from indirect social experiences through processes such as inductive generalization (Gershman & Niv, 2015; Tenenbaum et al., 2006) or meta-learning (Botvinick et al., 2019; Griffiths et al., 2019; Vilalta & Drissi, 2002) and guide the mental predictions of others, flexibly, contingent on the social setting.

As implied by the generative hypothesis, strategic decision-making requires the computation and representation of task-relevant contingencies between others' intentions and actions, in service of selecting the choice that is most likely to lead to the preferred outcome for the decision-maker. The putative causal mapping from strategic actions to outcomes echoes the long-standing hypothesis of the “cognitive map,” which is believed to support a range of flexible behaviors by capturing the complex and abstract relationships among task-related entities (e.g., cues, actions, and

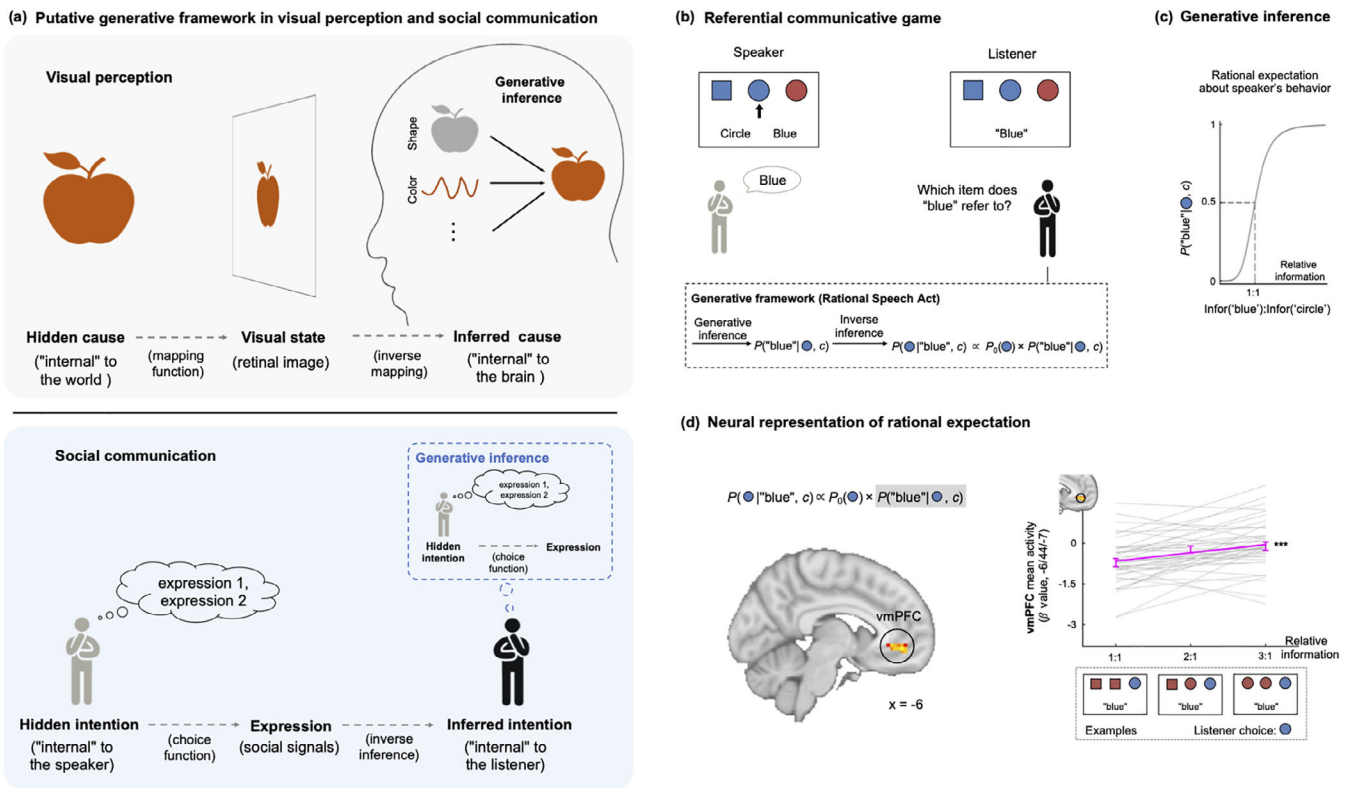


FIGURE 2 (a) Putative generative models in visual perception and social communication. In visual perception, the brain infers sensory causes (e.g., the apple in sight) from the bodily effects (e.g., the retinal image) by modeling the sensation-generating process and then inverting this model to derive the most probable cause of the sensation. For social communication, a listener decodes the hidden intention from a received expression by internally simulating the speaker's choice process. (b) Referential communication game involves random-matched pairs of a speaker and a listener. The speaker's goal is to refer to one of three objects presented in a context, by selecting an expression denoting either the color or the shape feature of the target. The listener, who faces the same context but does not know the target, is required to recover the target according to the expression received from the speaker. The computational model (Frank & Goodman, 2012; Franke & Degen, 2016) proposes an inverse inferential process, by which a listener simulates the intention-action contingency from the speaker's perspective and then inverts this process to identify the most probable target. (c) Listener's internal generative model predicts that, for best helping the audience recognize the intended meaning, speakers would compare between candidate expressions and select the expression that conveys the maximal amount of information in context. (d) Listener's vmPFC encodes the intention (e.g., referring to the blue circle)—action (e.g., selecting the expression "blue") contingencies of the speaker in a manner consistent with the model prediction, even in situations where such generative signals are not necessary for recovering the intention (e.g., three cases shown in the right panel). Adapted from Mi et al. (2021). vmPFC, ventromedial prefrontal cortex

outcomes) (Behrens et al., 2018; Boorman et al., 2021; Son et al., 2021). In strategic behavior, however, the putative generative processes have several unique features compared to those subserving more general goal-directed behavior. First, in contrast to many other reward-seeking behaviors, which only require access to one's own actions and outcomes, strategic decision-making entails putting oneself in other individuals' shoes and evaluating actions, payoffs, and other task-related features from others' perspectives. Failure in perspective taking may result in ineffective strategic decisions. Second, strategic generative processes can go beyond simple perspective taking and require cognitive operations that monitor and organize common knowledge (i.e., information shared among agents and known to be shared), as well as inferences about others' inferences (Camerer, 2011; De Freitas et al., 2019). Compared to nonstrategic social inferences (e.g., judging the intention of a movie character), strategic behavior likely requires multidimensional representations of task contingencies that involve joint actions and outcomes of agents involved in the interaction.

To characterize strategic generative processes in a quantitative and neurobiologically-plausible way, one potential method is to combine tools from decision neuroscience and game theory. A recent neuroimaging study offers an example of this method and initial neural evidence for the generative hypothesis in the context of interpersonal communication (Mi et al., 2021). This study investigates how the brain reads between the lines, or decodes intentions from communicative signals in context (Figure 2). Intentional communication offers an excellent test bed for studying novel

strategic behavior, not only because communication has long been considered a special form of a cooperative behavior between a speaker and a listener (Grice, 1975), but also because communicative decisions are often inherently novel, and achieving mutual understanding requires flexible choices and inferences against a near-infinite variety of communicative signals and contexts. The generative hypothesis in such communicative reasoning has been linked with the possibility that a listener recovers the intended meaning of an utterance by modeling the utterance generating process of the speaker in a given context. This study demonstrates that the listener's vmPFC, a region implicated in cognitive map representation (Behrens et al., 2018; Schuck et al., 2016; Wikenheiser & Schoenbaum, 2016; Wilson et al., 2014), encodes the probabilistic inference of what a rational speaker should say in order to convey the maximal amount of information given a communicative intention and context, in a manner consistent with the rational generative principle for speech act proposed in prior research (Franke & Degen, 2016; Goodman & Frank, 2016). In line with the proposed role of ToM, it is observed that the vmPFC representation of generative inferences is supported by signals from the TPJ and dmPFC, regions previously implicated in ToM. This case demonstrates a potential link between brain regions implicated in cognitive map representation and novel strategic inferences. It also illustrates how the use of behavioral models based on rational generative assumptions can be decisive in allowing researchers to directly test mechanistic hypotheses regarding how the brain constructs causal inferences flexibly in context.

In addition to predicting speakers' choices based on rational assumptions, it is also possible that the listener's brain generates an utterance selection process based on certain heuristics or bounded-rational assumptions of

BOX 4 Pro-sociality in strategic interactions

Pro-sociality has been intensively studied in the field of social decision neuroscience, including social exchanges that involve other-regarding considerations like generosity, fairness, reciprocity, and so on (Camerer, 2011; Fehr & Camerer, 2007; Fehr & Fischbacher, 2003; Fehr & Krajbich, 2014; see also Tusche & Bas, 2021 for guidelines of modeling altruistic decision-making processes). Much of this research rests on the assumption that prosocial behavior is driven by some internal preferences for other's well-being (aka other-regarding preferences), which have been associated with neural signals in reward circuits during decision-making processes (Ruff & Fehr, 2014). In strategic interactions, social preferences have been proposed to influence choices directly and indirectly.

For direct influences, there is growing evidence that choosing among different options in strategic interactions involves computing a subjective value for each option, taking into account not only self-interests but also the potential gains and losses of the interacting partner (Camerer, 2011; Fehr & Camerer, 2007; Fehr & Fischbacher, 2003; Fehr & Krajbich, 2014). For example, there is much evidence that, by influencing the subjective valuation process, other-regarding preferences underlie the egalitarian proposals observed in ultimatum games (Henrich et al., 2001), reciprocation in trust games (King-Casas et al., 2005), and cooperation in the public goods games (Park et al., 2019). Altruistic behavior in these games has been associated with activity in the ventral tegmental area, ventral striatum, and OFC (Decety & Yoder, 2017; Fehr & Krajbich, 2014; Rilling & Sanfey, 2011). Focal lesion to the OFC and noninvasive brain stimulation to the lateral prefrontal cortex are also found to be able to change altruistic strategic behavior in some strategic games (Krajbich et al., 2009; Ruff et al., 2013).

Social preferences may also play an indirect role, altering how one thinks about others during interpersonal interactions. For example, in public goods games, the anticipation of nonzero contribution from other players promotes cooperation and has been associated with activity in regions including the ACC and TPJ (Camerer, 2011; Fehr & Fischbacher, 2003; Park et al., 2019). Beside such first-order beliefs, social preferences may also affect strategic considerations by influencing second-order beliefs, which characterize the inference of other's inferences. For example, in the trust game, the discrepancy between a trustee's belief of how much the investor expects her to return and her actual return may be associated with aversive social emotions such as guilt, and is thought as a potential mechanism for promoting cooperation (Chang et al., 2011; Nihonsugi et al., 2015). Signals related to such second-order beliefs are observed in areas including the insula, dorsolateral and medial prefrontal cortex (Chang et al., 2011), and altering the activity in the dlPFC changes cooperative behavior (Nihonsugi et al., 2015).

communication. For example, rather than assuming the optimal delivery of information, a listener may consider the potential cognitive cost associated with searching for the most appropriate expression in context, thereby constraining the generative model by the assumption regarding limited cognitive resources (Lieder & Griffiths, 2019).

More broadly, substantial evidence has suggested that human behaviors often deviate from optimality in a systematic and predictable manner, raising intriguing questions regarding whether and how the internal generative processes would incorporate bounded rationality. For example, the brain may consider decision factors beyond self-payoff maximization, accounting for other-regarding incentives (e.g., preferences for fairness or generosity) based on social norms established among strategic players (de Quervain et al., 2004; Fehr & Camerer, 2007; Fehr & Fischbacher, 2004; Spitzer et al., 2007; see Zoh et al., 2022 for more discussions on norms under cooperative contexts; see also Box 4). It is also possible that, rather than assuming that all agents have perfect reasoning ability, the generative model may consider individual differences in the level of reasoning sophistication, possibly as described by the famous cognitive hierarchy theory (Camerer et al., 2004). Future research would be invaluable in exploring whether and how cognitive biases would influence strategic behavior through internal generative processes, and to what extent biases in social inferences would be similar to or distinct from biases in one's own cognitive and behavioral computations.

One possible mechanism by which indirect experiences support such generative processes is through affecting the internal knowledge related to how we see each other. For example, in the communicative context, the interpretation of a communicative signal might be different when the signal sender is a 3-year-old versus 30-year-old, or is a native versus non-native language speaker. Indeed, recent studies provide evidence that the perception of social partners and their relationships is organized in ways similar to map-like representations (Park et al., 2020; Park et al., 2021; Tavares et al., 2015). It is possible that, by providing structural frameworks for social perception, cognitive maps may facilitate the evaluation of and reasoning about the interacting partner in service of strategic choice selection. One particularly interesting case is related to stereotypes, which may provide a mechanism by which preexisting knowledge helps map the information related to a stranger's age, gender, ethnic group into task-relevant characteristics like trustworthiness, warmth, or competence (Fiske et al., 2007; Greenwald & Banaji, 1995), thereby affecting interpersonal choices. On the one hand, extant data have shown that stereotype-related information may be organized in ways mimicking cognitive maps in the hippocampus, entorhinal cortex, medial prefrontal cortex, and other brain regions (Park et al., 2020; Park et al., 2021). On the other hand, quantitative variations in stereotypes about others' warmth and competence are observed to modulate resource allocation behavior toward these individuals (Jenkins et al., 2018). It remains to connect these disparate lines of research, and explore whether and how strategic decisions are supported by information retrieved from map-like spaces that organize the social perception in terms of its positions in the space. Other important questions for future research include how humans and other social animals gradually acquire the generative models of others' behavior, and how such ability develops over the life span. What are the cognitive and social functions necessary for building generative model based on indirect knowledge for strategic behavior, and what neural developments support or facilitate the implementation of such generative processes?

4 | CONCLUSION

Strategic behavior is a topic of tremendous importance and at the intersection of social and biological sciences. Recent developments in decision neuroscience have highlighted the possibility that strategic behavior may be guided by not only basic reward processing, but also other-predictive signals generated from an internal model about others. This model may be either learned from repeated interpersonal interactions through sophisticated RL-like processes, or generalized from indirect social knowledge via processes related to cognitive maps or others that support novel decisions and inferences. This perspective thus ties together seemingly diverse strategic behaviors and implies that the key difference between adaptive and novel strategic decisions may involve the differential neurocomputational processes by which mental models are constructed. Future work is needed to rigorously test these possibilities and explore whether interpersonal abnormalities observed in psychiatric conditions such as autism (Chiu et al., 2008; Yoshida, Dziobek, et al., 2010), paranoia (Raihani & Bell, 2017), social anxiety (Sripada et al., 2009), or borderline personality disorder (King-Casas et al., 2008) are related to the impaired ability to learn or to generalize from past social experience.

ACKNOWLEDGMENTS

The authors thank the helpful comments and suggestions from three anonymous referees. They are also grateful to Cong Wang, Guihua Yu, and other members of the Zhu laboratory.

CONFLICT OF INTEREST

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

AUTHOR CONTRIBUTIONS

Yaomin Jiang: Conceptualization (supporting); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Hai-Tao Wu:** Conceptualization (supporting); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Qingtian Mi:** Conceptualization (supporting); visualization (equal); writing – original draft (equal); writing – review and editing (supporting). **Lusha Zhu:** Conceptualization (lead); funding acquisition (lead); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

DATA AVAILABILITY STATEMENT

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

ORCID

Yaomin Jiang  <https://orcid.org/0000-0002-2367-2374>

Hai-Tao Wu  <https://orcid.org/0000-0003-3458-1459>

Qingtian Mi  <https://orcid.org/0000-0002-5175-4780>

Lusha Zhu  <https://orcid.org/0000-0001-8717-6356>

RELATED WIREs ARTICLES

[Enhancing models of social and strategic decision making with process tracing and neural data](#)

[Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data](#)

[Decision neuroscience: neuroeconomics](#)

REFERENCES

- Abe, H., & Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron*, 70(4), 731–741. <https://doi.org/10.1016/j.neuron.2011.03.026>
- Apps, M. A., Green, R., & Ramnani, N. (2013). Reinforcement learning signals in the anterior cingulate cortex code for others' false beliefs. *NeuroImage*, 64, 1–9. <https://doi.org/10.1016/j.neuroimage.2012.09.010>
- Apps, M. A. J., Rushworth, M. F. S., & Chang, S. W. C. (2016). The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron*, 90(4), 692–707. <https://doi.org/10.1016/j.neuron.2016.04.018>
- Backus, A. R., Schoffelen, J.-M., Szebényi, S., Hanslmayr, S., & Doeller, C. F. (2016). Hippocampal-prefrontal theta oscillations support memory integration. *Current Biology*, 26(4), 450–457. <https://doi.org/10.1016/j.cub.2015.12.048>
- Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1, 1–10. <https://doi.org/10.1038/s41562-017-0064>
- Balsters, J. H., Apps, M. A., Bolis, D., Lehner, R., Gallagher, L., & Wenderoth, N. (2017). Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain*, 140(1), 235–246. <https://doi.org/10.1093/brain/aww287>
- Baram, A. B., Muller, T. H., Nili, H., Garvert, M. M., & Behrens, T. E. J. (2021). Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. *Neuron*, 109(4), 713–723.e7. <https://doi.org/10.1016/j.neuron.2020.11.024>
- Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45), 18327–18332. <https://doi.org/10.1073/pnas.1306572110>
- Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science*, 324(5931), 1160–1164. <https://doi.org/10.1126/science.1169694>
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, 456(7219), 245–249. <https://doi.org/10.1038/nature07538>
- Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, 100(2), 490–509. <https://doi.org/10.1016/j.neuron.2018.10.002>
- Bellman, R. (1957). *Dynamic programming*. Princeton University Press.
- Bhatt, M. A., Lohrenz, T., Camerer, C. F., & Montague, P. R. (2012). Distinct contributions of the amygdala and parahippocampal gyrus to suspicion in a repeated bargaining game. *Proceedings of the National Academy of Sciences of the United States of America*, 109(22), 8728–8733. doi:10.1073/pnas.1200738109
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1), B25–B31. [https://doi.org/10.1016/S0010-0277\(00\)00096-2](https://doi.org/10.1016/S0010-0277(00)00096-2)
- Bongioanni, A., Folloni, D., Verhagen, L., Sallet, J., Klein-Flügge, M. C., & Rushworth, M. F. S. (2021). Activation and disruption of a neural mechanism for novel choice in monkeys. *Nature*, 591(7849), 270–274. <https://doi.org/10.1038/s41586-020-03115-5>

- Boorman, E. D., Sweigart, S. C., & Park, S. A. (2021). Cognitive maps and novel inferences: A flexibility hierarchy. *Current Opinion in Behavioral Sciences*, 38, 141–149. <https://doi.org/10.1016/j.cobeha.2021.02.017>
- Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends in Cognitive Sciences*, 23, 408–422. <https://doi.org/10.1016/j.tics.2019.02.006>
- Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences. United States of America*, 107(32), 14431–14436. <https://doi.org/10.1073/pnas.1003111107>
- Camerer, C., & Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4), 827–874. <https://doi.org/10.1111/1468-0262.00054>
- Camerer, C. F. (2011). *Behavioral game theory*. Princeton University Press.
- Camerer, C. F., & Hare, T. A. (2014). The neural basis of strategic choice. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics* (pp. 479–492). Academic Press.
- Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3), 861–898.
- Campbell-Meiklejohn, D., Simonsen, A., Frith, C. D., & Daw, N. D. (2017). Independent neural computation of value from other people's confidence. *Journal of Neuroscience*, 37(3), 673–684. <https://doi.org/10.1523/JNEUROSCI.4490-15.2017>
- Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, 70(3), 560–572. <https://doi.org/10.1016/j.neuron.2011.02.056>
- Charness, G., Rustichini, A., & Ven, J. (2017). Self-confidence and strategic behavior. *Experimental Economics*, 21(1), 73–99. <https://doi.org/10.1007/s10683-017-9526-3>
- Charpentier, C. J., Iigaya, K., & O'Doherty, J. P. (2020). A neuro-computational account of arbitration between choice imitation and goal emulation during human observational learning. *Neuron*, 106(4), 687–699.e7. <https://doi.org/10.1016/j.neuron.2020.02.028>
- Chiu, P. H., Kayali, M. A., Kishida, K. T., Tomlin, D., Klinger, L. G., Klinger, M. R., & Montague, P. R. (2008). Self responses along cingulate cortex reveal quantitative neural phenotype for high-functioning autism. *Neuron*, 57(3), 463–473. <https://doi.org/10.1016/j.neuron.2007.12.020>
- Clark, H. H. (1996). *Using language*. Cambridge University Press.
- Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature*, 462(7), 51–57. <https://doi.org/10.1038/nature08366>
- Darwin, C. (2004). *On the origin of species, 1859*. Routledge.
- Daw, N. D. (2018). Are we of two minds? *Nature Neuroscience*, 21(11), 1497–1499. <https://doi.org/10.1038/s41593-018-0258-2>
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704–1711. <https://doi.org/10.1038/nn1560>
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), 876–879. <https://doi.org/10.1038/nature04766>
- Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience*. MIT Press.
- Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, 7(5), 889–904.
- De Freitas, J., Thomas, K., DeScioli, P., & Pinker, S. (2019). Common knowledge, coordination, and strategic mentalizing in human social life. *Proceedings of the National Academy of Sciences of the United States of America*, 1, 13751–13758. <https://doi.org/10.1073/pnas.1905518116>
- De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., & Love, B. C. (2017). Social information is integrated into value and confidence judgments according to its reliability. *Journal of Neuroscience*, 37, 6066–6074. <https://doi.org/10.1523/JNEUROSCI.3880-16.2017>
- de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305(5688), 1254–1258. <https://doi.org/10.1126/science.1100735>
- Decety, J., & Yoder, K. J. (2017). The emerging social neuroscience of justice motivation. *Trends in Cognitive Sciences*, 21(1), 6–14. <https://doi.org/10.1016/j.tics.2016.10.008>
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5), 178–190. [https://doi.org/10.1002/\(SICI\)1520-6505\(1998\)6:5<178::AID-EVAN5>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8)
- Dunne, S., & O'Doherty, J. P. (2013). Insights from the application of computational neuroimaging to social neuroscience. *Current Opinion in Neurobiology*, 23(3), 1–6. <https://doi.org/10.1016/j.conb.2013.02.007>
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets*. Cambridge University Press.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4), 848–881.
- Farquharson, R. (1969). *Theory of voting*. Blackwell.
- Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Sciences*, 11(10), 419–427. <https://doi.org/10.1016/j.tics.2007.09.002>
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791. <https://doi.org/10.1038/nature02043>
- Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, 8(4), 185–190. <https://doi.org/10.1016/j.tics.2004.02.007>
- Fehr, E., & Krajbich, I. (2014). Social preferences and the brain. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics* (2nd ed., pp. 193–218). Academic Press. <https://doi.org/10.1016/B978-0-12-416008-8.00011-5>
- FeldmanHall, O., & Nassar, M. R. (2021). The computational challenge of social learning. *Trends in Cognitive Sciences*, 1–13, 1045–1057. <https://doi.org/10.1016/j.tics.2021.09.002>

- Ferreira, P. L., Santos, F. C., & Pequito, S. (2021). Risk sensitivity and theory of mind in human coordination. *PLoS Computational Biology*, 17(7), e1009167. <https://doi.org/10.1371/journal.pcbi.1009167>
- Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences of the United States of America*, 113(34), E5072–E5081. <https://doi.org/10.1073/pnas.1610344113>
- Fisher, M., & Oppenheimer, D. M. (2021). Harder than you think: How outside assistance leads to overconfidence. *Psychological Science*, 32(4), 598–610. <https://doi.org/10.1177/0956797620975779>
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084), 998. <https://doi.org/10.1126/science.1218633>
- Franke, M., & Degen, J. (2016). Reasoning in reference games: Individual- vs. population-level probabilistic modeling. *PLoS One*, 11(5), e0154854–e0154825. <https://doi.org/10.1371/journal.pone.0154854>
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation*, 29(1), 1–49. https://doi.org/10.1162/NECO_a_00912
- Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games*. MIT Press.
- Gershman, S. J., & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science*, 7(3), 391–415. <https://doi.org/10.1111/tops.12138>
- Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585–595. <https://doi.org/10.1016/j.neuron.2010.04.016>
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings National Academy of Sciences United States of America*, 108(Suppl. 3), 15647–15654. <https://doi.org/10.1073/pnas.1014269108>
- Glimcher, P. W., & Fehr, E. (2014). *Neuroeconomics: Decision making and the brain*. Academic Press.
- Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11), 818–829. <https://doi.org/10.1016/j.tics.2016.08.005>
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). Academic Press.
- Griffiths, T. L., Callaway, F., Chang, M. B., Grant, E., Krueger, P. M., & Lieder, F. (2019). Doing more with less: Meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, 29, 24–30. <https://doi.org/10.1016/j.cobeha.2019.01.005>
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(18), 6741–6746. <https://doi.org/10.1073/pnas.0711099105>
- Haroush, K., & Williams, Z. M. (2015). Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell*, 160(6), 1233–1245. <https://doi.org/10.1016/j.cell.2015.01.045>
- Hawkins, R. X. D., Goodman, N. D., & Goldstone, R. L. (2019). The emergence of social norms and conventions. *Trends in Cognitive Sciences*, 23(2), 158–169. <https://doi.org/10.1016/j.tics.2018.11.003>
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review*, 91(2), 73–78.
- Hertz, U., Palminteri, S., Brunetti, S., Olesen, C., Frith, C. D., & Bahrami, B. (2017). Neural computations underpinning the strategic management of influence in advice giving. *Nature Communications*, 8(1), 1–12. <https://doi.org/10.1038/s41467-017-02314-5>
- Hill, C. A., Suzuki, S., Polania, R., & Ruff, C. C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature Neuroscience*, 20(8), 1142–1149. <https://doi.org/10.1038/nn.4602>
- Hinton, G. E. (2007). Learning multiple layers of representation. *Trends in Cognitive Sciences*, 11(10), 428–434. <https://doi.org/10.1016/j.tics.2007.09.004>
- Hunter, L. E., Meer, E. A., Gillan, C. M., Hsu, M., & Daw, N. D. (2021). Increased and biased deliberation in social anxiety. *Nature Human Behaviour*, 1–12, 146–154. doi:10.1038/s41562-021-01180-y
- Jamali, M., Grannan, B. L., Fedorenko, E., Saxe, R., Báez-Mendoza, R., & Williams, Z. M. (2021). Single-neuronal predictions of others' beliefs in humans. *Nature*, 591(7851), 610–614. <https://doi.org/10.1038/s41586-021-03184-0>
- Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, 20(8), 589–604. <https://doi.org/10.1016/j.tics.2016.05.011>
- Jenkins, A. C., Karashchuk, P., Zhu, L., & Hsu, M. (2018). Predicting human behavior toward members of different social groups. *Proceedings National Academy of Sciences United States of America*, 115, 9696–9701. <https://doi.org/10.1073/pnas.1719452115>
- Kikumoto, A., & Mayr, U. (2019). Balancing model-based and memory-free action selection under competitive pressure. *eLife*, 8, 8. <https://doi.org/10.7554/eLife.48810>
- King-Casas, B., & Chiu, P. H. (2012). Understanding interpersonal function in psychiatric illness through multiplayer economic games. *Biological Psychiatry*, 72(2), 119–125. <https://doi.org/10.1016/j.biopsych.2012.03.033>
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science*, 321(5890), 806–810. <https://doi.org/10.1126/science.1156902>
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, 308(5718), 78–83. <https://doi.org/10.1126/science.1108062>

- Kishida, K. T., King-Casas, B., & Montague, P. R. (2010). Neuroeconomic approaches to mental disorders. *Neuron*, 67(4), 543–554. <https://doi.org/10.1016/j.neuron.2010.07.021>
- Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.
- Konovalov, A., Hill, C., Daunizeau, J., & Ruff, C. C. (2021). Dissecting functional contributions of the social brain to strategic behavior. *Neuron*, 109(20), 3323–3337. doi:10.1016/j.neuron.2021.07.025
- Konovalov, A., & Ruff, C. C. (2022). Enhancing models of social and strategic decision making with process tracing and neural data. *Wiley Interdisciplinary Reviews. Cognitive Science*, 13(1), e1559. <https://doi.org/10.1002/wcs.1559>
- Kool, W., Cushman, F. A., & Gershman, S. J. (2018). Competition and cooperation between multiple reinforcement learning systems. In R. Morris, A. Bornstein, & A. Shenhav (Eds.), *Goal-directed decision making* (pp. 153–178). Academic Press. <https://doi.org/10.1016/B978-0-12-812098-9.00007-3>
- Kording, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10(7), 319–326. <https://doi.org/10.1016/j.tics.2006.05.003>
- Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, 79(5), 836–848. <https://doi.org/10.1016/j.neuron.2013.08.020>
- Krajbich, I., Adolphs, R., Tranel, D., Denburg, N. L., & Camerer, C. F. (2009). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *Journal of Neuroscience*, 29(7), 2188–2192. <https://doi.org/10.1523/JNEUROSCI.5086-08.2009>
- Lee, D., & Seo, H. (2015). Neural basis of strategic decision making. *Trends in Neurosciences*, 39(1), 1–9. <https://doi.org/10.1016/j.tins.2015.11.002>
- Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43, e1. <https://doi.org/10.1017/S0140525X1900061X>
- Ligneul, R., Obeso, I., Ruff, C. C., & Dreher, J.-C. (2016). Dynamical representation of dominance relationships in the human rostromedial prefrontal cortex. *Current Biology*, 26(23), 1–9. <https://doi.org/10.1016/j.cub.2016.09.015>
- Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is there a “social” brain? Implementations and algorithms. *Trends in Cognitive Sciences*, 24(10), 802–813. <https://doi.org/10.1016/j.tics.2020.06.011>
- Lockwood, P. L., & Klein-Flügge, M. C. (2021). Computational modelling of social cognition and behaviour—A reinforcement learning primer. *Social Cognitive and Affective Neuroscience*, 16(8), 761–771. <https://doi.org/10.1093/scan/nsaa040>
- Mark, S., Moran, R., Parr, T., Kennerley, S. W., & Behrens, T. E. J. (2020). Transferring structural knowledge across cognitive maps in humans and models. *Nature Communications*, 11(1), 1–12. <https://doi.org/10.1038/s41467-020-18254-6>
- Mars, R. B., Sallet, J., Neubert, F.-X., & Rushworth, M. F. S. (2013). Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 110(26), 10806–10811. <https://doi.org/10.1073/pnas.1302956110>
- Mi, Q., Wang, C., Camerer, C. F., & Zhu, L. (2021). Reading between the lines: Listener's vmPFC simulates speaker cooperative choices in communication games. *Science Advances*, 7(10), eabe6276. <https://doi.org/10.1126/sciadv.abe6276>
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, 29, 417–448. <https://doi.org/10.1146/annurev.neuro.29.051605.112903>
- Montague, P. R., & Lohrenz, T. (2007). To detect and correct: Norm violations and their enforcement. *Neuron*, 56(1), 14–18. <https://doi.org/10.1016/j.neuron.2007.09.020>
- Montague, P. R., Lohrenz, T., & Dayan, P. (2015). The three R's of trust. *Current Opinion in Behavioral Sciences*, 3, 102–106. <https://doi.org/10.1016/j.cobeha.2015.02.009>
- Moulin, H. (1986). *Game theory for the social sciences*. NYU press.
- Nihonsugi, T., Ihara, A., & Haruno, M. (2015). Selective increase of intention-based economic decisions by noninvasive brain stimulation to the dorsolateral prefrontal cortex. *Journal of Neuroscience*, 35(8), 3412–3419. <https://doi.org/10.1523/JNEUROSCI.3885-14.2015>
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53, 139–154. <https://doi.org/10.1016/j.jmp.2008.12.005>
- Niv, Y. (2019). Learning task-state representations. *Nature Neuroscience*, 22(10), 1–10. <https://doi.org/10.1038/s41593-019-0470-8>
- Noveck, I. A., & Reboul, A. (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences*, 12(11), 425–431. <https://doi.org/10.1016/j.tics.2008.07.009>
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563. <https://doi.org/10.1126/science.1133755>
- Nyarko, Y., & Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica*, 70(3), 971–1005. <https://doi.org/10.1111/1468-0262.00316>
- O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, 38(2), 329–337. [https://doi.org/10.1016/S0896-6273\(03\)00169-7](https://doi.org/10.1016/S0896-6273(03)00169-7)
- O'Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304(5669), 452–454. <https://doi.org/10.1126/science.1094285>
- Ong, W. S., Madlon-Kay, S., & Platt, M. L. (2020). Neuronal correlates of strategic cooperation in monkeys. *Nature Neuroscience*, 555, 1–22. <https://doi.org/10.1038/s41593-020-00746-9>
- Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. MIT Press.
- Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science*, 24(5), 751–761. <https://doi.org/10.1177/0956797612463080>
- Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, 5(2), 97–98. <https://doi.org/10.1038/nn802>

- Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature Neuroscience*, 24(9), 1292–1301. <https://doi.org/10.1038/s41593-021-00916-3>
- Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020). Map making: Constructing, combining, and inferring on abstract cognitive maps. *Neuron*, 107(6), 1226–1238. <https://doi.org/10.1016/j.neuron.2020.06.030>
- Park, S. A., Sestito, M., Boorman, E. D., & Dreher, J.-C. (2019). Neural computations underlying strategic social decision-making in groups. *Nature Communications*, 1–12, 5287. <https://doi.org/10.1038/s41467-019-12937-5>
- Parke, D. C., & Wellman, M. P. (2015). Economic reasoning and artificial intelligence. *Science*, 349(6), 267–272. <https://doi.org/10.1126/science.aaa8403>
- Pescetelli, N., & Yeung, N. (2021). The role of decision confidence in advice-taking and trust formation. *Journal of Experimental Psychology: General*, 150(3), 507–526. <https://doi.org/10.1037/xge0000960>
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042–1045. <https://doi.org/10.1038/nature05051>
- Radulescu, A., Shin, Y. S., & Niv, Y. (2021). Human representation learning. *Annual Review of Neuroscience*, 44, 253–273. <https://doi.org/10.1146/annurev-neuro-092920-120559>
- Raihani, N. J., & Bell, V. (2017). Paranoia and the social representation of others: A large-scale game theory approach. *Scientific Reports*, 7(1), 4544. <https://doi.org/10.1038/s41598-017-04805-3>
- Rao, R. (1999). An optimal estimation approach to visual perception and learning. *Vision Research*, 39(11), 1963–1989. [https://doi.org/10.1016/s0042-6989\(98\)00279-x](https://doi.org/10.1016/s0042-6989(98)00279-x)
- Ray, D., King-Casas, B., Montague, P. R., & Dayan, P. (2008). Bayesian model of behaviour in economic games. *Advances in Neural Information Processing Systems*, 21, 1345–1352.
- Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, 62, 23–48. <https://doi.org/10.1146/annurev.psych.121208.131647>
- Robson, S. E., Repetto, L., Gountouna, V.-E., & Nicodemus, K. K. (2020). A review of neuroeconomic gameplay in psychiatric disorders. *Molecular Psychiatry*, 25(1), 67–81. <https://doi.org/10.1038/s41380-019-0405-5>
- Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, 15(8), 549–562. <https://doi.org/10.1038/nrn3776>
- Ruff, C. C., Ugazio, G., & Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, 342(6157), 482–484. <https://doi.org/10.1126/science.1241399>
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, 55, 87–124. <https://doi.org/10.1146/annurev.psych.55.090902.142044>
- Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, 19(2), 65–72. <https://doi.org/10.1016/j.tics.2014.11.007>
- Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8. <https://doi.org/10.1016/j.cobeha.2014.07.005>
- Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., Lillicrap, T., & Silver, D. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature*, 588(7839), 604–609. <https://doi.org/10.1038/s41586-020-03051-4>
- Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6), 1402–1412. <https://doi.org/10.1016/j.neuron.2016.08.019>
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>
- Seo, H., Cai, X., Donahue, C. H., & Lee, D. (2014). Neural correlates of strategic reasoning during competitive games. *Science*, 346(6207), 340–343. <https://doi.org/10.1126/science.1256254>
- Sliwa, J., & Freiwald, W. A. (2017). A dedicated network for social interaction processing in the primate brain. *Science*, 356(6339), 745–749. <https://doi.org/10.1126/science.aam6383>
- Soldà, A., Ke, C., von Hippel, W., & Page, L. (2021). Absolute versus relative success: Why overconfidence creates an inefficient equilibrium. *Psychological Science*, 32(10), 1662–1674. <https://doi.org/10.1177/09567976211007414>
- Son, J. Y., Bhandari, A., & FeldmanHall, O. (2021). Cognitive maps of social features enable flexible inference in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 118(39), e2021699118. <https://doi.org/10.1073/pnas.2021699118>
- Spalding, K. N., Schlichting, M. L., Zeithamova, D., Preston, A. R., Tranel, D., Duff, M. C., & Warren, D. E. (2018). Ventromedial prefrontal cortex is necessary for normal associative inference and memory integration. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 38(15), 3767–3775. <https://doi.org/10.1523/JNEUROSCI.2501-17.2018>
- Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007). The neural signature of social norm compliance. *Neuron*, 56(1), 185–196. <https://doi.org/10.1016/j.neuron.2007.09.011>
- Sripada, C. S., Angstadt, M., Banks, S., Nathan, P. J., Liberzon, I., & Phan, K. L. (2009). Functional neuroimaging of mentalizing during the trust game in social anxiety disorder. *Neuroreport*, 20(11), 984–989. <https://doi.org/10.1097/WNR.0b013e32832d0a67>
- Stewart, A. J., Mosleh, M., Diakonova, M., Arechar, A. A., Rand, D. G., & Plotkin, J. B. (2019). Information gerrymandering and undemocratic decisions. *Nature*, 573(7772), 1–9. <https://doi.org/10.1038/s41586-019-1507-6>
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning*. MIT Press.

- Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., Cheng, K., & Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron*, 74(6), 1125–1137. <https://doi.org/10.1016/j.neuron.2012.04.030>
- Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A map for social navigation in the human brain. *Neuron*, 87(1), 231–243. <https://doi.org/10.1016/j.neuron.2015.06.011>
- Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, 10(7), 309–318. <https://doi.org/10.1016/j.tics.2006.05.009>
- Tervo, D. G. R., Proskurin, M., Manakov, M., Kabra, M., Vollmer, A., Branson, K., & Karpova, A. Y. (2014). Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell*, 159(1), 21–32. <https://doi.org/10.1016/j.cell.2014.08.037>
- Tomasello, M. (2010). *Origins of human communication*. MIT Press.
- Tomasello, M. (2019). *Becoming human*. Belknap Press.
- Tusche, A., & Bas, L. M. (2021). Neurocomputational models of altruistic decision-making and social motives: Advances, pitfalls, and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 12(6), e1571. <https://doi.org/10.1002/wcs.1571>
- van den Bos, W., Talwar, A., & McClure, S. M. (2013). Neural correlates of reinforcement learning and social preferences in competitive bidding. *Journal of Neuroscience*, 33(5), 2137–2146. <https://doi.org/10.1523/JNEUROSCI.3095-12.2013>
- Vilalta, R., & Drissi, Y. (2002). A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2), 77–95. <https://doi.org/10.1023/A:1019956318069>
- Weibull, J. W. (1997). *Evolutionary game theory*. MIT Press.
- Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Neuroscience*, 17, 513–523. <https://doi.org/10.1038/nrn.2016.56>
- Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81(2), 267–279. <https://doi.org/10.1016/j.neuron.2013.11.005>
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128.
- Wittmann, M. K., Lockwood, P. L., & Rushworth, M. F. (2018). Neural mechanisms of social cognition in primates. *Annual Review of Neuroscience*, 41, 99–118. <https://doi.org/10.1146/annurev-neuro-080317-061450>
- Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Computational Biology*, 4(12), e1000254–e1000214. <https://doi.org/10.1371/journal.pcbi.1000254>
- Yoshida, W., Dziobek, I., Kliemann, D., Heekeren, H. R., Friston, K. J., & Dolan, R. J. (2010). Cooperation and heterogeneity of the autistic mind. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(26), 8815–8818. <https://doi.org/10.1523/JNEUROSCI.0400-10.2010>
- Yoshida, W., Seymour, B., Friston, K. J., & Dolan, R. J. (2010). Neural mechanisms of belief inference during cooperative games. *Journal of Neuroscience*, 30(32), 10744–10751. <https://doi.org/10.1523/JNEUROSCI.5895-09.2010>
- Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, 75(1), 168–179. <https://doi.org/10.1016/j.neuron.2012.05.010>
- Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience*, 15(6), 695–707. <https://doi.org/10.1093/scan/nsaa089>
- Zhu, L., Jiang, Y., Scabini, D., Knight, R. T., & Hsu, M. (2019). Patients with basal ganglia damage show preserved learning in an economic game. *Nature Communications*, 10(1), 1–10. <https://doi.org/10.1038/s41467-019-08766-1>
- Zhu, L., Mathewson, K. E., & Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proceedings National Academy of Sciences United States of America*, 109(5), 1419–1424. <https://doi.org/10.1073/pnas.1116783109>
- Zoh, Y., Chang, S. W., & Crockett, M. J. (2022). The prefrontal cortex and (uniquely) human cooperation: A comparative perspective. *Neuropsychopharmacology*, 47(1), 119–133. <https://doi.org/10.1038/s41386-021-01092-5>

How to cite this article: Jiang, Y., Wu, H.-T., Mi, Q., & Zhu, L. (2022). Neurocomputations of strategic behavior: From iterated to novel interactions. *WIREs Cognitive Science*, 13(4), e1598. <https://doi.org/10.1002/wcs.1598>