

Review

Not peer-reviewed version

Decision-Making in Repeated Games: Insights from Active Inference

[Hui Yuan](#) , [Ligang Wang](#) ^{*} , Wenbin Gao , Ting Tao , Chunlei Fan

Posted Date: 28 October 2025

doi: 10.20944/preprints202510.2029.v1

Keywords: repeated games; decision-making; computational modeling; active inference



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Decision-Making in Repeated Games: Insights from Active Inference

Hui Yuan ^{1,2}, Ligang Wang ^{1,2,*}, Wenbin Gao ^{1,2}, Ting Tao ^{1,2} and Chunlei Fan ^{1,2}

¹ State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

² Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: wanglg@psych.ac.cn

Abstract

This review systematically explores the potential of the active inference framework in illuminating the cognitive mechanisms of decision-making within repeated games. Characterized by multi-round interactions and social uncertainty, repeated games more closely resemble real-world social scenarios, where the decision-making process involves interconnected cognitive components such as inference, policy selection, and learning. Unlike traditional reinforcement learning models, active inference, grounded in the free energy minimization principle, unifies perception, learning, planning, and action within a single generative model. Belief updating is achieved by minimizing variational free energy, while the exploration-exploitation dilemma is balanced by minimizing expected free energy. Formulated based on partially observable Markov decision processes, the framework naturally incorporates social uncertainty, and its hierarchical structure allows for simulating mentalizing processes, thereby offering a unified account of social decision-making. Future research can further validate its effectiveness through model simulation and behavioral fitting.

Keywords: repeated games; decision-making; computational modeling; active inference

1. Introduction

Decision-making in repeated games represents a crucial aspect of human intelligence and has been the focus of extensive research across fields such as psychology and artificial intelligence (Akata et al., 2025; van Dijk & De Dreu, 2021). Participants in repeated games face multiple rounds of interpersonal interactions that more closely resemble the real social world - a world often portrayed as more unpredictable and uncertain than the non-social world (FeldmanHall & Shenhav, 2019). Decision-making within social contexts presents significant complexity and difficulty. This stems not only from the difficulty of identifying choices that maximize self-interest given the unpredictability of others' behaviors, but also from the necessity to reconcile the self-interest with the interests of others, which requires a trade-off between cooperation and competition (Lee, 2008). Decision-making in such contexts presents significant challenges for both artificial intelligence and human agents. Consequently, a deeper investigation into its cognitive mechanisms is crucial not only for enhancing human decision-making capabilities but also for advancing AI towards greater intelligence, enhanced adaptability to complex social environments, and more effective collaboration with humans.

Traditional psychological methods are limited in their ability to assess such complex dynamic processes. Computational modeling, as a quantitative methodology characterized by rigor, scientific precision, and interpretability, has been extensively applied in domains such as computational psychiatry (Hitchcock et al., 2022; Montague et al., 2012), and demonstrates emerging utility in social psychology (Cushman, 2024; Hackel & Amodio, 2018). It enables the simulation of complex cognitive processes underlying behavioral phenomena through mathematical formalisms, provides rigorous scientific characterization of the dynamic mechanisms governing human behavior, and derives latent variables from behavioral data that are not directly observable (Montague, 2018). Thereby, it

establishes a novel theoretical framework for advancing understanding of psychological and neural mechanisms, overcoming the limitations of traditional research methods.

Various families of computational models have been developed to characterize the decision-making process. Unlike reinforcement learning, which frames decision-making as an adaptive behavior driven by the principle of maximizing external rewards (Bartoc, 1998), active inference conceptualizes decisions as a process of minimizing free energy (Friston et al., 2016). Here, the decision-making agent explores the environment to reduce uncertainty, minimizing the discrepancy between anticipated outcomes and preferred outcomes, while dynamically integrating perception and policy selection. The active inference framework offers a promising model for simulating human cognitive processes in partially observable environments, such as social contexts. This approach could significantly deepen our understanding of how human intelligence navigates adaptive decision-making within dynamic social settings.

2. Game Theory and Repeated Games

Within social environments, human decision-making necessitates explicit consideration of others' behaviors. Games provide a valuable methodological framework for quantitatively assessing such social decisions, with structured interaction scenarios allowing for the study of social cognitive processes such as cooperation, competition, and trust formation (Maurer et al., 2018; Ng & Au, 2016; Zhang et al., 2017). Unlike single-round games in which participants have no shared past or future, repeated games more closely resemble real-life social interactions. These interactions involve multiple sequential encounters where participants receive feedback between rounds. The feedback enables participant update their beliefs of the opponent and adjust their strategy in time, resulting in dynamically evolving behavior over time (Cochard et al., 2004).

The payoff matrix of the game describes the strategies available to players and their corresponding outcomes. Based on this matrix, strategies and their expected utilities can be quantitatively analyzed. Assuming that all players act rationally to maximize their individual payoffs, the optimal strategies for all participants called Nash equilibria can be derived (Nash, 1950). Depending on the nature of interdependence among players' interests, repeated games can be categorized into three distinct types: conflict games, cooperation games, and coordination games (Schelling, 1960).

Conflict games, also known as zero-sum games, are strategic situations in which the interests of players are inherently antagonistic. In such games, the total sum of payoffs and losses across all players always equals zero, meaning one player's gain necessarily implies another's loss (Schelling, 1960). Under this framework, cooperation among players is structurally impossible; instead, decision-making is premised on competition. Participants focus on concealing their own strategic intentions while attempting to infer opponents' strategies and psychological states to optimize their own outcomes. Conflict games manifest ubiquitously across real-world scenarios, encompassing competitive sports (e.g., table tennis), board games (e.g., chess), as well as the game of rock-paper-scissors. Within psychological research, such adversarial frameworks are frequently employed as experimental paradigms to investigate deception, inequality, and related strategic performance (Lacomba et al., 2017; Zhang et al., 2017).

Cooperation games, also known as mixed-motive games, characterize interactions where individuals face both competitive and collaborative incentives, resulting in payoff structures that are neither purely adversarial nor fully aligned. These games inherently involve resource allocation between self and others, compelling players to navigate a critical trade-off: whether to cooperate to maximize collective benefits or defect to secure personal gains (Miller Moya, 2007). Unlike zero-sum conflict games, participants in cooperative frameworks exhibit stronger motivations to understand opponents' strategies (Wang & Kwan, 2023). Cooperation games such as the Prisoner's dilemma, trust game, and dictator game are widely employed to investigate the formation of cooperation and defection (Bonowski & Minnameier, 2022; Loennqvist & Walkowitz, 2019; Press & Dyson, 2012). The Prisoner's dilemma stands as a canonical example, wherein two players independently choose

between "cooperation" and "defection" without communication. Its payoff structure is characterized by four critical outcomes: mutual cooperation yields moderate rewards for both players; mutual defection results in moderate penalties; while unilateral defection provides the defector with maximum gains at the expense of the cooperator, who incurs the highest penalty. This matrix is shown in **Figure 1**. In the one-shot Prisoner's dilemma, defection constitutes the dominant strategy that maximizes individual payoff under rationality. However, in the repeated game, persistent defection risks triggering retaliatory defection from opponents in subsequent rounds, reducing long-term cumulative gains. It is within this interaction mechanism that cooperation gradually forms (Colman et al., 2018).

| | cooperation | defection |
|-------------|-------------|-----------|
| cooperation | 3, 3 | 0, 5 |
| defection | 5, 0 | 1, 1 |

Figure 1. Decision and Payoff Matrix of the Prisoner's Dilemma Game.

The coordination game is a type of game where the interests of participants are consistently aligned. In such a game, the payoffs and losses of the participants are entirely identical, meaning that both parties will inevitably either succeed together or fail together (Schelling, 1960). Under the rules of this game, due to the congruence of interests between the two sides, cooperation becomes their inevitable choice. The players only need to consider how to coordinate with each other, rather than deliberating on how to address the issue of interest distribution between themselves and others. Coordination games, such as matching games (e.g., the Heads-Tails game), involve scenarios where players receive a positive payoff only if they select the same option; whereas mismatched choices result in no payoff for either. While less frequently studied in decision-making processes research, these games are commonly used to investigate inter-brain synchronization during cooperation (Cui et al., 2012; Pan et al., 2017).

On the continuum of games, pure coordination games and zero-sum games represent the extremes, while mixed-motive games, incorporating elements of both cooperation and conflict, occupy the intermediate space (Schelling, 1960). However, within the context of repeated games, regardless of the players' incentive structure, the opponent's behavior directly impacts the decision-maker's choices and outcomes. The opponent's hidden states and actions introduce social uncertainty that exist over and above the non-social uncertainty. Social uncertainty refers to a person's inability to precisely predict their own future states and actions due to uncertainty about others' states and actions (FeldmanHall & Shenhav, 2019). The repeated game continuum provides a powerful framework for investigating decision-making processes under such social uncertainty.

3. Decision-Making in Repeated Games

Extensive research has focused on decision-making processes under non-social uncertainty. Individuals prioritize their attention on cues that can reduce uncertainty (Walker et al., 2019) and pay greater attention to options with higher uncertainty (Stojic, Orquin, et al., 2020). Increased uncertainty enhances individuals' learning rates (Speekenbrink & Shanks, 2010) and promotes their exploratory behaviors (Stojic, Schulz, et al., 2020). While the mechanisms by which non-social uncertainty influences decision-making can be extended to social uncertainty to some extent, social contexts inherently exhibit greater unpredictability. This is because others' behavioral motives are difficult to directly observe and evolve dynamically over time. Concurrently, in contexts involving repeated interactions with others—such as repeated games—suboptimal decisions may exert long-lasting negative effects. These effects can hinder the formation of cooperation with others and ultimately compromise one's own interests.

To resolve social uncertainty in order to make adaptive decisions, individuals develop specialized mechanisms in which social inference plays an important role. Inference can be categorized into automatic inference and controlled inference based on the degree of cognitive control and the corresponding effort costs (FeldmanHall & Shenhav, 2019). When encountering opponents in a game, individuals rapidly and automatically form initial impressions—including whether they are trustworthy, threatening or risk-seeking—based on opponent's features like skin color and clothing (Hughes et al., 2017). When information about others is scarce, people quickly resort to established social norms, assuming others are more likely to make decisions based on principles like cooperation and trust (Fleischhut et al., 2022). Automatic inference consumes minimal cognitive costs and can strongly constrain subsequent predictions about opponents. In contrast, controlled inference requires greater cognitive control to infer the opponent's motives and specific behaviors in a more nuanced manner, thereby updating the initial impression. People often engage in perspective-taking, adopting the viewpoint of others or their situational context to imagine or infer their perspectives and attitudes (Devaine et al., 2014; Galinsky et al., 2005). This process narrows the scope of predicting others' behaviors and further reduces social uncertainty.

Based on inferences about their opponent, individuals need to engage in appropriate policy selection in games. Traditional behavioral game theory posits that policy selection is based on the principle of expected utility maximization (Von Neumann & Morgenstern, 1944). Individuals are assumed to rationally weigh the future rewards of different action sequences and choose the policy that yields the highest payoff. However, this principle faces explanatory limitations in social decision-making contexts characterized by high uncertainty (Colman, 2003). It relies on the individual's ability to form and update precise probabilistic beliefs about others' behavior, which poses a significant challenge when the opponent's strategy is unknown or dynamically changing. The recently developed active inference framework offers a novel perspective. It proposes that policy selection is grounded in the principle of free energy minimization – that is, minimizing prediction error (Parr & Friston, 2019). This process entails both minimizing the discrepancy between actual outcomes and the individual's preferred outcomes, and actively reducing uncertainty about the environment. Consequently, the driver of policy selection expands from merely pursuing reward to a dual pursuit of both pragmatic and epistemic value.

In repeated games, the outcomes resulting from policy selections directly guide an individual's subsequent decisions. Participants can directly observe opponents' responses to their behavioral choices, receive feedback between game rounds, and engage in learning (Behrens et al., 2007). They integrate new feedback evidence with prior inferences by weighting these information sources, thereby updating their predictions about the opponent (Speekenbrink & Shanks, 2010). Depending on whether the feedback are consistent with prior predictions, it can broaden or narrow the belief distribution concerning the opponent's likely behaviors (FeldmanHall & Shenhav, 2019). As the game unfolds, participants continuously gather information about the opponent, learning their behavioral patterns. The learning rate is influenced by uncertainty; at the beginning of repeated games, when social uncertainty is highest, the learning rate is also at its peak (Courville et al., 2006; Speekenbrink

& Shanks, 2010). Concurrently, participants face a critical exploration-exploitation trade-off during learning (Gershman, 2019; Krafft et al., 2021). They must weigh the choice between exploring unknown strategies to observe the opponent's response and gain information about the value of such behaviors, versus exploiting known strategies that maximize immediate payoff based on past experience (Speekenbrink & Konstantinidis, 2015). Excessive exploration may lead to reduced efficiency, while prematurely exploiting existing experience may result in missing out on better strategies.

Social gaming situations evoke high degrees of uncertainty, during which the decision-making process, as illustrated in **Figure 2**, involves interrelated mechanisms of inference, policy selection and learning. Automatic inference and controlled inference operate simultaneously as two poles of a continuum, constraining predictions regarding the intentions and behaviors of opponents. These predictions guide participants' strategic choices based on different principles and are continuously updated through learning in repeated games, thereby further reducing social uncertainty. Human social cognition resembles a sophisticated computational system, allowing individuals to predict, respond to, and coordinate with others' behaviors, thereby underpinning the stability of social interactions.

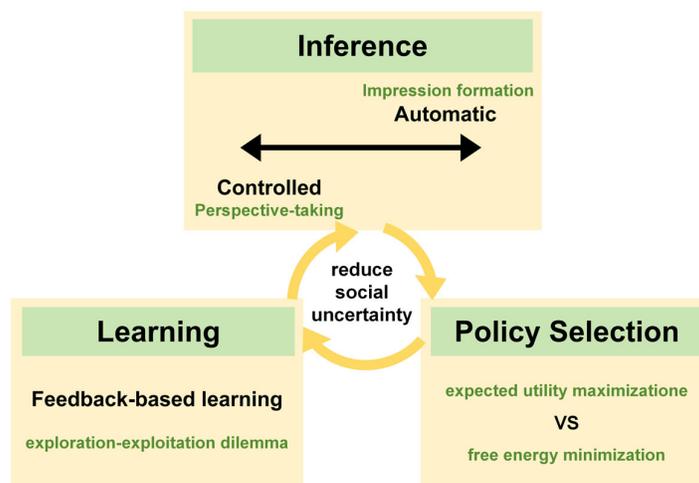


Figure 2. Decision-Making Process in Repeated Games.

4. Computational Modeling for Decision-Making

With the advancement of computational science, computational ideas pervade many areas of science and plays an integrative explanatory role in cognitive sciences and neurosciences (Montague et al., 2012). While focusing on human social behavior, social psychology is committed to explaining social life based on the general principles of human minds (Cushman & Gershman, 2019). Computational ideas has thus been increasingly adopted and utilized by social psychology researchers. Computational modeling offers a rigorous framework for describing abstract theories underlying human sociality in a clear and interpretable manner (Cushman, 2024). Since decision-making in repeated games involves recursive cognitive processes such as inference and learning—where recursivity, characterized by cyclic causal relationships, represents a core concept in computational modeling (Qinglin & Yuan, 2021). Additionally, computational modeling establishes a system of interacting relationships, allowing models of inference and learning to be integrated to capture more complex cognitive processes. Thus, employing computational modeling to investigate the underlying social psychological mechanisms in repeated games is highly feasible.

Current computational models for decision-making research can be primarily categorized into two major classes: reinforcement learning models and Bayesian models. Reinforcement learning models, grounded in the rational agent assumption, are widely used in the field to simulate how individuals learn from outcome feedback during interactive behaviors (Tomov et al., 2021; Tump et al., 2024; Yifrah et al., 2021). These models conceptualize the individual-environment interaction as a Markov Decision Process (MDP), in which the agent can observe all possible states of the environment and influence it through actions. As the state of the environment transitions, a reward function provides feedback to the individual, who in turn adapts their behavior based on the principle of maximizing cumulative rewards (Puterman, 1994). Researchers have integrated various cognitive strategies with reinforcement learning frameworks to develop a range of decision-making models. For instance, the Prospect-Valence Learning model posits that individuals evaluate the expected utility of different options and update expected valences using reinforcement learning rules to guide decisions (Ahn et al., 2008). Similarly, the Valence-Plus-Perseverance model combines heuristic strategies with reinforcement learning, proposing that choices are influenced by previous actions and their outcomes, which are integrated with expected valence to inform subsequent decision-making (Worthy et al., 2013).

The rational agent assumption of reinforcement learning models is not fully consistent with individuals' decision-making behaviors in daily life, and the environment in which individuals operate is not fully observable—instead, it is fraught with uncertainty. Therefore, unlike reinforcement learning models, Bayesian models for decision-making are based on the Partially Observable Markov Decision Process (POMDP), which formalizes the state of the decision-making environment as only partially observable (Itoh & Nakamura, 2007). Bayesian models assume that individuals exhibit bounded rationality and hold their own prior beliefs about environmental states; they can update these beliefs based on feedback obtained from interactions to form posterior, and then make decisions accordingly. In comparison, the Bayesian framework conceptualizes decision-making as a process of belief updating based on both internal and external information under uncertainty, thereby offering a more reasonable explanation for decision-making behavior in everyday social interactions (FeldmanHall & Shenhav, 2019). Meanwhile, the Bayesian framework does not require individuals to behave in a strictly Bayes-optimal manner. Rather, approximate Bayesian inference processes can effectively account for human decision-making behavior.

5. Basic Concepts of Active Inference

Active inference, as a specific implementation and computational framework of approximate Bayesian inference, is founded on a key premise: that organisms are endowed with an internal generative model of the statistical regularities of their environment. Equipped with this model, an organism can infer the hidden causes of its perception and can select optimal actions to achieve desired outcomes. Simultaneously, the active inference framework is underpinned by two core concepts (Smith et al., 2022). The first is that decision-makers are not simply passive Bayesian observers. Instead, they actively engage with the environment to gather information and seek preferred observations. The second is Bayesian inference, which refers to the uncertainty-weighted updating of prior beliefs based on the observed distribution of possible outcomes, as formalized by Bayes' theorem:

$$p(s|o) = \frac{p(o|s)p(s)}{p(o)} \quad (1)$$

In Bayes' theorem, the left-hand side, $p(s|o)$ represents the posterior belief about possible states (s)—that is, the updated belief distribution after incorporating a new observation (o). Here, s is an abstract variable that can represent any entity about which beliefs can be formed. On the right-hand side, $p(s)$ denotes the prior belief about s before acquiring the new observation. The term $p(o|s)$ is the likelihood term, representing the probability of observing o given that the state is s . The denominator $p(o)$ represents the model evidence, also known as the marginal likelihood, which indicates the total

probability of observing o . It serves as a normalization constant, ensuring that the posterior belief constitutes a valid probability distribution.

For instance, in a cooperation game, we need to infer whether an opponent is more inclined to cooperate or defect. Here, s represents the opponent's behavioral type. Assuming the opponent can be coarsely categorized into two types: s_1 represents cooperative type and s_2 represents defecting type. The variable o corresponds to the observed actual behavior of the opponent in each round of the game. Prior to the game, we can form an initial impression of the opponent through social inference. For example, if we believe the opponent is inclined to cooperate, we may assign a higher prior probability to s_1 , such as $p(s_1) = 0.8$, and consequently $p(s_2) = 0.2$. Suppose during the feedback phase of the game, we observe our opponent's betrayal. In this context, $p(o|s_1)$ signifies the probability that a cooperative-type opponent would defect; this value is expected to be low, for instance, $p(o|s_1) = 0.3$. Conversely, $p(o|s_2)$ denotes the probability that a defecting-type opponent would defect, which would be higher, e.g., $p(o|s_2) = 0.7$. The marginal likelihood $p(o)$, which acts as a normalizing constant, is calculated as the total probability of observing a defection across all possible opponent types: $p(o) = p(o|s_1) \times p(s_1) + p(o|s_2) \times p(s_2) = (0.3 \times 0.8) + (0.7 \times 0.2) = 0.38$. The posterior belief that the opponent is cooperative is then updated via Bayes' theorem: $p(s_1|o) = [p(o|s_1) \times p(s_1)] / p(o) = (0.3 \times 0.8) / 0.38 \approx 0.632$. Similarly, the posterior belief that the opponent is defecting is: $p(s_2|o) = [p(o|s_2) \times p(s_2)] / p(o) = (0.7 \times 0.2) / 0.38 \approx 0.368$. Based on this Bayesian update, after observing one instance of defection, the player's belief about the opponent's behavioral pattern shifts: the probability assigned to the cooperative type decreases from 80% to approximately 63.2%, while the probability of the defecting type increases from 20% to 36.8%. In subsequent rounds of the game, this posterior belief becomes the new prior. The player continues to perform Bayesian updates based on the opponent's new actions, iteratively refining their beliefs to better approximate the opponent's true behavioral pattern.

In this simple example, we were able to easily compute Bayes' theorem numerically. However, beyond the simplest belief distributions, the marginal likelihood $p(o)$ in Bayes' theorem is computationally intractable. It requires summing the probability of observations under all possible states. As the number of state dimensions increases, the number of terms to sum grows exponentially. For instance, in a game, an opponent's type might be defined by multiple parameters such as cooperative tendency, risk aversion, and capability. In such cases, calculating $p(o)$ becomes an integration over an ultra-high-dimensional space, which is infeasible to perform directly. Since computing posterior beliefs through exact Bayesian inference is infeasible in complex models, approximation techniques are required to solve this problem. This computational infeasibility is the fundamental reason for the central concept of variational free energy (VFE) in active inference.

Since the exact posterior distribution $p(s|o)$ is computationally intractable, a simple approximate posterior distribution $q(s)$ is introduced. By adjusting parameters through optimization algorithms, $q(s)$ is made to approximate $p(s|o)$ infinitely closely, transforming the uncomputable Bayesian inference into an optimization problem of minimizing a computable function. The metric used to quantify the discrepancy between two distributions is the Kullback–Leibler Divergence. The closer two distributions match, the smaller the KL divergence. Through mathematical derivation based on the relevant equation, the divergence between $q(s)$ and $p(s|o)$ can be expressed as follows:

$$\begin{aligned}
D_{KL}[q(s)\|p(s|o)] &= \sum_s q(s) \ln \frac{q(s)}{p(s|o)} \\
&= \sum_s q(s) [\ln q(s) - \ln p(s|o)] \\
&= \sum_s q(s) \left[\ln q(s) - \ln \frac{p(o,s)}{p(o)} \right] \\
&= \sum_s q(s) [\ln q(s) - \ln p(o,s) + \ln p(o)] \\
&= \sum_s q(s) \ln \frac{q(s)}{p(o,s)} + \ln p(o) \\
&= E_{q(s)} \left[\ln \frac{q(s)}{p(o,s)} \right] + \ln p(o)
\end{aligned} \tag{2}$$

According to the equation(2), in order to make $q(s)$ as close as possible to $p(s|o)$, it is necessary to minimize $D_{KL}[q(s)\|p(s|o)]$. Since $\ln p(o)$ is a constant term independent of $q(s)$, minimizing the KL divergence is equivalent to minimizing $E_{q(s)} \left[\ln \frac{q(s)}{p(o,s)} \right]$, which is defined as the variational free energy (F). Since $D_{KL}[q(s)\|p(s|o)] \geq 0$, it can be inferred through equation transformation that the variational free energy $F \geq -\ln p(o)$. $-\ln p(o)$ represents the surprise of sensory input, where $p(o)$ denotes the overall probability of obtaining an observation, also known as model evidence. A lower value of $p(o)$ corresponds to a higher degree of surprise in the sensory input. For example, the probability of seeing an elephant in an office is very low, so the surprise associated with such an observation is high. As the variational free energy (F) is an upper bound on surprise, minimizing the variational free energy consequently minimizes surprise indirectly.

The equation for variational free energy and its further derivation are shown below. It is worth noting that in the active inference framework, variational free energy is typically computed conditioned on different policies (π). This is because active inference emphasizes that agents actively infer future observations based on potential courses of action, hence the approximate posterior and other quantities are conditioned on policies.

$$\begin{aligned}
F_\pi &= E_{q(s|\pi)} \left[\ln \frac{q(s|\pi)}{p(o,s|\pi)} \right] \\
&= E_{q(s|\pi)} [\ln q(s|\pi) - \ln p(o,s|\pi)] \\
&= E_{q(s|\pi)} [\ln q(s|\pi) - \ln p(s|\pi)] - E_{q(s|\pi)} [\ln p(o|s,\pi)] \\
&= D_{KL}[q(s|\pi)\|p(s|\pi)] - E_{q(s|\pi)} [\ln p(o|s)]
\end{aligned} \tag{3}$$

According to the last line of the equation(3), minimizing variational free energy is equivalent to minimizing $D_{KL}[q(s|\pi)\|p(s|\pi)]$ while maximizing $E_{q(s|\pi)} [\ln p(o|s)]$. $D_{KL}[q(s|\pi)\|p(s|\pi)]$ refers to the difference between the approximate posterior distribution $q(s|\pi)$ and the prior distribution $p(s|\pi)$ under a given policy π . It represents the cost incurred to align posterior beliefs with prior expectations, serving as a measure of complexity. Minimizing this term means the agent tends to choose policies that minimize changes in its beliefs. Excessively high complexity leads to a phenomenon analogous to overfitting in statistics, where the agent adjusts its model to accommodate the randomness of a particular observation, thereby reducing its overall predictive capability. $E_{q(s|\pi)} [\ln p(o|s)]$ is the expected log-likelihood, i.e., the expected logarithmic probability of observing o given state s . It reflects how accurately the model predicts observations based on beliefs. Therefore, minimizing variational free energy is achieved by simultaneously minimizing model complexity and maximizing predictive accuracy.

The active inference framework treats perception and learning as processes of minimizing variational free energy (Friston, 2010). Perception corresponds to updating posterior beliefs in real-time based on each new observation, providing the best interpretation of sensory input; learning corresponds to gradually adjusting model parameters over long-term observations to align with accumulated experience. In the processes of perception and learning, the agent is not solely concerned

with finding the best-fitting posterior. It also strives to update its beliefs in the most parsimonious way, avoiding excessive deviations from prior beliefs, thereby achieving a balance between accuracy and complexity.

The active inference addresses not only the processing of past and present information but also encompasses the planning and action selection concerning future states. Similar to the principles underlying perception and learning, the goal of planning and action selection is to choose a policy π that minimizes variational free energy in the future. The key distinction lies in the extension of the variational free energy to incorporate expected future observations, yielding the expected free energy (EFE). The specific derivation of the expected free energy (G) is shown below. The first two lines of its equation closely resemble those of the variational free energy, with the only difference being the inclusion of future observations o under expectation. After decomposing expected free energy into two components related to information seeking and reward seeking in the third line, the fourth line introduces the agent's preferences C . As a decision-making model, active inference similarly requires encoding preferences. Unlike reinforcement learning, which encodes preferences as an external reward function, active inference internalizes them as part of the generative model by incorporating preferred observations $p(o|C)$ into the expected free energy.

$$\begin{aligned}
 G_{\pi} &= E_{q(o,s|\pi)} \left[\ln \frac{q(s|\pi)}{p(o,s|\pi)} \right] \\
 &= E_{q(o,s|\pi)} [\ln q(s|\pi) - \ln p(o,s|\pi)] \\
 &= E_{q(o,s|\pi)} [\ln q(s|\pi) - \ln p(s|o,\pi)] - E_{q(o|\pi)} [\ln p(o|\pi)] \\
 &\approx E_{q(o,s|\pi)} [\ln q(s|\pi) - \ln q(s|o,\pi)] - E_{q(o|\pi)} [\ln p(o|C)] \\
 &= -E_{q(o,s|\pi)} [\ln q(s|o,\pi) - \ln q(s|\pi)] - E_{q(o|\pi)} [\ln p(o|C)]
 \end{aligned} \tag{4}$$

In the equation(4), the first term $E_{q(o,s|\pi)} [\ln q(s|o,\pi) - \ln q(s|\pi)]$ in the final line represents the epistemic value, which expresses the expected log difference between the approximate posterior (based on the policy and the observations it generates) and the prior (based on the policy). A greater difference between the prior and the posterior indicates that the agent acquires more new information, implying a greater information gain from the policy and, consequently, a higher epistemic value. According to the equation, a higher epistemic value leads to a lower expected free energy for that policy. The second term $E_{q(o|\pi)} [\ln p(o|C)]$ represents the pragmatic value, which expresses the expected log probability of the agent's preferred observations. A higher probability of preferred outcomes indicates that the agent is more likely to obtain rewards, resulting in a greater pragmatic value. Similarly, a higher pragmatic value also leads to a lower expected free energy.

Planning and action selection are conceptualized as a process of minimizing expected free energy in active inference—that is, the agent seeking a policy that maximizes the sum of pragmatic value and epistemic value (Hodson et al., 2024; Parr & Friston, 2019). Consequently, during behavioral selection, the agent not only seeks to maximize reward returns by exploiting known resources but also pursue exploring unknown information to reduce uncertainty. Expected free energy offers a principled solution to the exploration-exploitation dilemma in repeated games (Gijzen et al., 2022). Exploration seeking epistemic value and exploitation seeking pragmatic value are regarded as two equally important aspects of minimizing expected free energy. The exploration-exploitation trade-off transitions from a sequential decision-making problem in reinforcement learning to the optimization of a single objective function associated with expected free energy minimization. The choice to explore or exploit depends on current levels of uncertainty and the level of expected reward (Smith et al., 2022). It is worth emphasizing that the epistemic value term within expected free energy formally instantiates a mechanism for directed exploration, analogous to curiosity, driving the agent to autonomously and actively seek observations that reduce uncertainty about hidden states (Friston et al., 2015; Parr & Friston, 2017).

6. Why Active Inference? Advantages for Decision-Making in Repeated Games

Active inference, a theoretical framework grounded in generative models typically formalized as Partially Observable Markov Decision Processes (POMDP), offers depth and breadth for understanding decision-making in repeated games. Its strength lies not only in its ability to unify multiple cognitive processes under the single principle of free energy minimization, but also in its capacity to more closely capture the core essence of human social interaction. Future research can validate the effectiveness of active inference models through computational model comparison and simulation. Furthermore, by fitting these models to behavioral game data and neuroimaging data, we can rigorously test the psychological and neural significance of their parameters, thereby bridging the gap between computation, behavior, and brain function.

6.1. The Free-Energy Principle: A Unified Framework for Cognitive Integration and Behavioral Optimization

The free-energy principle posits that any self-organizing system (such as the brain) within a changing environment must minimize its free energy to preserve the homeostasis required for survival (Friston et al., 2006). Free energy, serving as an upper bound on surprise, is a tractable measure of prediction error. The most fundamental theoretical advantage of the active inference framework lies precisely in this free-energy principle. It transcends the paradigm of traditional models that treat perception, learning, planning, and action selection as separate modules, offering instead a unified and biologically plausible computational framework. This framework conceptualizes all these cognitive processes as different manifestations of a single principle: free energy minimization (Friston, 2010). This characteristic endows it with exceptional explanatory power for modeling the complex decision-making processes inherent in repeated games. Within the context of repeated games, a player's perception of their opponent (inferring their hidden states), immediate action selection (choosing to cooperate or defect), and learning from multi-round interactions (updating beliefs about the opponent's behavioral patterns) are no longer treated as separate cognitive processes. Instead, they are integrated into a unified process of active inference under this single framework.

Conventional cognitive science research often presupposes different optimization objectives for distinct cognitive functions such as perception and action. For instance, perception optimizes for accuracy, while action optimizes for utility. In social decision-making, the accuracy of individuals' inferences about another person's thoughts and feelings is critically important, as it pertains to prediction, control, and decision outcomes (Vorauer et al., 2025). Researchers focus on how to optimize the accuracy of interpersonal perception within social interactions (Kenny & Albright, 1987). At the action level, from a game theory perspective, the goal of an individual in a game is to continually optimize their behavior to approximate an optimal strategy, seeking to maximize their own utility (Camerer, 2003). Faced with these perception and action that have different optimization targets, the active inference framework allows them to revolve around the same fundamental objective: minimizing the model's prediction error—the minimization of free energy. In terms of perception, minimizing free energy involves updating the internal state of others' thoughts and feelings to reduce the discrepancy between the "anticipated thoughts of others" and the "actually observed interactive cues," ultimately enhancing the accuracy of interpersonal perception. In terms of action, minimizing free energy involves adjusting one's own behavior to make the "actual outcome of the interaction" align more closely with the "preferred anticipated outcome," thereby increasing behavioral utility.

Furthermore, for the core issue of the exploration-exploitation dilemma in repeated games, the free energy principle can similarly provide an endogenous solution, achieving a dynamic balance in behavioral optimization. Traditional reinforcement learning models often rely on manually tuned external parameters to balance exploratory and exploitative behaviors. For instance, in the ϵ -greedy strategy, ϵ represents the probability of exploration, while $1 - \epsilon$ corresponds to exploitation, with ϵ typically decaying over the learning process (Vermorel & Mohri, 2005). The setting and adjustment

of ε are usually based on empirical or mathematical considerations rather than a well-founded cognitive mechanism. In contrast, active inference internalizes this trade-off through expected free energy minimization, framing it as a unified optimization problem with a more principled cognitive foundation (Friston et al., 2015). In the mathematical formulation of expected free energy, pragmatic value drives exploitative behavior, whereas epistemic value promotes exploration (Kirsh & Maglio, 1994). Exploration and exploitation are integrated into a single objective function (surprise minimization), where the goal of action selection is to maximize the sum of epistemic and pragmatic values (Friston, 2010). Thus, compared to classical expected utility maximization, surprise minimization under the free energy principle does not contradict it but rather extends it by incorporating epistemic value. This allows the model to dynamically adapt to the game context. During early stages of interaction, when social uncertainty is high, epistemic value dominates, leading to pronounced exploratory behavior. As beliefs about the opponent become more precise, uncertainty decreases, the weight of epistemic value diminishes, and pragmatic value begins to guide decision-making, resulting in a natural shift toward exploitation. This dynamic balance—emerging from the natural evolution of internal belief states rather than external parameter tuning—significantly enhances the realism and explanatory power of the model in describing human behavior in repeated games. Moreover, the free energy principle provides a powerful account of commonly observed behaviors such as risk-taking in high-reward contexts and curiosity-driven exploration in the absence of extrinsic incentives (Smith et al., 2022). By incorporating the exploration-exploitation trade-off into the unified free energy principle, the active inference framework demonstrates strong explanatory power in both behavioral optimization and cognitive integration.

6.2. Simulating the Nature of Social Interaction

The core challenge of social decision-making stems from social uncertainty, which is ubiquitous in the social world—not merely in repeated game scenarios. During social interactions, the thoughts and intentions of others are largely hidden, making it difficult to infer their behavior and its implications for us (FeldmanHall & Shenhav, 2019; Kappes et al., 2019). Active inference, grounded in the framework of Partially Observable Markov Decision Processes (POMDP), naturally incorporates this uncertainty and unobservability into its models (Smith et al., 2022). Furthermore, active inference intrinsically incorporates the behavioral drive to resolve this uncertainty into its mathematical formulation. Extensive prior research has focused on exploring the influence of rewards and punishments on social decision-making, revealing that reward-related computations and neural circuits play a crucial role in guiding choices, similar to non-social decision-making (Ruff & Fehr, 2014). However, another equally important guiding factor has been overlooked: the motivation to reduce social uncertainty (Alchian, 1950). In social interactions, people often engage in exploratory behaviors to reduce uncertainty; these actions do not confer direct rewards and may even entail risks and costs. The emergence of such behavior indicates that beyond the practical rewards brought by decision-making, understanding others' thoughts and reducing social uncertainty themselves also possess intrinsic value—a value distinct from external rewards (Loewenstein, 1994). In social interactions, individuals possess a proactive desire to explore others' thoughts and intentions, known as interpersonal curiosity (Way & Taffe, 2025). To satisfy interpersonal curiosity, individuals need to reduce uncertainty. The expected free energy of active inference effectively incorporates the epistemic value of behavioral choices. Epistemic value, also known as intrinsic value, mathematically corresponds to information gain or a reduction of uncertainty (Friston et al., 2015). According to the theory of expected free energy, an individual's behavioral choices aim not only to maximize pragmatic value but are equally driven by the maximization of intrinsic epistemic value. This computational mechanism clearly explains why individuals exhibit interpersonal curiosity during social interactions.

In social interactions, individuals need to infer others' thoughts, perspectives, emotional states, and behavior patterns, and more. This involves both automatic inference that form initial impressions and controlled inference like perspective-taking (Devaine et al., 2014; Hughes et al., 2017). When

modeling this mentalizing process in repeated game scenarios, the hierarchical model of active inference offers an elegant framework (Friston et al., 2017; Proietti et al., 2023). Its Bayesian network representation is shown in **Figure 3d**. To facilitate understanding of the complex hierarchical model, **Figure 3** illustrates the evolution of various active inference models, progressing from simple to complex. In the figure, circles represent variables, squares represent factors that modulate relationships, and arrows indicate dependencies between variables. **Figure 3a** depicts a generative model for static perception, involving only a single time point, analogous to standard Bayesian inference. Consistent with the earlier description of Bayes' theorem, s represents the abstract hidden state, o denotes the observable outcome, D represents the prior belief about the hidden state s , and A is the likelihood function, specifying the probability of observing o given the state s . **Figure 3b** shows a generative model for dynamic perception. Compared to 3a, it incorporates two or more time points and includes a state transition matrix B , which describes how the hidden state s evolves over time. **Figure 3c** builds upon 3b by adding policy selection, representing dynamic perception with action choices. Here, π denotes a policy, where different policies correspond to different state transition matrices. G represents the expected free energy, and C represents the preferred outcomes. **Figure 3d** illustrates the hierarchical model, which comprises two levels of generative models. In this architecture, the hidden state s^2 at the higher level provides the prior belief for the hidden state s^1 at the lower level. Conversely, the posterior belief of the lower-level model is treated as an observation at a given time point for the higher-level model. The likelihood matrix A^2 of the higher-level model mediates this bidirectional information flow between levels. This structure allows the higher-level model to update at a slower temporal scale compared to the lower-level model, which is why this framework is also referred to as the deep temporal model.

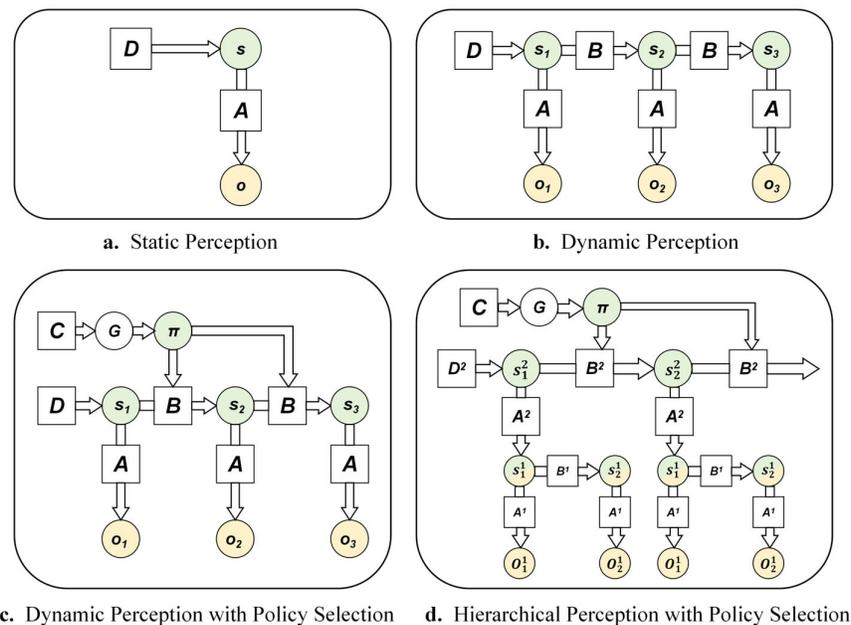


Figure 3. Bayesian network representation of various active inference models.

In repeated game scenarios, understanding the thoughts and goals behind others' actions typically requires observing a series of specific behaviors within the game. This is analogous to inferring a higher-level intention behind a shot (e.g., forehand or backhand) in tennis by observing the movement characteristics of various parts of the opponent's body (Proietti et al., 2023). Within the hierarchical active inference model, the agent employs its lower-level to process the specific choices in each round of the game, while the upper level is used to characterize the opponent's higher-level,

more stable attributes—such as intentions, strategies, or personality traits. For example, in multi-round prisoner's dilemma games, the prior beliefs D^2 in the upper-level model represent the automatically inferred initial impression of the opponent. The hidden states s^2 represent the opponent's different strategic modes (such as cooperative, deceptive, tit-for-tat, etc.). The opponent's strategy may change slowly over time, and the individual undergoes slow updating through perspective-taking and learning processes, which aligns with the characteristics of the upper-level model. The high-level beliefs about the opponent's strategy pattern, combined with the high-level likelihood matrix A^2 , generate predictions regarding the opponent's specific behavior in the next round. These predictions serve as prior beliefs for the lower-level model, guiding its own specific behavioral choices. The hierarchical model offers considerable flexibility, allowing for specific adjustments based on different game contexts. It enables the modeling of an individual's high-order beliefs, making it possible to simulate mentalizing during social interactions, thereby providing a closer approximation to human social decision-making processes.

6.3. Future Directions: Model Simulation and Behavioral Fitting

The explanatory power of active inference is demonstrated not only at the theoretical level, such as through the free energy principle, but also in its provision of a computationally tractable and testable generative model that effectively bridges theory and data. Through simulation studies, researchers have shown that active inference models generate predictions for perception and behaviour in domains like action understanding, cognitive control, and sport anticipation that closely match empirically observed patterns (Harris et al., 2022; Proietti et al., 2025; Proietti et al., 2023). Other researchers have employed model fitting to behavioural data from cognitive processes such as approach-avoidance conflict, interoceptive inference, and directed exploration, extracting key model parameters that serve as effective biomarkers for distinguishing between healthy individuals and those with psychopathology (Smith et al., 2021; Smith, Kuplicki, et al., 2020; Smith, Schwartenbeck, et al., 2020). Beyond computational psychiatry, studies fitting active inference models to behavioural data from sport anticipation tasks have also yielded parameters that effectively differentiate experts from novices (Harris et al., 2023). However, few researchers have yet applied active inference models to the research of social cognition using either simulation or behavioural fitting methods. The theoretical advantages of the active inference framework for understanding social decision-making in repeated games likewise require rigorous testing.

Model simulation and behavioral fitting are two complementary methodologies, exhibiting a progressive relationship that spans from theoretical verification to empirical testing. Model simulation constructs models based on theory to generate simulated data. It is a deductive process within an idealized environment, capable of testing a theory's internal consistency and generative capacity. Future research could define the state space (hidden states s , observations o), likelihood matrix A , and state transition matrix B for generative models based on specific game tasks. By adjusting parameters such as prior beliefs and the precision of expected free energy, researchers can observe whether the agent's policy patterns match those empirically observed. For instance, a hierarchical active inference model can be established based on the Prisoner's Dilemma game rules. By manipulating the priors of the higher-level model to simulate agents with different initial impressions of their opponent (e.g., cooperative vs. defecting), we could investigate whether the process of forming cooperation differs, thereby testing if the model can reproduce empirically observed strategic patterns. Building upon the preliminary validation of a model's effectiveness through simulation, behavioral fitting further applies the model to empirical behavioral data. Starting from real data, this approach compares the model's goodness of fit or estimates its parameters. As a model inversion process, behavioral fitting tests the theory's external validity and explanatory power. By collecting behavioral data (e.g., specific choices, reaction times, subjective reports) from participants in game experiments, the active inference model can be compared against classic computational models like reinforcement learning. This comparison tests whether the active inference model can effectively explain human behavioral data and evaluates its goodness of fit. The

parameters estimated from fitting can serve as computational biomarkers to distinguish between different groups and measure individual differences in social decision-making. Future work may design experimental tasks encompassing different types of games to collect human behavioral data. Using the same active inference model, fitting could be achieved by adjusting specific structural parameters based solely on differing game types. If the model successfully explains human behavior across diverse game contexts, it would demonstrate its potential and universality as a unified theory of social decision-making. Furthermore, integrating neuroimaging data with computational model parameters will help unveil the neurocomputational mechanisms underlying social decision-making, promoting the deeper application of active inference in social cognitive neuroscience.

It must be acknowledged, however, that applying active inference to social decision-making research presents challenges. Active inference remains a predominantly theoretical framework, and empirical application is still in its early stages. Its high computational complexity and large number of parameters make the modeling process relatively complex for researchers in psychology. Furthermore, existing guidelines for constructing active inference models are primarily based on relatively simple experimental tasks (Smith et al., 2022); transferring them to complex repeated game is a formidable undertaking. Concurrently, researchers from the philosophy of science have questioned the free energy principle (Colombo & Wright, 2021), arguing that it is too abstract and generalized, unable to replace research into the specific mechanisms underlying cognition and behavior. This serves as a crucial reminder that while we recognize the powerful theoretical explanatory power of active inference, equal emphasis must be placed on empirical research to validate and substantiate the concrete implementation of its computational models.

7. Conclusions

This review has systematically explored the potential of the active inference framework to illuminate the cognitive mechanisms of decision-making in repeated games. Unlike one-shot games, repeated games more closely resemble real-life social interactions, serving as a classic paradigm for studying social decision-making. The difficulty in observing others' thoughts and motivations introduces social uncertainty, which significantly increases the complexity of decision-making in repeated games. Individuals must engage in real-time learning and strategic adjustments based on feedback across multiple rounds of interaction, a process that involves interconnected cognitive components such as inference, strategy generation, and learning. Faced with such complex cognitive processes, traditional reinforcement learning models, while adept at capturing experience-driven value updating, have limitations in handling incomplete information and inferring opponents' strategies. Bayesian inference can accurately describe belief updating and uncertainty management, yet traditional exact Bayesian methods often struggle to fully integrate with action policy selection. Against this backdrop, the active inference framework, via the free-energy principle, unifies perception, learning, planning, and action within a single generative model structure. Perception and learning are achieved by minimizing variational free energy, balancing model complexity minimization with accuracy maximization. Planning and action selection are realized by minimizing expected free energy, maximizing both epistemic and pragmatic value. Compared to other models, the active inference framework offers a more powerful approach to tackling the critical exploration-exploitation dilemma inherent in repeated games. Formulated based on partially observable Markov decision processes (POMDP), it naturally incorporates social uncertainty. Furthermore, its hierarchical structure allows for a plausible representation of the mentalizing processes central to social interaction. Future research can further test the framework's empirical validity through complementary methods of model simulation and behavioral fitting. In summary, the active inference framework provides a unified and explanatory perspective for understanding human behavior in repeated games. It offers a rigorous methodology for uncovering the hidden internal dynamics of social decision-making, positioning itself as a potent tool for advancing computational social psychology.

Author Contributions: Conceptualization, H.Y. and L.W.; writing—original draft preparation, H.Y.; writing—review and editing, H.Y. and L.W.; supervision and discussion, T.T. and C.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science and Technology Major Project (2022ZD0116403).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Ahn, W.-Y., Busemeyer, J. R., Wagenmakers, E.-J., & Stout, J. C. (2008). Comparison of Decision Learning Models Using the Generalization Criterion Method. *Cognitive Science*, 32(8), 1376–1402, Article Pii 906363840. <https://doi.org/10.1080/03640210802352992>
- Akata, E., Schulz, L., Coda-Forno, J., Oh, S. J., Bethge, M., & Schulz, E. (2025). Playing repeated games with large language models [; Early Access]. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-025-02172-y>
- Alchian, A. A. (1950). Uncertainty, evolution, and economic theory. *Journal of political economy*, 58(3), 211–221.
- Bartoc, A. G. (1998). Course Notes Reinforcement Learning I: Introduction All Rights Reserved 1 Learning from Interaction.
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–1221. <https://doi.org/10.1038/nn1954>
- Bonowski, T., & Minnameier, G. (2022). Morality and trust in impersonal relationships. *Journal of Economic Psychology*, 90, Article 102513. <https://doi.org/10.1016/j.joep.2022.102513>
- Camerer, C. (2003). Behavioural studies of strategic thinking in games [Review]. *Trends in Cognitive Sciences*, 7(5), 225–231. [https://doi.org/10.1016/s1364-6613\(03\)00094-9](https://doi.org/10.1016/s1364-6613(03)00094-9)
- Cochard, F., Van, P., & Willinger, M. (2004). Trusting behavior in a repeated investment game. *Journal of Economic Behavior & Organization*, 55(1), 31–44. <https://doi.org/10.1016/j.jebo.2003.07.004>
- Colman, A. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences*, 26(2), 139–+. <https://doi.org/10.1017/s0140525x03000050>
- Colman, A. M., Pulford, B. D., & Krockow, E. M. (2018). Persistent cooperation and gender differences in repeated Prisoner's Dilemma games: Some things never change. *Acta Psychologica*, 187, 1–8. <https://doi.org/10.1016/j.actpsy.2018.04.014>
- Colombo, M., & Wright, C. (2021). First principles in the life sciences: the free-energy principle, organicism, and mechanism. *Synthese*, 198(SUPPL 14), 3463–3488. <https://doi.org/10.1007/s11229-018-01932-w>
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world [; Proceedings Paper]. *Trends in Cognitive Sciences*, 10(7), 294–300. <https://doi.org/10.1016/j.tics.2006.05.004>
- Cui, X., Bryant, D. M., & Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage*, 59(3), 2430–2437. <https://doi.org/10.1016/j.neuroimage.2011.09.003>
- Cushman, F. (2024). Computational Social Psychology [Review]. *Annual Review of Psychology*, 75, 625–652. <https://doi.org/10.1146/annurev-psych-021323-040420>
- Cushman, F., & Gershman, S. (2019). Editors' Introduction: Computational Approaches to Social Cognition [Editorial Material]. *Topics in Cognitive Science*, 11(2), 281–298. <https://doi.org/10.1111/tops.12424>
- Devaine, M., Hollard, G., & Daunizeau, J. (2014). The Social Bayesian Brain: Does Mentalizing Make a Difference When We Learn? *Plos Computational Biology*, 10(12), Article e1003992. <https://doi.org/10.1371/journal.pcbi.1003992>

17. FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world [Review]. *Nature Human Behaviour*, 3(5), 426–435. <https://doi.org/10.1038/s41562-019-0590-x>
18. Fleischhut, N., Artinger, F. M., Olschewski, S., & Hertwig, R. (2022). Not all uncertainty is treated equally: Information search under social and nonsocial uncertainty. *Journal of Behavioral Decision Making*, 35(2). <https://doi.org/10.1002/bdm.2250>
19. Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., & Pezzulo, G. (2016). Active inference and learning [Review]. *Neuroscience and Biobehavioral Reviews*, 68, 862–879. <https://doi.org/10.1016/j.neubiorev.2016.06.022>
20. Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, 6(4), 187–214. <https://doi.org/10.1080/17588928.2015.1020053>
21. Friston, K. J. (2010). The free-energy principle: a unified brain theory? [Review]. *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
22. Friston, K. J., Editor, R. R. G., Parr, T., Price, C., & Bowman, H. (2017). Deep temporal models and active inference [Review]. *Neuroscience and Biobehavioral Reviews*, 77, 388–402. <https://doi.org/10.1016/j.neubiorev.2017.04.009>
23. Friston, K. J., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3), 70–87. <https://doi.org/10.1016/j.jphysparis.2006.10.001>
24. Galinsky, A., Ku, G., & Wang, C. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination [; Proceedings Paper]. *Group Processes & Intergroup Relations*, 8(2), 109–124. <https://doi.org/10.1177/1368430205051060>
25. Gershman, S. J. (2019). Uncertainty and Exploration. *Decision-Washington*, 6(3), 277–286. <https://doi.org/10.1037/dec0000101>
26. Gijsen, S., Grundei, M., & Blankenburg, F. (2022). Active inference and the two-step task. *Scientific Reports*, 12(1), Article 17682. <https://doi.org/10.1038/s41598-022-21766-4>
27. Hackel, L. M., & Amodio, D. M. (2018). Computational neuroscience approaches to social cognition [Review]. *Current Opinion in Psychology*, 24, 92–97. <https://doi.org/10.1016/j.copsyc.2018.09.001>
28. Harris, D. J., Arthur, T., Broadbent, D. P., Wilson, M. R., Vine, S. J., & Runswick, O. R. (2022). An Active Inference Account of Skilled Anticipation in Sport: Using Computational Models to Formalise Theory and Generate New Hypotheses [Review]. *Sports Medicine*, 52(9), 2023–2038. <https://doi.org/10.1007/s40279-022-01689-w>
29. Harris, D. J., North, J. S., & Runswick, O. R. (2023). A Bayesian computational model to investigate expert anticipation of a seemingly unpredictable ball bounce. *Psychological Research-Psychologische Forschung*, 87(2), 553–567. <https://doi.org/10.1007/s00426-022-01687-7>
30. Hitchcock, P. F., Fried, E. I., & Frank, M. J. (2022). Computational Psychiatry Needs Time and Context [Review]. *Annual Review of Psychology*, 73, 243–270. <https://doi.org/10.1146/annurev-psych-021621-124910>
31. Hodson, R., Mehta, M., & Smith, R. (2024). The empirical status of predictive coding and active inference [Review]. *Neuroscience and Biobehavioral Reviews*, 157, Article 105473. <https://doi.org/10.1016/j.neubiorev.2023.105473>
32. Hughes, B. L., Zaki, J., & Ambady, N. (2017). Motivation alters impression formation and related neural systems. *Social Cognitive and Affective Neuroscience*, 12(1), 49–60. <https://doi.org/10.1093/scan/nsw147>
33. Itoh, H., & Nakamura, K. (2007). Partially observable Markov decision processes with imprecise parameters. *Artificial Intelligence*, 171(8), 453–490. <https://doi.org/https://doi.org/10.1016/j.artint.2007.03.004>
34. Kappes, A., Nussberger, A.-M., Siegel, J. Z., Rutledge, R. B., & Crockett, M. J. (2019). Social uncertainty is heterogeneous and sometimes valuable [Letter]. *Nature Human Behaviour*, 3(8), 764–764. <https://doi.org/10.1038/s41562-019-0662-y>
35. Kenny, D. A., & Albright, L. (1987). *Accuracy in interpersonal perception: A social relations analysis* [doi:10.1037/0033-2909.102.3.390]. American Psychological Association.
36. Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513–549. [https://doi.org/https://doi.org/10.1016/0364-0213\(94\)90007-8](https://doi.org/https://doi.org/10.1016/0364-0213(94)90007-8)

37. Krafft, P. M., Shmueli, E., Griffiths, T. L., Tenenbaum, J. B., & Pentland, A. (2021). Bayesian collective learning emerges from heuristic social learning. *Cognition*, 212, Article 104469. <https://doi.org/10.1016/j.cognition.2020.104469>
38. Lacomba, J. A., Lagos, F., Reuben, E., & van Winden, F. (2017). Decisiveness, peace, and inequality in games of conflict. *Journal of Economic Psychology*, 63, 216–229. <https://doi.org/10.1016/j.joep.2016.11.003>
39. Lee, D. (2008). Game theory and neural basis of social decision making [Review]. *Nature Neuroscience*, 11(4), 404–409. <https://doi.org/10.1038/nn2065>
40. Loennqvist, J.-E., & Walkowitz, G. (2019). Experimentally Induced Empathy Has No Impact on Generosity in a Monetarily Incentivized Dictator Game. *Frontiers in Psychology*, 10, Article 337. <https://doi.org/10.3389/fpsyg.2019.00337>
41. Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1), 75–98. <https://doi.org/10.1037/0033-2909.116.1.75>
42. Maurer, C., Chambon, V., Bourgeois-Gironde, S., Leboyer, M., & Zalla, T. (2018). The influence of prior reputation and reciprocity on dynamic trust-building in adults with and without autism spectrum disorder. *Cognition*, 172, 1–10. <https://doi.org/10.1016/j.cognition.2017.11.007>
43. Miller Moya, L. M. (2007). Coordination and collective action. *Revista Internacional De Sociologia*, 65(46), 161–183.
44. Montague, P. R. (2018). *Computational Phenotypes Revealed by Interactive Economic Games* (Publication Number 978-0-12-809826-4; 978-0-12-809825-7) [; Book Chapter, <Go to ISI>://WOS:000416850000013
45. Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry [Review]. *Trends in Cognitive Sciences*, 16(1), 72–80. <https://doi.org/10.1016/j.tics.2011.11.018>
46. Nash, J. F. (1950). Equilibrium points in *n*-person games. *Proceedings of the National Academy of Sciences*, 36(1), 48–49. <https://doi.org/doi:10.1073/pnas.36.1.48>
47. Ng, G. T. T., & Au, W. T. (2016). Expectation and cooperation in prisoner's dilemmas: The moderating role of game riskiness. *Psychonomic Bulletin & Review*, 23(2), 353–360. <https://doi.org/10.3758/s13423-015-0911-7>
48. Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers: An fNIRS-based hyperscanning study. *Human Brain Mapping*, 38(2), 831–841. <https://doi.org/10.1002/hbm.23421>
49. Parr, T., & Friston, K. J. (2017). Uncertainty, epistemics and active inference. *Journal of the Royal Society Interface*, 14(136), Article 20170376. <https://doi.org/10.1098/rsif.2017.0376>
50. Parr, T., & Friston, K. J. (2019). Generalised free energy and active inference. *Biological Cybernetics*, 113(5–6), 495–513. <https://doi.org/10.1007/s00422-019-00805-w>
51. Press, W. H., & Dyson, F. J. (2012). Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26), 10409–10413. <https://doi.org/doi:10.1073/pnas.1206569109>
52. Proietti, R., Parr, T., Tessari, A., Friston, K., & Pezzulo, G. (2025). Active inference and cognitive control: Balancing deliberation and habits through precision optimization [Review]. *Physics of Life Reviews*, 54, 27–51. <https://doi.org/10.1016/j.pprev.2025.05.008>
53. Proietti, R., Pezzulo, G., & Tessari, A. (2023). An active inference model of hierarchical action understanding, learning and imitation [Review]. *Physics of Life Reviews*, 46, 92–118. <https://doi.org/10.1016/j.pprev.2023.05.012>
54. Puterman, M. L. (1994). Markov Decision Processes: Discrete Stochastic Dynamic Programming.
55. Qinglin, G., & Yuan, Z. (2021). Psychological and neural mechanisms of trust formation: A perspective from computational modeling based on the decision of investor in the trust game [计算模型视角下信任形成的心理和神经机制--基于信任博弈中投资者的角度]. *Advances in Psychological Science*, 29(1), 178–189, Article 1671-3710(2021)29:1<178:Jsmxsj>2.0.Tx;2-3.
56. Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making [Review]. *Nature Reviews Neuroscience*, 15(8), 549–562. <https://doi.org/10.1038/nrn3776>
57. Schelling, T. (1960). *The Strategy of Conflict* Cambridge.
58. Smith, R., Friston, K. J., & Whyte, C. J. (2022). A step-by-step tutorial on active inference and its application to empirical data. *Journal of Mathematical Psychology*, 107, Article 102632. <https://doi.org/10.1016/j.jmp.2021.102632>

59. Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., McDermott, T. J., Taylor, S., Khalsa, S. S., Paulus, M. P., & Aupperle, R. L. (2021). Long-term stability of computational parameters during approach-avoidance conflict in a transdiagnostic psychiatric patient sample. *Scientific Reports*, *11*(1), Article 11783. <https://doi.org/10.1038/s41598-021-91308-x>
60. Smith, R., Kuplicki, R., Feinstein, J., Forthman, K. L., Stewart, J. L., Paulus, M. P., Khalsa, S. S., & Investigators, T. (2020). A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *Plos Computational Biology*, *16*(12), Article e1008484. <https://doi.org/10.1371/journal.pcbi.1008484>
61. Smith, R., Schwartenbeck, P., Stewart, J. L., Kuplicki, R., Ekhtiari, H., Paulus, M. P., & Investigators, T. (2020). Imprecise action selection in substance use disorder: Evidence for active learning impairments when solving the explore-exploit dilemma. *Drug and Alcohol Dependence*, *215*, Article 108208. <https://doi.org/10.1016/j.drugalcdep.2020.108208>
62. Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and Exploration in a Restless Bandit Problem. *Topics in Cognitive Science*, *7*(2), 351–367. <https://doi.org/10.1111/tops.12145>
63. Speekenbrink, M., & Shanks, D. R. (2010). Learning in a Changing Environment. *Journal of Experimental Psychology-General*, *139*(2), 266–298. <https://doi.org/10.1037/a0018620>
64. Stojic, H., Orquin, J. L., Dayan, P., Dolan, R. J., & Speekenbrink, M. (2020). Uncertainty in learning, choice, and visual fixation. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(6), 3291–3300. <https://doi.org/10.1073/pnas.1911348117>
65. Stojic, H., Schulz, E., Analytis, P., & Speekenbrink, M. (2020). It's New, but Is It Good? How Generalization and Uncertainty Guide the Exploration of Novel Options. *Journal of Experimental Psychology-General*, *149*(10), 1878–1907. <https://doi.org/10.1037/xge0000749>
66. Tomov, M. S., Schulz, E., & Gershman, S. J. (2021). Multi-task reinforcement learning in humans. *Nature Human Behaviour*, *5*(6), 764–+. <https://doi.org/10.1038/s41562-020-01035-y>
67. Tump, A. N., Deffner, D., Pleskac, T. J., Romanczuk, P., & Kurvers, R. H. J. M. (2024). A Cognitive Computational Approach to Social and Collective Decision-Making. *Perspectives on Psychological Science*, *19*(2), 538–551. <https://doi.org/10.1177/17456916231186964>
68. van Dijk, E., & De Dreu, C. K. W. (2021). Experimental Games and Social Decision Making. *Annual Review of Psychology*, *72*(Volume 72, 2021), 415–438. <https://doi.org/https://doi.org/10.1146/annurev-psych-081420-110718>
69. Vermorel, J., & Mohri, M. (2005). Multi-armed bandit algorithms and empirical evaluation. In J. Gama, R. Camacho, P. Brazdil, A. Jorge, & L. Torgo (Eds.), *Machine Learning: Ecm1 2005, Proceedings* (Vol. 3720, pp. 437–448).
70. Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.
71. Vorauer, J. D., Hodges, S. D., & Hall, J. A. (2025). Thought-Feeling Accuracy in Person Perception and Metaperception: An Integrative Perspective [Review]. *Annual Review of Psychology*, *76*, 413–441. <https://doi.org/10.1146/annurev-psych-011624-024416>
72. Walker, A. R., Luque, D., Le Pelley, M. E., & Beesley, T. (2019). The role of uncertainty in attentional and choice exploration. *Psychonomic Bulletin & Review*, *26*(6), 1911–1916. <https://doi.org/10.3758/s13423-019-01653-2>
73. Wang, H., & Kwan, A. C. (2023). Competitive and cooperative games for probing the neural basis of social decision-making in animals. *Neuroscience and Biobehavioral Reviews*, *149*, Article 105158. <https://doi.org/10.1016/j.neubiorev.2023.105158>
74. Way, N., & Taffe, R. (2025). Interpersonal Curiosity: A Missing Construct in the Field of Human Development. *Human Development*, *69*(2), 79–90. <https://doi.org/10.1159/000542162>
75. Worthy, D. A., Pang, B., & Byrne, K. A. (2013). Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task. *Frontiers in Psychology*, *4*, Article 640. <https://doi.org/10.3389/fpsyg.2013.00640>

76. Yifrah, B., Ramaty, A., Morris, G., & Mendelsohn, A. (2021). Individual differences in experienced and observational decision-making illuminate interactions between reinforcement learning and declarative memory. *Scientific Reports*, 11(1), Article 5899. <https://doi.org/10.1038/s41598-021-85322-2>
77. Zhang, M., Liu, T., Pelowski, M., & Yu, D. (2017). Gender difference in spontaneous deception: A hyperscanning study using functional near-infrared spectroscopy. *Scientific Reports*, 7, Article 7508. <https://doi.org/10.1038/s41598-017-06764-1>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.