**RESEARCH ARTICLE**

WILEY

# Neural correlates of recursive thinking during interpersonal strategic interactions

## Shanshan Zhen[1]  |  Rongjun Yu[2,3,4]

[1]Department of Psychology, National University of Singapore, Singapore, Singapore

[2]Department of Management, School of Business, Hong Kong Baptist University, Hong Kong, China

[3]Department of Sport, Physical Education and Health, Faculty of Social Sciences, Hong Kong Baptist University, Hong Kong, China

[4]Department of Physics, Faculty of Science, Hong Kong Baptist University, Hong Kong, China

**Correspondence**
Rongjun Yu, Department of Management, Hong Kong Baptist University, Hong Kong, China.
Email: rongjunyu@hkbu.edu.hk

## Abstract

To navigate the complex social world, individuals need to represent others' mental states to think strategically and predict their next move. Strategic mentalizing can be classified into different levels of theory of mind according to its order of mental state attribution of other people's beliefs, desires, intentions, and so forth. For example, reasoning people's beliefs about simple world facts is the first-order attribution while going further to reason people's beliefs about the minds of others is the second-order attribution. The neural substrates that support such high-order recursive reasoning in strategic interpersonal interactions are still unclear. Here, using a sequential-move interactional game together with functional magnetic resonance imaging (fMRI), we showed that recursive reasoning engaged the frontal-subcortical regions. At the stimulus stage, the ventral striatum was more activated in high-order reasoning as compared with low-order reasoning. At the decision stage, high-order reasoning activated the medial prefrontal cortex (mPFC) and other mentalizing regions. Moreover, functional connectivity between the dorsomedial prefrontal cortex (dmPFC) and the insula/hippocampus was positively correlated with individual differences in high-order social reasoning. This work delineates the neural correlates of high-order recursive thinking in strategic games and highlights the key role of the interplay between mPFC and subcortical regions in advanced social decision-making.

**KEYWORDS**

game theory, neuroimaging, recursive reasoning, sequential-move game, strategic thinking, theory of mind

## 1 | INTRODUCTION

Strategic reasoning about other's mental states and predicting their behaviors in terms of such states are the main determinants of social competence. The ability to reason about beliefs, desires, and intentions of another individual like one's own is referred to as mentalizing or theory of mind (ToM) (Premack & Woodruff, 1978), which is crucial to human social reasoning and social success. ToM reasoning entails a hierarchical classification of mental state

attribution that allows individuals to recursively infer other's mental states (Hedden & Zhang, 2002; Perner & Wimmer, 1985). Reasoning other's beliefs about real events are first-order ToM reasoning. For example: "I think that you believe that the box contains a pencil." However, first-order reasoning cannot adequately describe human interactions. In complex social interactions, people need to take into account that the other person also holds beliefs about the minds of other people, and so on (Bhatt & Camerer, 2011). Social inferences based on the ToM often take a recursive form, such as

the second-order ToM reasoning: "I think that you think that I think...."

A classic false-belief paradigm, the "Sally-Anne" task, is known for examining the first-order mental state attributions. For example, a participant was asked about where Sally thought her marble was after Anne moved the marble when Sally left the room (Wimmer & Perner, 1983). Research in developmental psychology found that children could generally pass this standard false-belief task by age 4 (Astington & Hughes, 2013), but would take a few more years of maturation to pass a similar task that required the second-order ToM (Fu, Xiao, Killen, & Lee, 2014; Keysar, Lin, & Barr, 2003; Symeonidou, Dumontheil, Chow, & Breheny, 2016; Tager-Flusberg & Sullivan, 1994). Even healthy adults are not perfect in performing all kinds of strategic tasks requiring high-order ToM reasoning, suggesting the presence of cognitive limits or difficulties in iterated thinking (Bhatt & Camerer, 2005; Brandenburger & Li, 2015; Camerer, 2009; Coricelli & Nagel, 2009; Hedden & Zhang, 2002).

The neural basis of first-order ToM has been extensively studied using false-belief tasks. Increased brain activity was found in the medial prefrontal cortex (mPFC), temporoparietal junction (TPJ), superior temporal sulcus (STS), and precuneus (PC) when human subjects represented other's false beliefs (Carrington & Bailey, 2009; Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011; Gallagher et al., 2000; Gallagher & Frith, 2003; Saxe & Kanwisher, 2003; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Spunt et al., 2015). These false belief tasks usually involve a situation in which one's models of the world are different from models of others' mental states. Typically, the tasks require participants to be observers, that is, having no interactions with characters in the scenarios so their own benefits of the task would not be affected by their predictions of the characters. This kind of false belief task is not suitable to investigate neural mechanisms underlying recursive reasoning in social interaction, especially when the depth of ToM can affect reasoners' social outcomes. It is intriguing to find out whether high-level strategic thinking recruits the putative ToM neural network (e.g., mPFC, TPJ, STS) or involves other special brain regions adapted specifically for high-order reasoning. The answer to this research question may help us understand whether ToM is a domain-general concept, or it is a multidimensional construct. It may also contribute to our understanding of the variability in the human ability to make inferences about the minds of others in complex social interactions (Conway, Catmur, & Bird, 2019). Following this, it is also critical to identify the neural correlates of the heterogeneity observed in human complex strategic behavior, which may help to explain why some people are able to deeply think about others to gain more successful social outcomes while some make decisions without considering others' behavior.

Only a few studies have examined higher-order ToM reasoning using a range of simple yet interactive games. In strategic interpersonal games, players have to infer each other's beliefs as one player's outcome is dependent on the other player's choice and vice versa. For example, (Coricelli & Nagel, 2009) showed that subjects who reasoned using high-level ToM to play the "beauty contest" game had increased activity in the mPFC relative to subjects who reasoned using low-level

ToM. Another neuroimaging study found that activity in anterior mPFC tracked the degree of changes in the prediction of other's strategies given one's own play in an inspection game (Hampton, Bossaerts, & O'Doherty, 2008). Yoshida, Seymour, Friston, and Dolan (2010) further found that activity in mPFC coded participants' uncertainty about co-players' actions while the level of recursion they engaged in during decision-making was represented in the dorsolateral prefrontal cortex (dlPFC). In addition, Zhu, Mathewson, and Hsu (2012) using a competitive game found that prediction errors for reward were coded in the ventral striatum (VS) while prediction errors for belief updating on the opponent's actions were coded in the ventral striatum as well as the rostral anterior cingulate cortex (ACC). Taken together, these findings seem to point out the important role of the prefrontal cortex in encoding beliefs of others to predict future behaviors during social interaction.

However, in the studies of social interaction, mentalizing is often embedded in probabilistic social learning tasks in which opponents' choices are not fully predictable due to their social preferences (e.g., over short-term gains or long-term gains) and insufficient computational resources (e.g., decision-makers not certain about their partners' level of strategic reasoning ability). Game theory often assumes players are more rational than they may be capable of in reality (Camerer, 1991). In this case, mentalizing and probabilistic learning may be mingled together and hard to differentiate. In deterministic games, however, players can apply a more analytic, rule-based mechanism of recursive reasoning, which may substantially differ from the probabilistic one. It is unclear whether the pattern of neural activity in probabilistic social learning holds true in deterministic situations. Here, we aimed to investigate the neural correlates of recursive reasoning in a deterministic game in which the unique optimal option chosen by the opponents can be inferred by participants before making choices at each step.

Therefore, we adapted a sequential-move game from previous studies to investigate the neural correlates of recursive reasoning within the context of a marble drop game (Hedden & Zhang, 2002; Meijering, Van Rijn, Taatgen, & Verbrugge, 2012). In this marble drop game, two players make the decision sequentially and their outcomes are affected by each other's choice. Importantly, perfect information is embedded in the game in which both players have common information about the possible payoffs for each action so that the outcome of each action is predictable. This game provides a tool to investigate whether players in strategic interactions mentally process other's thoughts and beliefs, and if so, how the mental processing unfolds. To achieve the goal of the task, the first player at the first decision point (i.e., the first trapdoor in Figure 1a) has to apply second-order ToM reasoning: think about what Player 2 at the second decision point believes about Player 1's decision at the final decision point.

In a sequential-move game with finite reasoning steps, the game theory assumes that players should use backward induction, that is, the process of thinking backward from the final decision point to decide how to achieve the highest payoff (Aumann, 1995; Osborne & Rubinstein, 1994). Although backward reasoning is an efficient way to maximize payoff because the optimal outcome is known at each
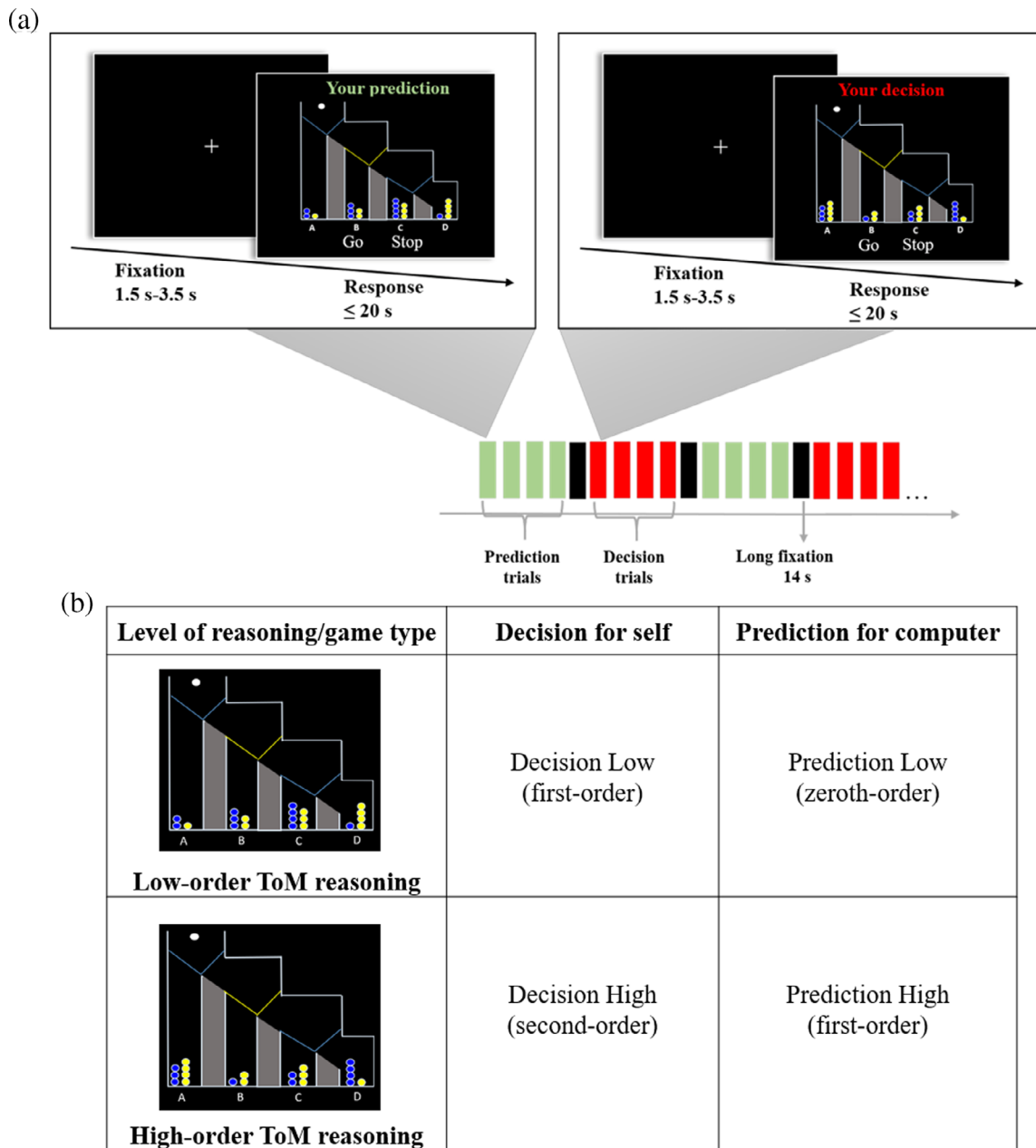
(a)



(b)

| Level of reasoning/game type | Decision for self | Prediction for computer |
|---|---|---|
| **Low-order ToM reasoning** | Decision Low (first-order) | Prediction Low (zeroth-order) |
| **High-order ToM reasoning** | Decision High (second-order) | Prediction High (first-order) |

**FIGURE 1** Experimental design. (a) Task procedure. There are two types of trials in this game (i.e., Prediction for computer and Decision for self). On Prediction trials, we presented the green word "Your prediction" on top of the screen with the stimuli. During these trials, participants were asked to record their prediction about the decision of the computer at the second decision point/trapdoor by choosing an option labeled "Go" or "Stop." On the contrary, on Decision trials, we presented the red word "Your decision" on top of the screen with the stimuli. When seeing this, participants were asked to make a choice about what they should do at the first decision point/trapdoor. The order of trials was pseudo-randomly determined for each participant: four trials under the same type of the game were presented consecutively, followed by a long ITI (i.e., a fixation crosshair) for 14 s; after that, another four trials from the other game type were presented sequentially. Every four trials consisted of two trials of low-order ToM reasoning and two high-order ToM reasoning. (b) Experimental conditions. There were four unique types of trials based on the combination of level of reasoning (low- vs. high-order ToM reasoning) and game type (decision for self vs. prediction for computer). The levels of reasoning were designed by manipulating the payoff structures

decision point, it is not the only kind of strategy adopted under second-order attribution in strategic games (Bergwerff, Meijering, Szymanik, Verbrugge, & Wierda, 2014). By using the two-players marble drop game, Bergwerff et al. (2014) proposed that a combination of forward and backward reasoning yields a faster solution (i.e., a smaller

number of reasoning steps) than backward induction does. Based on previous studies, we counted the reasoning steps (see the methods) to attain each possible payoff using forward reasoning plus backtracking to model the neurocognitive subprocesses underlying different levels of strategic thinking. In our model of strategic thinking,

a high ToM agent would first think by forward reasoning about how other people get the highest possible payoff, and then figured out how that would affect their own highest possible payoff by using backward reasoning.

The normative solutions of sequential-move games require recursive modeling of other players' thinking to its full depth. However, empirical evidence shows that individuals often fail to fully employ back-and-forth recursive reasoning, possibly due to the limits in working memory and cognitive resources (Carlson, Moses, & Breton, 2002; Schiebener & Brand, 2015). In the present study, we included two types of games that required different depths of recursive reasoning to investigate the neural representation of high-order ToM and low-order ToM reasoning. Items used in the high-order ToM condition were carefully designed to require the second-order ToM reasoning to make the best move. In the low-order ToM condition, participants did not have to think about what Player 2 might think about what they would do at the last decision point because the payoff for Player 2 in bins C and D were either lower or higher than bins A and B (see the methods). In this case, Player 2 could make an optimal decision regardless of Player 1's decision at the last decision point. We also included a manipulation in which participants had to make predictions about player 2's move. In the high-order ToM condition, player 2 had to reason about what Player 1 would do at the next stage. In this case, Player 2 was required first-order reasoning. In the low-order condition, Player 2's decision was independent of player 1's next move and only zeroth-order reasoning (i.e., considering only one's own desires and facts) was needed.

Since generating the internal predictions of others is essential for constructive sequential decisions for oneself when playing the game, we specifically examined this prediction process by asking participants to predict others' next move. This third-person (i.e., anticipation of other's move) perspective allowed us to determine whether the second-order ToM network (second vs. first order) for self-decision is also involved in the reasoning for other agent (first vs. 0th order) during the time of anticipation. Our design allowed us to contrast high-order reasoning with low-order reasoning in both first-person and third-person perspectives, and also allowed us to examine whether the neural activity in response to each level of ToM reasoning modulated by the number of reasoning steps. We predicted that the mPFC would be more engaged in the condition that requires high-order reasoning, given its important role in encoding the computation, maintenance, and proactive use of the perspective signal to guide the selection of appropriate action (Hillebrandt, Dumontheil, Blakemore, & Roiser, 2013). Previous studies posited that mentalizing and reward-based learning may share similar Bayesian inference mechanisms (Ahn, Krawitz, Kim, Busemeyer, & Brown, 2013; Aitchison & Lengyel, 2017; Baker & Tenenbaum, 2014; Behrens, Hunt, Woolrich, & Rushworth, 2008; Devaine, Hollard, & Daunizeau, 2014; Khalvati et al., 2019; Robalino & Robson, 2012; Zhu et al., 2012). Hence, we hypothesized that the striatum, a region known to encode both reward and social prediction errors (Báez-Mendoza & Schultz, 2013; Daniel & Pollmann, 2014; Fliessbach et al., 2007; Joiner, Piva, Turrin, & Chang, 2017; Pagnoni, Zink, Montague, & Berns, 2002), might be

involved in the condition that requires high-order reasoning, especially at the stimulus stage. We further predicted that the mPFC would interact with other regions, such as the hippocampus and insula, areas involved in memory-based learning and strategic uncertainty (Eichenbaum, 2017; Nagel, Brovelli, Heinemann, & Coricelli, 2018; Seger & Cincotta, 2006; Tavares et al., 2015), to determine individual differences in the ability to implement high-level reasoning.

## 2 | MATERIALS AND METHODS

### 2.1 | Participants

Thirty-one right-handed volunteers (13 women; mean age ± SD, 22.92 ± 2.60 years) were paid to participate in the fMRI study. Participants reported no neurological or psychiatric history and had a normal or corrected-to-normal vision. Participants were given Singapore dollar S\$20 ($\sim$S\$1 = US\$0.73) for showing-up and a performance-based bonus (see below). All participants provided informed consent and the study was approved by the local Institutional Review Board. One of these subjects (woman, age 20) was excluded for image analysis due to technical problems in fMRI data collection. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

### 2.2 | Experimental design and task

To examine neural correlates of the process of ToM reasoning in a sequential two-player task, we modified the three-stage turn-tasking game (Hedden & Zhang, 2002; Meijering et al., 2012) using an fMRI event-related design. In this game (see the example in Figure 1b), the white marble was dropped onto trapdoors that were controlled by one of the players: if the trapdoor was blue, the blue player controlled the trapdoor; if the trapdoor was yellow, the yellow player controlled it. Both players were to take turns to remove the color-coded trapdoors to control the path of the white marble towards a bin that maximized payoff. In this experiment, the color of the trapdoor was fixed across participants, which means Player 1 (participant in this game) always controlled the blue trapdoor at the first and third trapdoors/ decision points, and Player 2 always controlled the yellow trapdoor at the second decision point. The white marble was led either to another trapdoor or a bin containing payoffs. There were always four bins (labeled A, B, C, and D) and four possible payoffs for both players, ranking from worst to best, that is, ranging from 1 to 4. Each bin contained many blue and yellow marbles. The number of marbles of the player's own color determined their own payoff. The goal of this game for a player was to obtain as many marbles (points) as possible, irrespective of the points the other player earned. Therefore, to achieve the goal, both players had to remove the trapdoors (either "go" or "stop") sequentially to control the ending point of the white marble to result in a unique payoff for each player.

On each trial, the game began at bin A, and Player 1 had to decide whether to stop the white marble at bin A or let the white marble go to bin B. If Player 1 decided to go to bin B, it then became Player 2's turn. Similarly, Player 2 then decided whether to stop at bin B or go to bin C. If Player 2 decided to go to bin C, the turn passed back to Player 1 who had to make the final decision of whether to stop at bin C or go to bin D. In either case, the game ended after this final decision of Player 1. However, the game ended immediately if at any stage either player decided to stop in the current bin during their turn. Therefore, the outcome of each game could be any one of four bins based on the combination of actions of two players. It is worth noting that if players did not believe that their partner would adopt strategic thinking by taking the full path of the game into account, it would be no longer true that players' their own best response would reflect their strategic reasoning ability. To make sure participants' performance would not be influenced by their assumption of others' rationality, participants were informed that they were playing against a fully rational computer-simulated agent. They were told that such agent cared about their own payoffs and had a unique goal to achieve their possible highest payoffs while playing the game with them. This is also in line with the game theory which assumes that all players are rational and know that others are also rational. Accordingly, whether the opponent is a human being or a computer-simulated agent should not make difference because the opponent is assumed to always choose the optimal choice in game theory. Researchers have shown that ToM performance in this paradigm is not affected by whether the opponent is a virtual partner or a purported human opponent (Hedden & Zhang, 2002).

## 2.3 | Procedure

Since the more complex choice was at the first decision point at which player 1's decision depended upon what Player 2 might do if the game progresses to bin B, participants were always assigned to the role of Player 1. By doing this, we were able to discriminate the use of level of reasoning of the participants. In addition, to classify the use of the level of reasoning based on strategic mental models from the perspective of self and other, participants were also asked to predict what the other player might do at the second decision point. Participants were told to imagine that the white marble had rolled to the second decision point. Participants were informed that the other agent would play rationally, and the computer-simulated agent would also assume the participant was rational. Participants understood that both players' goals were to maximize their individual payoffs and the game was a non-cooperative game (e.g., the agent did not play with or against you).

In total, there were four unique types of trials based on the combination of the level of reasoning (low- vs. high-order ToM reasoning) and game type (decision for self vs. prediction for computer). Sixteen unique payoff structures were used for low- and high-order reasoning separately, resulting in 32 trials for each type of game (see the payoff

structures for more details). In total, there were 64 trials in this task. The presentation of the task is shown in Figure 1a. Trials started with a fixation intertrial interval (ITI) of 1.5–3.5 s (jittered), after which a four-bin game appeared. On Prediction trials, we presented the green word "Your prediction" on top of the screen with the stimuli. In these trials, participants were asked to record their prediction about the decision of the computer by choosing an option labeled "Go" or "Stop." On the contrary, on Decision trials, we presented the red word "Your decision" on top of the screen with the stimuli. When seeing this, participants were asked to choose what they should do at the first decision point. To give the participants enough time to think about their choices and conduct the experiment within the time constraint of the scanner, participants were told to give an answer within 20 s, upon which the stimuli disappeared without recording the response.

The position (left/right) of the two options such as "Go" and "Stop" was counterbalanced across trials. The order of trials was pseudo-random: four trials under the same type of the game were presented consecutively, followed by a long ITI (i.e., a fixation crosshair) for 14 s; after that, another four trials from the other game type were presented sequentially. We specifically used this design to make sure the participants would not be exhausted from switching decisions for different agents (i.e., self and computer) and could concentrate on the reasoning for the same agent. Every four trials consisted of two trials of low-order ToM reasoning and two high-order ToM reasoning, and the order of the level of reasoning (i.e., high and low) in every four trials was randomized across participants. In addition, the order of the two game types was counterbalanced across participants (the order was: prediction followed by decision or vice versa).

Participants were told that for each game in which they achieved their goal, they would be rewarded one point, and they would be paid according to the total number of points they accumulated (1 point = 50 cents). Thus, they had a monetary incentive to maximize the number of points. To familiarize the participants with the setup of the sequential two-player game before they entered the scanner, participants were first asked to make a decision for one trivial two-bin game that did not require ToM reasoning and 1 three-bin game that require first-order ToM reasoning. Later on, all participants were asked to familiarize the rules of the four-bin game and completed a short version of the task which included all four types of trials in equal quantities (i.e., one for each condition and all these four trials were not used for the formal study). This practice allowed the participants to play the game without learning from history.

## 2.4 | Payoff structures

The payoff structures of the current experiment were provided in Table S1. Notably, the configuration of payoffs determined the complexity of the reasoning required of the players (both Player 1 and Player 2). For example, Player 1 would immediately remove the left-side trapdoor if bin A contained his/her highest number of marbles. In

this case, player 1 would ignore the rest of the possible actions and did not have to reason about Player 2's beliefs. Therefore, such payoff structures were excluded since they did not involve any recursive reasoning. More importantly, to identify the neural basis related to high-order reasoning and low-order reasoning, two different payoff structures were selected so that the depth of ToM reasoning mastered by the participants could be derived from the manipulations of the game (for the manipulations of the payoff structures of high- and low-order ToM reasoning, please refer to the supplementary materials).

## 2.5 | Number of steps for strategic reasoning

To model the cognitive process of the strategy in reasoning about the mental states of others, we constructed the number of steps that an agent used to yield the optimal outcome based on previous papers (Bergwerff et al., 2014; Meijering et al., 2012). The current experiment provided complete and perfect information to both players. Namely, the optimal outcome was known at each decision point. Thus, one could apply forward and backward reasoning to determine the optimal outcome for each player when it was the player's turn to move (Bergwerff et al., 2014; Hedden & Zhang, 2002). According to our current design of the experiment, we proposed a way to measure the reasoning steps for four types of trials. If Player 1 made the decision based on ToM reasoning, Player 1 would reason about the mental content of Player 2 and realized that Player 2 had a goal of her/his own. Similar to Player 1, Player 2 would predict the behavior of Player 1 when making her/his decision. A high ToM agent would think about how other people could get to the highest possible payoff and how that would turn out to affect his/her highest possible payoff.

In this case, forward reasoning (find out which bin contained the highest possible payoff in a forward sequence) and backtracking (predicted the sequential move in a backward sequence) were required to determine whether the highest possible payoff was accessible to self and other. Specifically, in our design, the participant who employed this strategy would begin by taking the perspective of the opponent and find out the bin that contained the highest possible payoff for the opponent (forward reasoning), and then used backward reasoning to find out whether that particular bin was reachable by predicting the trapdoors the opponent would want to open, to finally reach a decision on whether to move the trapdoor at the first decision point. Following this strategy, we quantified the reasoning steps necessary for each trial with respect to the number of recursive thinking between two players. Identifying the neural correlates underlying steps of different levels of reasoning may help to explain the cognitive process engaged in interpersonal ToM reasoning. Table S2 shows the number of steps of forward reasoning plus backtracking strategy, computed by counting the number of times a value was attended. Following a previous study (Bergwerff et al., 2014), each time two values in different bins were compared and it was counted as two steps. This was because the participants had to attend both values. Counting details for the two trials of high- and low-order reasoning were provided in the supplementary materials.

## 2.6 | fMRI data acquisition

Imaging was performed on a Siemens PRISMA 3T MRI scanner (Siemens, Erlangen, Germany) with a 32-channel head coil. Functional images were collected using a 2-dimensional multiband single-shot T2*-weighted, echo-planar imaging (EPI) sequence (78 slices, 2 mm-thickness; TR = 1,000 ms; TE = 31.40 ms; flip angle = 60°; FOV = 220 mm; matrix size = 88 × 88; voxel size: $2.5 \times 2.5 \times 2.0$ mm$^3$). We used a multiband acceleration factor of six and scans were acquired in an oblique axial plane covering the whole brain. A high-resolution ($1.0 \times 1.0 \times 1.0$ mm$^3$) T1-weighted MPRAGE anatomical image (TR = 2,300 ms; TE = 2.22 ms; flip angle = 8; 1 mm-thickness) was acquired for registration and normalization of functional data to the standard brain. The total acquisition time for MPRAGE was 4 min and 52 s. The average duration of the functional scan time for each participant was 14 min. Head padding was used for each subject to minimize head motion. To dampen the scanner noise, headphones and earplugs were used.

## 2.7 | fMRI data preprocessing

SPM12 (www.fil.ion.ucl.ac.uk/spm/) was used for preprocessing and general linear model (GLM) analysis. The first eight volumes were discarded, and the remaining volumes were realigned to the first volume. The anatomical image was then co-registered to the mean functional image after motion correction and segmented, generating parameters for normalization. All volumes were subsequently normalized to the Montreal Neurological Institute (MNI) space and resampled to $2 \times 2 \times 2$ mm$^3$ isotropic voxel. The normalized functional images were then spatially smoothed with Gaussian kernel with 8 mm FWHM. A high-pass temporal filter with a cutoff of 128 s was applied to remove low-frequency drifts.

## 2.8 | fMRI data analysis

For each subject, a GLM was built (Friston et al., 1994). Only trials with correct response were included in the GLM analysis for each participant (number of subjects = 30, mean trial ± SD for Decision Low: 14 ± 2; Decision High: 10 ± 3; Prediction Low: 15 ± 1; Prediction Second: 13 ± 3, of 16 trials for each condition). At the first level, the individual data were analyzed by constructing sets of delta (stick) functions at the onset of the stimulus and at the time of the response (durations = 0). Because the current experiment was self-paced, the trial-by-trial variations of the reaction time served as the jittered inter-stimulus interval. Such a design allowed us to separate neural signals from the stimulus and response phases.

Each phase had four task-related regressors of interest: decision high, decision low, prediction high, and prediction low. This model allowed us to investigate: (a) the neural differences between decision high and decision low at both phases; (b) the neural differences between prediction high and prediction low at both phases. All

regressors were further modulated by the number of steps that were computed using the abovementioned procedures to allow us to investigate how the brain activity engaged in different levels of strategic thinking was modulated by the reasoning step. All these regressors were convolved with a canonical hemodynamic response function (HRF). Following this, the multicollinearity of the multiple regression factor matrix was tested. The variance inflation factor (VIF) for each regressor in the GLM model was low (all VIFs <2; VIF > 5 or 10 suggesting potential multicollinearity), indicating weak multicollinearity in the current design matrix (Mumford, Poline, & Poldrack, 2015). To regress out movement-related effects, the six scan-to-scan motion parameters were included as additional regressors. Parameter estimates from contrasts in single participant models were entered into the second level random-effect analysis to carry out a one-sample $t$ test on whether the activation during a contrast was significantly different from zero (Penny & Holmes, 2004). As a robustness check, we also built a model with durations = 2 s for each regressor and found qualitatively similar results to the above main model.

To further investigate the neural correlates of the heterogeneity observed in the ability of human strategic reasoning, we ran two whole-brain regression analyses using mean percentages of accuracies of participants' task performance in the high-order reasoning games as covariates associated with the corresponding contrasts that were estimated from the first-level analysis: (a) the neural response of decision high versus decision low with individual task performance in the decision high condition as a covariate and (b) the neural response of prediction high versus prediction low with individual task performance in the prediction high condition as a covariate.

To investigate functional connectivity patterns during ToM reasoning, we implemented generalized psychophysiological interaction (gPPI) analyses using a toolbox (https://www.nitrc.org/projects/gppi) (McLaren, Ries, Xu, & Johnson, 2012). Given that we found the relationship between individual high-order reasoning ability and the (decision high vs. decision low) effect at the left insula and left hippocampus (see the results), we investigated the task-dependent functional connectivity that the insula and hippocampus had with other regions using PPI (Friston et al., 1997; O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). The seed regions were created using a 10-mm diameter sphere in the left insula (peak MNI coordinate: −38, −6, −2) and left hippocampus (peak MNI coordinate: −18, −20, −10) as defined by the whole-brain regression analysis. In addition, to reveal the functional connectivity patterns in the prediction trials, another seed region was defined as a 10-mm sphere in the right caudate (peak coordinate MNI: 16, 18, 16) using the peak voxel from the corresponding significant cluster to the second-level regression effect with the accuracy in the prediction high condition.

For each first-level analysis, we first extracted the de-convolved time series within the seed region as the physiological regressor. We then define all regressors in the designed GLM as the psychological regressors. These two terms were multiplied as the PPI regressors. All these regressors were convolved with the canonical HRF to model the BOLD signal. In addition, six head-motion parameters were

included as covariates to account for residual motion effects. We further carried out whole-brain regression analyses between gPPI connectivity estimates and participants' task performance. The contrasts between the PPI regressors (decision high vs. decision low and prediction high vs. prediction low) for each participant were used separately at the second-level analysis with their corresponding behavioral index as a covariate.

To identify brain activity with considerable individual differences in complex strategic processing, we discussed results surviving after a voxel-level height threshold at $p < .005$ and cluster-level family-wise error (FWE) correction at $p < .05$. We also reported results at an uncorrected voxel-wise threshold of $p < .001$ with a cluster-wise threshold of $p < .05$ after FWE correction (See the tables for all brain activations). In addition, small volume correction (SVC) with peak FWE corrected $p$-values ($p < .05$) was used on a priori regions of interest previously implicated in mentalizing and interpersonal strategic thinking: ventral striatum (MNI, x, y, z = ± 10, 10, −6 mm), mPFC (x, y, z = ± 5, 60, 5 mm) (Yoshida et al., 2010), insula (x, y, z = ± 42, 0, 0 mm) (Bhatt & Camerer, 2005), dorsal anterior cingulate cortex (dACC) (x, y, z = ± 6, 30, 18 mm) (Apps, Green, & Ramnani, 2013), inferior frontal gyrus (x, y, z = ± 40, 10, 28 mm) (Van der Meer, Groenewold, Nolen, Pijnenborg, & Aleman, 2011), hippocampus (x, y, z = ± 22, −17, −15 mm) (Tavares et al., 2015), and left and right caudate defined by the corresponding automated anatomical labeling mask (Tzourio-Mazoyer et al., 2002). For each coordination-based ROI, a 10 mm radius was used. False discovery rate (FDR, $p < .05$ level) was applied for multiple corrections among ROIs where appropriate.

# 3 | RESULTS

## 3.1 | Behavioral results

Thirty-one subjects' data were included in the behavioral data analysis, although one participant's brain data were excluded in the fMRI analysis due to technical problems. We omitted no response trials for analyses (0.6% of all trials across all participants). A correct response in each trial was identified if the participant's choice matched the optimal decision in each trial (see Table S1). That is, to make a correct choice, participants should move the trapdoor correctly to obtain the highest possible marbles for themselves (decision for self) or others (prediction for computer) in each trial. Figure 2a depicts the mean percentage of accuracy of participants' task performance for the four conditions. To examine the mean accuracy scores in four conditions of the game, a two-way repeated measure ANOVA using the game type (decision/prediction) and the level of reasoning (low/high) as independent factors and the mean percentage of accuracy score as the dependent variable was performed. The analysis revealed a marginally significant interaction effect between the game type and the level of reasoning ($F_{[1,30]} = 3.270$, $p = .081$, $\eta_p^2 = 0.098$). Further, post hoc analysis showed that the mean percent accuracy scores in the condition of low-order reasoning (decision low, mean ± $SE$: 87.634 ± 1.650%) were significantly higher than the scores in the condition of
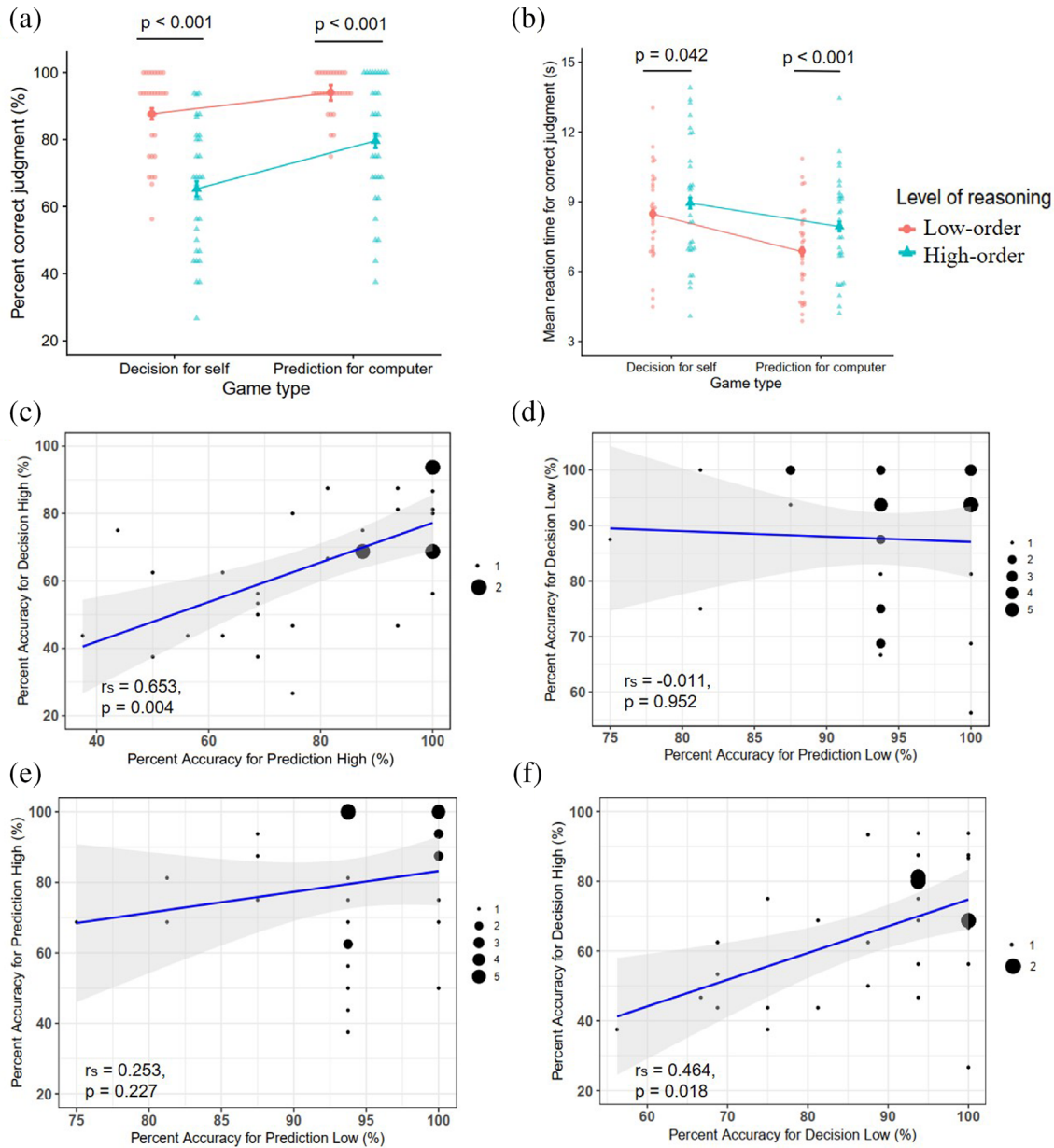
**FIGURE 2** Behavioral performance. (a) Mean percentage accuracy and (b) Mean reaction time as a function of game type and level of reasoning. Error bars represent the within-subjects *SE* (Morey, 2008). Each dot represents trial-averaged data per participant, for each condition. (c–f) Scatterplots for the relationship between mean percentage accuracies in four conditions (prediction high, decision high, prediction low, and decision low). The size of a black circle is proportional to the number of observations. The gray area represents the 95% confidence interval of the linear regression line

high-order reasoning (decision high: 65.282 ± 2.284%, $p < .001$) when making decisions for self at the first decision point. Similarly, the accuracy scores in the condition of low-order reasoning (prediction low: 93.952 ± 2.275%) was significantly higher than the scores in the condition of high-order reasoning (prediction high: 79.637 ± 2.145%, $p < .001$) when predicting the computer's choice at the second decision point. As expected, significant main effects of game type ($F_{[1,30]} = 34.937$, $p < .001$, $\eta_p^2 = 0.538$) and level of reasoning ($F_{[1,30]} = 63.508$, $p < .001$, $\eta_p^2 = 0.679$) were found. These results suggest that the depth of reasoning affects individuals' speculative

behavior, regardless of thinking from the perspectives of self or other. To assess possible learning effects, we compared the first half of the experiment with the second half and found no significant effects involving the session, $p$ values >.15, which suggests that human participants may not be able to increase their performance in ToM reasoning over repeated trials.

Additionally, correlation tests were conducted to examine the relationships between the accuracy scores of the four types of trials. The FDR ($p < .05$ level) was used to adjust for the number of associations we tested. First of all, we investigated whether one's ability to

predict the opponent's behavior was correlated with one's own ability in making optimal decisions in the same type of trials. The percent accuracy in the prediction high condition was positively correlated with the percent accuracy in the decision high condition across all participants (Figure 2c, Spearman correlation coefficient: $r_s$ = .653, FDR-corrected $p$ = .004, $n$ = 31), indicating that the mental model of the opponent's knowledge and strategy in the high-level ToM reasoning condition predicts subsequent self-related decision. However, the percent accuracy in the prediction low condition was not correlated with the percent accuracy in the decision low condition (Figure 2d, $r_s$ = −.011, FDR-corrected $p$ = .952, $n$ = 31). As the manipulation of the trial type of the prediction low requires the zeroth-order ToM which considers only one's own desires, beliefs, and goals, the performance in the Prediction Low condition may not be able to effectively facilitate self-reasoning on a higher level. People could predict the behavior of an opponent based only on the outcome of the opponent. As shown in Figure 2d, most of the participants' accuracy scores in the prediction low are larger than 85%, which may suggest that the reasoning in this condition is not as complicated as other conditions. We also found that the percent accuracy in the prediction low condition was not correlated with the percent accuracy in the prediction high condition (Figure 2e, $r_s$ = .253, FDR-corrected $p$ = .227, $n$ = 31), while the percent accuracy in the decision low condition was positively correlated with that in the decision high condition (Figure 2f, $r_s$ = .464, FDR-corrected $p$ = .018, $n$ = 31). These correlation results suggest that zeroth-order reasoning may tap into a distinct mechanism from other levels of reasoning. Higher-level ToM could not be predicted given their ability in zeroth-order ToM reasoning.

In addition to the mean percentage of accuracy score, a two-way repeated-measures ANOVA using the game type and the level of reasoning as independent factors and the mean reaction time (RT) of choosing the correct responses as the dependent variable was conducted. Our results showed a significant interaction effect between the game type and the level of reasoning ($F_{[1,30]}$ = 6.573, $p$ = .016, $\eta_p^2$ = 0.180, Figure 2b). Further, post hoc analysis showed that the mean reaction time in the decision low condition (8.475 ± 0.156) was significantly faster than the RT in the decision high condition (8.940 ± 0.216, $p$ = .042). Similarly, the RT in the prediction low condition (6.853 ± 0.169) was significantly faster than the RT in the prediction high condition (7.935 ± 0.192, $p$ < .001). In addition, significant main effects of game type ($F_{[1,30]}$ = 32.682, $p$ < .001, $\eta_p^2$ = 0.521) and level of reasoning ($F_{[1,30]}$ = 17.015, $p$ < .001, $\eta_p^2$ = 0.362) were found. These results suggest that the complexity of the task can affect the time needed for recursive thinking.

Furthermore, linear mixed-effect models using the LME4 package in R statistical software (version 3.6.1; R Core Team, 2019) with subject ID treated as a random intercept were carried out to examine whether the number of reasoning steps was predictive of the reaction time of the correct response in different types of trials. For each condition, the model included the number of steps as a fixed effect and reaction time as the dependent variable. The main effect of the reasoning step was found for each condition (decision high: β = .208, $SE$ = .091, $t$ = 2.294, $p$ = .023; decision low: β = .158, $SE$ = .069, $t$ = 2.306, $p$ = .022; prediction

high: β = .728, $SE$ = 0.156, $t$ = 4.656, $p$ < .001; prediction low: β = .559, $SE$ = 0.126, $t$ = 4.454, $p$ < .001). All four main models were compared with their own null model in which we only put the random effect of the participant. The Akaike information criterion (AIC; Akaike, 1974) was calculated for each model to see whether our main model for each condition was better (with lower AIC) than its null model. As expected, the AIC of all models with the number of steps was lower than the AIC of their null models. These results suggest that people were able to reason sufficiently deeply based on the internal reasoning strategy.

## 3.2 | Neuroimaging results

### 3.2.1 | Brain activation at the response phase

Figure 3a,b depict activation patterns at the response phase. Table 1 lists all significant peak activations related to the experimental factors. Compared with decision low, decision high at the decision phase showed greater activations in the left mPFC (Figure 3a), right STS, and left posterior cingulate cortex (PCC). In contrast, Prediction High showed greater activations in the left middle frontal gyrus (MFG), extending to the left mPFC (Figure 3b) when compared with prediction low. Another large cluster was found in the bilateral PCC, left TPJ, and left STS. We performed conjunction analysis between the decision high versus decision low and prediction high versus prediction low and found significant activation in the mPFC (peak MNI coordinate: x, y, z = 0, 60, 8, pFWE <.05, SVC, voxel size = 132).

When making decisions for self in the higher-order reasoning (decision high), the right dACC/dmPFC and right insula showed greater activations when the latent reasoning step of strategic thinking increased (Figure 4a, left panel). Greater activation was also found in the dACC/dmPFC when the reasoning step was increased in the decision low condition (Figure 4a, right panel). With regard to the neural patterns in reasoning steps in the prediction conditions, we found that increasing reasoning step in prediction high activated the right inferior frontal gyrus (IFG, Figure 4b, left panel), while increasing reasoning step in Prediction Low activated the bilateral MFG, bilateral ventromedial prefrontal cortex (vmPFC), left inferior temporal gyrus (ITG), left precuneus, and bilateral inferior parietal lobule (IPL, Figure 4b, right panel).

### 3.2.2 | Brain activation at the stimulus phase

At the stimulus phase, we examined the effects of the decision high versus decision low contrast and prediction high versus prediction low contrast. Table 1 summarizes the neural activity during the stimulus phase across participants. Notably, greater activity in the bilateral ventral striatum (VS) was found in both contrasts (decision high vs. decision low and prediction high vs. prediction low, see Figure 3c, d), while the reverse contrasts (decision low vs. decision high and prediction low vs. prediction high) yielded no significant effects. No suprathreshold clusters were found for the parametric modulation of
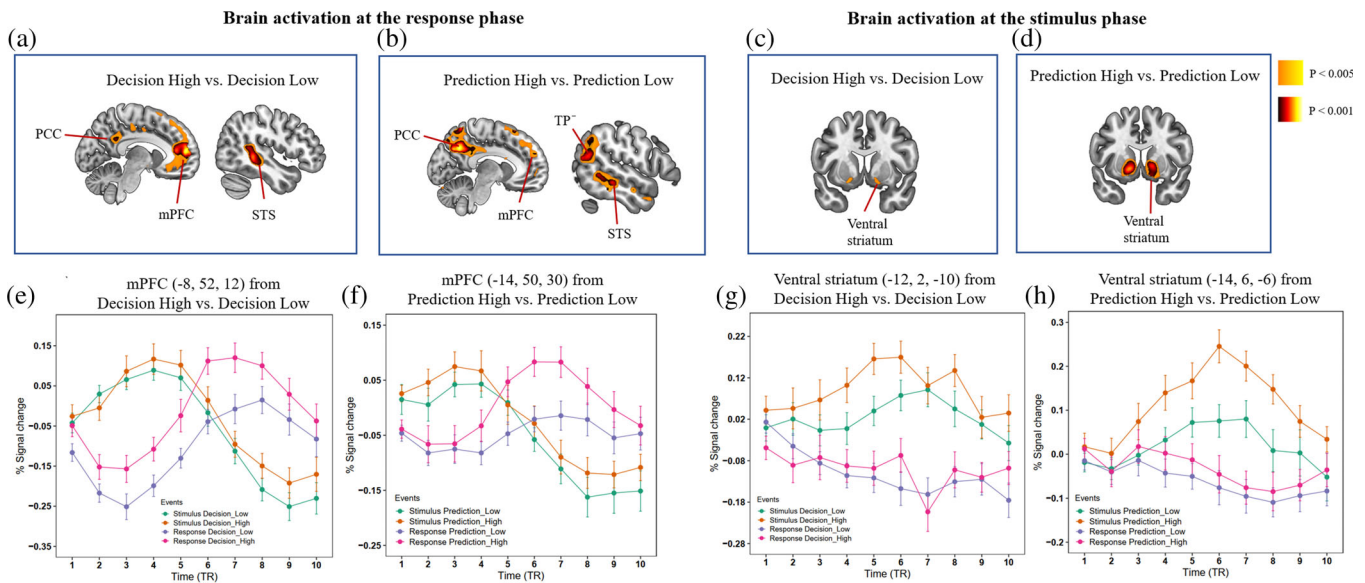
**FIGURE 3** Brain activation for the contrasts. (a) Comparison of the brain activity between decision high and decision low at the response phase. (b) Comparison of the brain activity between prediction high and prediction low at the response phase. (c) Comparison of the brain activity between decision high and decision low at the stimulus phase. (d) Comparison of the brain activity between prediction high and prediction low at the stimulus phase. The color bars represent statistical $t$-values. Results were displayed using an uncorrected voxel-wise threshold of $p < .005$ (warm color) and $p < .001$ (red-hot color) to show the full extent of the activations. (e–h) Finite impulse response (FIR) event-related time courses for the four brain areas of interest

reasoning steps at the stimulus phase. Besides, brain activity during the stimulus stage did not show any significant correlations with the participants' task performance. In addition, we created the mean time courses across all participants shown in Figure 3e–h. The time courses were related to the onset of the events and were derived from four regions of interest (ROIs) which were selected based on the results of the main GLM (peak MNI coordinate, mPFC: −8, 52, 12 from decision high vs. decision low at the response stage; mPFC: −14, 50, 30 from prediction high vs. prediction low at the response stage; VS: −12, 2, −10 from the decision high vs. decision low at the stimulus stage; VS: −14, 6, −6 from the prediction high vs. prediction low at the stimulus stage). These four ROIs were defined by 6-mm spheres with corresponding MNI coordinates as the center using Marsbar (Brett, Anton, Valabregue, & Poline, 2002). The mPFC significantly reacted for the decision high and prediction high conditions at the response stage while the ventral striatum built rapidly for the decision high and prediction high conditions at the stimulus stage. These time courses showed that activity patterns at the response stage were different from those at the stimulus stage, suggesting that the two stages were dissociated at the neural level.

### 3.2.3 | Neural correlates of individual differences at the response phase

To further investigate whether the neural activation reflected individual differences in the reasoning ability (i.e., mean percentage of behavioral accuracy) at the high-order strategic thinking, the whole-

brain regression analyses at the group level were performed. The results showed that individual differences in strategic reasoning ability were negatively correlated with the activity in the bilateral hippocampus, bilateral insula, midcingulate cortex (MCC), and occipital cortex in the contrast between decision high and decision low (Figure 5a). That is, participants who had better performance in high-order reasoning showed less activation in the bilateral hippocampus, bilateral insula, and MCC when deciding for self at the first decision point. In contrast, a positive correlation was found between individuals' strategic reasoning ability and activity in the right caudate when comparing prediction high with prediction low (Figure 5b), suggesting that participants who had better performance in predicting other's strategic behaviors showed greater activation in the right caudate. Table 2 lists all significant peak activations for the whole-brain regression analyses.

The individual differences in the strategic reasoning ability were also found in the functional connectivity between the seed regions of the left hippocampus (peak MNI coordinate: −18, −20, −10), left insula (peak MNI coordinate: −38, −6, −2) and right caudate (peak MNI coordinate: 16, 18, 16) and other brain areas during decision-making. With increases in individuals' accuracy of strategic decision, the left insula showed more positive functional connectivity with the left dACC/dmPFC, right superior frontal gyrus (SFG), and right thalamus in the contrast between decision high and decision low (Figure 5c, left panel). The results suggest that participants who had better performance in high-order reasoning showed stronger levels of coupling between the left insula and left dACC/dmPFC when making decisions in the high-order reasoning condition compared with the low-order reasoning condition. In addition, stronger functional

**TABLE 1** Brain activations in the general linear model (GLM) analysis of the task

| Contrast | Region | BA | R/L/M | MNI coordinates x | y | z | t score | Voxels |
|---|---|---|---|---|---|---|---|---|
| **Brain activation at the stimulus phase** | | | | | | | | |
| Decision high versus decision low | Ventral striatum[a] | – | R | 16 | 10 | −14 | 3.32 | 23 |
| | Ventral striatum[a] | – | L | −12 | 2 | −10 | 3.83 | 28 |
| | | | | | | | | |
| Prediction high versus prediction low | Ventral striatum[a] | – | R | 12 | 2 | −2 | 4.68 | 261 |
| | Ventral striatum[b] | – | L | −14 | 6 | −6 | 5.20 | 1,140 |
| | | | | | | | | |
| High versus low | Fusiform gyrus[b] | 37 | R | 34 | −40 | −16 | 6.03 | 1,945 |
| | Occipital cortex | 18 | R | 10 | −78 | −2 | 4.06 | |
| | Ventral striatum[b] | – | L | −18 | 8 | −10 | 5.87 | 1,442 |
| | Ventral striatum[b] | – | R | 12 | 4 | −8 | 5.13 | |
| | | | | | | | | |
| No suprathreshold clusters were found for the above reverse contrasts | | | | | | | | |
| **Brain activation at the decision phase** | | | | | | | | |
| Decision high versus decision low | Medial prefrontal cortex[b] | 10 | L | −8 | 52 | 12 | 5.64 | 4,056 |
| | Superior temporal sulcus[b] | 21 | R | 46 | −38 | 0 | 4.86 | 1,427 |
| | Posterior cingulate cortex | 23 | L | −6 | −48 | 28 | 4.06 | 1,142 |
| | | | | | | | | |
| Prediction high versus prediction low | Posterior cingulate cortex[b] | 31 | L | −10 | −58 | 36 | 4.87 | 8,040 |
| | Posterior cingulate cortex[b] | 31 | R | 10 | −54 | 32 | 4.80 | |
| | Temporoparietal junction[b] | 39 | L | −48 | −60 | 18 | 4.34 | |
| | Occipital cortex | – | R | 34 | −76 | 2 | 4.55 | 2,430 |
| | Cerebellum[b] | – | R | 32 | −78 | −26 | 4.35 | |
| | Superior temporal sulcus[b] | 21 | L | −58 | −32 | −10 | 4.49 | 1,156 |
| | Middle frontal gyrus[b] | 6 | L | −38 | 8 | 54 | 4.37 | 2,274 |
| | Medial prefrontal cortex | 9 | L | −14 | 50 | 30 | 3.89 | |
| | | | | | | | | |
| High versus low | Precuneus[b] | 7 | R | 8 | −68 | 34 | 5.52 | 22,914 |
| | Posterior cingulate cortex[b] | 23 | L | −6 | −22 | 32 | 5.49 | |
| | Temporoparietal junction[b] | 39 | L | −54 | −64 | 18 | 5.46 | |
| | Medial prefrontal cortex[b] | 10 | L | −4 | 52 | 20 | 4.77 | |
| | Cerebellum | – | L | −44 | −62 | −40 | 5.16 | |
| | Middle temporal gyrus[b] | 21 | L | −62 | −34 | −14 | 4.74 | |
| | Cerebellum[b] | – | R | 36 | −78 | −30 | 4.90 | 5,814 |
| | Superior temporal sulcus[b] | 22 | R | 66 | −22 | −4 | 4.83 | |
| | | | | | | | | |
| No suprathreshold clusters were found for the above reverse contrasts | | | | | | | | |
| | | | | | | | | |
| Increasing step for decision high | Dorsal anterior cingulate cortex/dorsomedial prefrontal cortex[a] | 32 | R | 14 | 30 | 16 | 3.24 | 63 |
| | Insula[a] | – | R | 38 | 2 | −8 | 3.32 | 14 |

**TABLE 1** (Continued)

| Contrast | Region | BA | R/L/M | MNI coordinates x | y | z | t score | Voxels |
|---|---|---|---|---|---|---|---|---|
| Decreasing step for decision high | Posterior cingulate cortex | 31 | L | −6 | −66 | 36 | 4.33 | 1,804 |
| | Posterior cingulate cortex | 23 | R | 10 | −56 | 26 | 3.45 | |
| | | | | | | | | |
| Increasing step for decision low | Dorsal anterior cingulate cortex/dorsomedial prefrontal cortex[a] | 32 | M | 0 | 30 | 22 | 3.12 | 50 |
| | | | | | | | | |
| Increasing step for prediction high | Inferior frontal gyrus[a] | 44 | R | 34 | 10 | 24 | 3.49 | 79 |
| | | | | | | | | |
| Increasing step for prediction low | Middle frontal gyrus[b] | 8 | L | −34 | 24 | 44 | 6.39 | 4,555 |
| | Ventromedial prefrontal cortex | 24 | L | −8 | 38 | 2 | 3.90 | |
| | Ventromedial prefrontal cortex | 10 | R | 8 | 52 | −2 | 3.27 | |
| | Inferior temporal gyrus | 20 | L | −56 | −24 | −24 | 5.81 | 1,399 |
| | Inferior parietal lobule[b] | 39 | L | −52 | −54 | 38 | 5.30 | 3,173 |
| | Precuneus | 7 | L | −6 | −62 | 62 | 3.54 | |
| | Middle frontal gyrus | 9 | R | 34 | 22 | 34 | 4.71 | 1,353 |
| | Cerebellum[b] | – | R | 34 | −68 | −46 | 4.69 | 1,759 |
| | Inferior parietal lobule | 7 | R | 40 | −46 | 58 | 4.19 | 1,210 |

*Note:* Results were reported surviving after a voxel-level height threshold at $p < .005$ and cluster-level family-wise error (FWE) correction at $p < .05$. Coordinates reported were based on the Montreal Neurological Institute (MNI) coordinate system.

Abbreviations: BA, Brodmann area; L, left; M, middle, R, right.

[a]Indicates significance after small volume correction with peak FWE corrected $p$-values ($p < .05$). False discovery rate (FDR, $p < .05$ level) was applied for multiple corrections among ROIs where appropriate.

[b]Indicates results survived at an uncorrected voxel-wise threshold of $p < .001$ with a cluster-wise threshold of $p < .05$ after FWE correction.



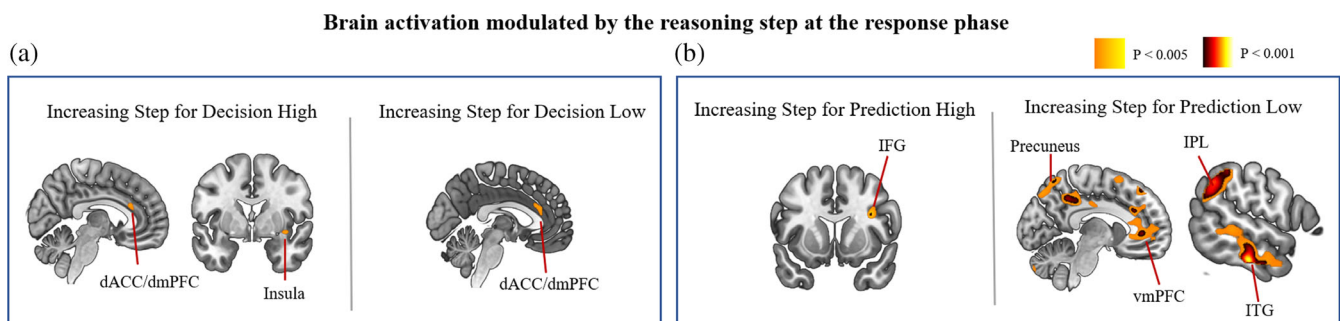**Brain activation modulated by the reasoning step at the response phase**

**FIGURE 4** Brain activation modulated by the reasoning step at the response phase. (a) Left panel: Parametric modulation of increasing number of reasoning steps of the high-order ToM reasoning during decision for self; Right panel: Parametric modulation of increasing number of reasoning steps of the low-order ToM reasoning during decision for self. (b) Left panel: Parametric modulation of increasing number of reasoning steps of the high-order ToM reasoning during prediction for the other; Right panel: Parametric modulation of increasing number of reasoning steps of the low-order ToM reasoning during prediction for the other. The color bars represent statistical $t$-values. Results were displayed using an uncorrected voxel-wise threshold of $p < .005$ (warm color) and $p < .001$ (red-hot color) to show the full extent of the activations

connectivity was also found between the left hippocampus and the bilateral dACC/dmPFC when contrasting between decision high and decision low as individuals' high-order reasoning ability increased (Figure 5c, right panel). When contrasting prediction high with

prediction low, positive functional connectivity was seen between right caudate and left insula as the reasoning ability to predict other's strategic behaviors increased (Figure 5d). Table 3 lists all significant peak activations for the PPI second-level analyses.
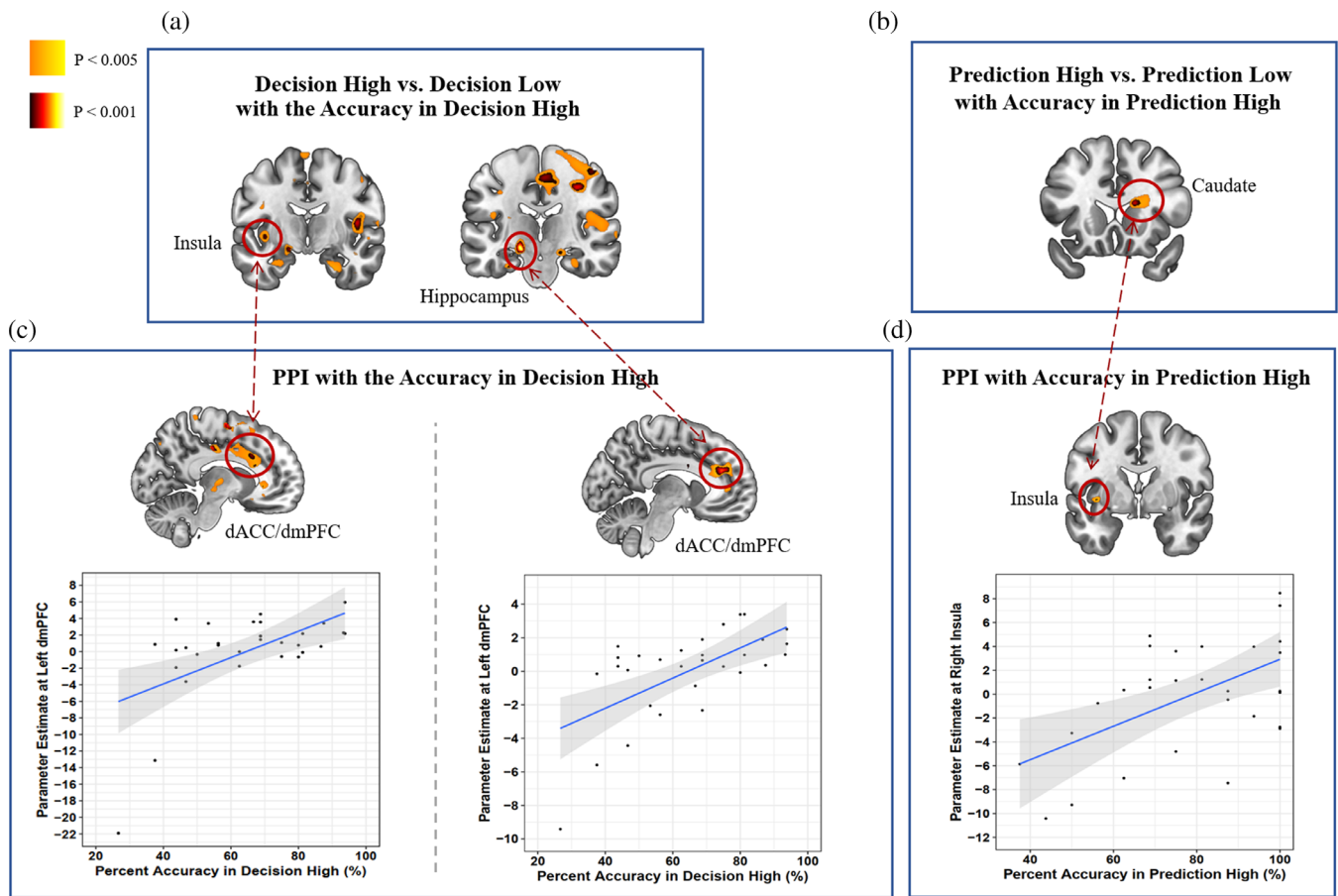
**FIGURE 5** Relationship between brain function and individual differences in high-order reasoning at the response phase. (a) Activity in the bilateral insula and the left hippocampus negatively associated with individuals' high-order reasoning ability when comparing decision high with decision low. (b) Activity in the right caudate positively associated with individuals' high-order reasoning ability when comparing prediction high with prediction low. (c) Top left panel: In the displayed image, the seed region was the left insula (peak MNI coordinate: −38, −6, −2). The positive functional connectivity between the left insula and dACC/dmPFC (peak MNI coordinate: −6, −2, 36) in the contrast of decision high versus decision low correlated with the individuals' high-order reasoning ability. Top right panel: The positive functional connectivity between the left hippocampus (seed region; peak MNI coordinate: −18, −20, −10) and dACC/dmPFC (peak MNI coordinate: −8, 36, 24) in the contrast of decision high versus decision low correlated with the individuals' high-order reasoning ability. (d) The positive functional connectivity between the right caudate (seed region; peak MNI coordinate: 16, 18, 16) and insula (peak MNI coordinate: −38, 0, −2) in the contrast of prediction high versus prediction low correlated with the individuals' high-order reasoning ability. The scatterplots below each panel depict the correlations between parameter estimates at the activated brain regions and individual differences in high-order reasoning. The gray shaded area represents the 95% confidence interval of the linear regression line. The color bars represent statistical $t$ values. Results were displayed using an uncorrected voxel-wise threshold of $p < .005$ (warm color) and $p < .001$ (red-hot color) to show the full extent of the activations

## 4 | DISCUSSION

In the current study, we used a sequential-move interactional game to manipulate and detect the use of level of reasoning in both first-person and third-person perspectives. Our behavioral data showed that high-order ToM reasoning indeed produced longer reaction time and lower accuracy in comparison with low-order ToM reasoning, irrespective of making the decision for self or making the prediction for others. In the decision task, our neural data revealed that high-order reasoning for oneself at the first decision point activated the ToM network, including mPFC, PCC, and STS (Amodio & Frith, 2006; Fletcher et al., 1995; Frith & Frith, 2006; Gallagher et al., 2000; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004; Schurz et al., 2014; Spunt &

Adolphs, 2014). Importantly, performance in high-order condition was negatively predicted by activity in the insula, hippocampus, and MCC. Furthermore, individuals with better performance in the high-order condition also showed stronger coupling between dmPFC/dACC and insula as well as coupling between dmPFC/dACC and hippocampus. In the prediction task, performance in high-order condition was positively correlated with activity in caudate and the functional connectivity between caudate and insula.

The difference between high-order and low-order conditions is that the former requires second-order ToM reasoning to obtain the highest possible payoff. We showed that the high-order condition is more demanding in the mental process of solving the game, as evidenced by the longer reaction time and lower accuracy.

**TABLE 2** Brain regions associated with individual differences in high-order reasoning at the decision phase

| Contrast | Region | BA | R/L/M | MNI coordinates | | | t score | Voxels |
|---|---|---|---|---|---|---|---|---|
| | | | | x | y | z | | |
| At the decision phase | | | | | | | | |
| **(Decision high vs. decision low) with accuracy of decision high as a covariate** | | | | | | | | |
| Positive correlation | No suprathreshold clusters were found | | | | | | | |
| | | | | | | | | |
| Negative correlation | Hippocampus[a] | – | L | −18 | −20 | −10 | 5.51 | 220 |
| | Hippocampus[a] | – | R | 18 | −20 | −14 | 3.63 | 72 |
| | Insula[a] | – | L | −38 | −6 | −2 | 3.88 | 63 |
| | Insula[a] | – | R | 44 | −6 | 6 | 3.56 | 24 |
| | Occipital cortex | 31 | R | 10 | −70 | 26 | 4.58 | 3,368 |
| | Occipital cortex[b] | 17 | L | −10 | −74 | 6 | 4.53 | |
| | Midcingulate cortex | 24 | R | 14 | −16 | 44 | 4.35 | 2,045 |
| | | | | | | | | |
| **(Prediction high vs. prediction low) with accuracy of prediction high as a covariate** | | | | | | | | |
| Positive correlation | Caudate[a] | – | R | 16 | 18 | 16 | 3.88 | 43 |
| | | | | | | | | |
| Negative correlation | No suprathreshold clusters were found | | | | | | | |

*Note:* The results were based on whole-brain regression analysis (Voxel-level threshold $p < .005$ uncorrected, cluster-level $p < .05$ FWE correction). Coordinates reported were based on the Montreal Neurological Institute (MNI) coordinate system.
Abbreviations: BA, Brodmann area; L, left; M, middle, R, right.
[a]Indicates significance after small volume correction with peak FWE corrected *p*-values ($p < .05$). False discovery rate (FDR, $p < .05$ level) was applied for multiple corrections among ROIs where appropriate.
[b]Indicates results survived at an uncorrected voxel-wise threshold of $p < .001$ with a cluster-wise threshold of $p < .05$ after FWE correction.

Interestingly, the accuracies in decision-making for oneself in the high-order and low-order conditions were correlated with each other but no such correlation was found for accuracies in making predictions about others. In the high-order condition, performances in making decisions and predictions were correlated. These findings suggest that both high-order thinking processes of decision and prediction may share common cognitive capacities. Our neuroimaging results support this argument by showing that the mPFC was involved in encoding the processes of high-order strategic thinking in both first-person and third-person perspectives.

At the stimulus stage, we found that the ventral striatum was more activated in the high-order ToM condition. The striatum, together with other dopamine-rich brain regions, are involved in representing reward prediction errors, the mismatch between expected rewards and actual outcomes (Fiorillo, Tobler, & Schultz, 2003; Schultz, 1998; Tobler, Fiorillo, & Schultz, 2005). In the reward prediction error framework, individuals infer the underlying reward structures by continually using these errors to update their predictions. Recent Bayesian theories of mind propose that individuals predict others' actions, and reconstruct others' belief state and reward function by using Bayesian inference that is conditioned from past observations of others' behaviors (Baker, Saxe, & Tenenbaum, 2011). These Bayesian learning models for ToM point out that mentalizing and reward-based learning may share similar Bayesian inference mechanisms (Devaine et al., 2014). Essentially, the neural basis of ToM can

be investigated as a neural prediction problem (Collette, Pauli, Bossaerts, & O'Doherty, 2017; Koster-Hale & Saxe, 2013). The striatum, with its special involvement in goal-directed learning and prediction error representation, is in a preferential position to be a candidate neural substrate of the "Bayesian ToM" and may encode the difference between the observable behaviors and invisible mental predictions. Interestingly, we only found strong striatal activity when stimuli were shown but not when participants made decisions. Thus, the specific functional significance of striatum in mentalizing has to be investigated further.

Both decision-making and prediction in the high ToM condition activated the mentalizing network, including mPFC, PCC, and STS (Rilling, King-Casas, & Sanfey, 2008; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004; Spunt & Adolphs, 2014; Spunt & Lieberman, 2012). These findings confirm that the mentalizing related brain regions are recruited when individuals think recursively in strategic games. Meta-analysis has shown that these regions are involved in mentalizing across tasks (Molenberghs, Johnson, Henry, & Mattingley, 2016; Schurz et al., 2014). Spunt and his colleagues further showed that the mentalizing system, including mPFC, TPJ, STS, and PCC, was sensitive to the high demand of explaining observed actions (Spunt & Adolphs, 2014; Spunt & Lieberman, 2012; Spunt, Satpute, & Lieberman, 2011), supporting our findings that the mentalizing network was involved in explaining and predicting the actions of self and others in complex situations. A majority of previous

**TABLE 3** Functional connectivity between regions as a function of individual differences in high-order reasoning at the decision phase

| | | | | | MNI coordinates | | | | |
| | | | | R/ | | | | t | |
| Seed | Contrast | Region | BA | L/M | x | y | z | score | Voxels |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| At the decision phase | | | | | | | | | |
| Functional connectivity of the decision high versus decision low contrast with accuracy of decision high as a covariate | | | | | | | | | |
| Insula | Positive | Thalamus | – | R | 14 | −8 | 2 | 4.99 | 2,136 |
| (−38, −6, −2) | | Superior frontal gyrus | – | R | 10 | 6 | 64 | 3.81 | 1,064 |
| | | Dorsal anterior cingulate cortex/dorsomedial prefrontal cortex[a] | 32 | L | −10 | 22 | 20 | 3.59 | 35 |
| | | | | | | | | | |
| | Negative | No suprathreshold clusters were found | | | | | | | |
| | | | | | | | | | |
| Hippocampus | Positive | Dorsal anterior cingulate cortex/dorsomedial prefrontal cortex[a] | 32 | L | −8 | 36 | 24 | 4.21 | 194 |
| (−18, −20, −10) | | Dorsal anterior cingulate cortex/dorsomedial prefrontal cortex[a] | 32 | R | 2 | 32 | 24 | 4.04 | 127 |
| | | | | | | | | | |
| | Negative | No suprathreshold clusters were found | | | | | | | |
| Functional connectivity of the prediction high vs. prediction low contrast with accuracy of prediction high as a covariate | | | | | | | | | |
| Caudate | Positive | Insula[a] | – | L | −38 | 0 | −2 | 3.53 | 25 |
| (16, 18, 16) | | | | | | | | | |
| | Negative | No suprathreshold clusters were found | | | | | | | |

*Note:* The results were based on whole-brain regression analysis (Voxel-level threshold $p < .005$ uncorrected, cluster-level $p < .05$ FWE correction). Coordinates reported were based on the Montreal Neurological Institute (MNI) coordinate system.

Abbreviations: BA, Brodmann area; L, left; M, middle; R, right.

[a]Indicates $p < .05$ FWE after small volume correction. False discovery rate (FDR, $p < .05$ level) was applied for multiple corrections among ROIs where appropriate.

fMRI studies have mainly compared a condition with ToM demand to a control condition that does not require the involvement of ToM. This manipulation allowed researchers to identify ToM related neural activities but did not permit the investigation of specific brain regions during high-order mentalizing processes. In the current study, our fMRI findings revealed that high-order reasoning, when compared with low-order reasoning also activated the mentalizing network, suggesting that the degree of mentalizing is encoded in the same mentalizing regions. Interestingly, our results showed the mPFC activated for predicting others was more dorsal than the mPFC recruited for making one's own choices. This resonates with previous findings showing that dorsal mPFC is more engaged for mentalizing about others' beliefs whereas the ventral mPFC is more specific to self-related processing (Hu et al., 2016; Sui & Humphreys, 2015). It has also been proposed that mPFC processes the motivational components of ToM whereas the TPJ is more engaged in the cognitive components of ToM, for example, representing other's beliefs (Koster-Hale et al., 2017; Molenberghs et al., 2016; Schurz et al., 2014). Our study supports this argument by showing that the goal-directed mentalizing process (decision high vs. decision low) mainly activated the mPFC.

In the decision task, we found that the number of reasoning steps was correlated with activity in dmPFC/dACC. In the prediction task, the number of reasoning steps was associated with activity in IFG, vmPFC, IPL, and ITG. In general, these regions are often involved in high demanding cognitive tasks and may represent the complexity of calculation and planning (Grotheer, Jeska, & Grill-Spector, 2018; Gruber, Indefrey, Steinmetz, & Kleinschmidt, 2001; Hubbard, Piazza, Pinel, & Dehaene, 2005; Kahnt, Heinzle, Park, & Haynes, 2011; Piazza, Pinel, Le Bihan, & Dehaene, 2007; Shenhav, Cohen, & Botvinick, 2016). Here the number of steps was calculated based on our model of forward reasoning plus backward tracking, and may or may not represent the actual thinking steps taken by all participants. Nevertheless, the calculated number of reasoning steps provides a proxy to indicate the complexity and cognitive load of strategic thinking. The parametric correlation with activities in regions associated with cognitive control and calculation suggests that a sophisticated level of mentalizing is represented in the same brain regions as those commonly engaged in demanding cognitive tasks.

When comparing the high-order ToM condition with the low-order ToM condition in the decision task, our study showed that individual differences in high-order social reasoning were negatively

correlated with activity in the insula and hippocampus. Moreover, the functional connectivity between the insula/hippocampus and dACC/dmPFC predicted the performance in high-order ToM condition. These findings suggest that individuals with better high-order reasoning activated the insula/hippocampus less but showed stronger coupling between the insula/hippocampus and dACC/dmPFC. One possibility is that the insula/hippocampus in these high-performing individuals may function more efficiently in the reasoning process. The insula was shown to be more active in pure coordination games which rely on intuition than in dominance-solvable games which depend on deliberation (Kuo, Sjöström, Chen, Wang, & Huang, 2009). Our findings may suggest that individuals who are better at high-order reasoning may be able to recruit insula more efficiently when engaging in recursive thinking. The hippocampus is known to contribute to relational encoding which involves flexible comparison and association between distinct items (Greene, Gross, Elsinger, & Rao, 2006). Moreover, the hippocampus is shown to subserve the unconscious relational inference, which facilitates the combination of past experience to update knowledge (Reber, Luechinger, Boesiger, & Henke, 2012). In our task, the participant must encode and hold in mind the results of reasoning for each step. The hippocampus may be involved in reasoning decisions that depend on the memory of steps. Our research points out that the hippocampus may have a role in deliberation and reasoning in strategic decisions. In the prediction task, we found that the activity in caudate and the connectivity between caudate and insula predicted the accuracy of prediction in high-order ToM trials. The caudate is involved in the quick generation of the best next move (Wan et al., 2011; Wan et al., 2012). Recent human meta-analysis indicates that the caudate is functionally connected with insula and plays a primary role in integrating information from the associative regions (Robinson et al., 2012). The increased caudate-insula coupling may help people integrate the prediction of others' movement with self-strategic action in the high-order ToM reasoning.

There are several worth mentioning limitations in the current study. First, participants only needed to reason pure rational opponents' next move. It is unclear whether our findings can be extended to uncertain social interactions in which opponents' choices are probabilistic. Interpersonal uncertainty, a common character in real-life social interactions, was minimized in our study (Yoshida et al., 2010). Another related concern is that our task may simply tap onto the general reasoning process rather than strategic thinking in social contexts. A meta-analysis of neuroimaging studies on deductive reasoning showed that general reasoning activates the lateralized brain system, for example, posterior parietal cortex for relational arguments, IFG for categorical arguments, and precentral gyrus for propositional arguments (Prado, Chadha, & Booth, 2011). Nevertheless, future studies may directly compare mentalizing steps with cognitive reasoning steps to further examine the specificity of strategic thinking in the brain. Second, performance in this task is associated with monetary rewards, which may raise the question of whether these brain activity patterns merely encode rewarding experiences associated with reasoning. Although the intrinsic joy of reasoning cannot be directly measured, our design does allow us to dissociate monetary reward with

reasoning levels by orthogonalizing the two parameters. The payoffs in the high and low reasoning conditions were comparable, suggesting that our fMRI findings cannot be explained by the magnitude of rewards. Finally, it is worth noting that our model of thinking strategies is an idealization, as participants may use mixed strategies. In principle, a player can use different strategies to find the optimal solution and many possible strategies could yield the same behavioral patterns. We had defined this as an inverse problem where the strategy is to infer the thinking processes based on choices to be made. Moreover, failure to use high-order reasoning may be due to the inability to understand another's intentions of a certain order or the lack of motivation to apply strategic thinking of an appropriate order that is demanding. Future computational work may try to capture the full complexity of thinking processes (Toelch & Dolan, 2015).

To sum up, our research showed that the mPFC-subcortical network is involved in high-order strategic reasoning, in both the processes of decision-making for self and predicting others' choices. One tentative neural model of strategic reasoning, although still awaiting further validation, is that the striatum represents Bayesian-like inferences at the stimulus stage while the mPFC, together with the insula/hippocampus, encodes the depth of reasoning and associated relational inference when individuals think about the next move. Our work highlights the key role of the interplay between mPFC and subcortical regions in advanced social decision-making.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

*Shanshan Zhen* https://orcid.org/0000-0002-0732-9429
*Rongjun Yu* https://orcid.org/0000-0003-0123-1524

## REFERENCES

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2013). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Decision*, *1*, 8–23.

Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277.

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.

Astington, J. W., & Hughes, C. (2013). Theory of mind: Self-reflection and social understanding. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology* (Vol. 2, pp. 398–424). Oxford, Englannd: Oxford University Press.

Aumann, R. J. (1995). Backward induction and common knowledge of rationality. *Games and Economic Behavior*, *8*(1), 6–19.

Báez-Mendoza, R., & Schultz, W. (2013). The role of the striatum in social behavior. *Frontiers in Neuroscience*, *7*, 233.

Baker, C. L., Saxe, R. R., & Tenenbaum, J. B. (2011). Bayesian theory of mind: Modeling joint belief-desire attribution. *Proceedings of the thirty-second annual conference of the cognitive science society*, 2469–2474.

Baker, C. L., & Tenenbaum, J. B. (2014). Modeling human plan recognition using Bayesian theory of mind. In *Plan, Activity, and Intent Recognition: Theory and Practice* (pp. 177–204). New York, NY: Elsevier.

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature*, 456(7219), 245–249.

Bergwerff, G., Meijering, B., Szymanik, J., Verbrugge, R., & Wierda, S. (2014). Computational and algorithmic models of strategies in turn-based games. *Proceedings of the 36th annual conference of the cognitive science society* (p. 1778–1783).

Bhatt, M., & Camerer, C. F. (2005). Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games and Economic Behavior*, 52(2), 424–459.

Bhatt, M., & Camerer, C. F. (2011). The cognitive neuroscience of strategic thinking. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford handbook of social neuroscience*. Oxford, England: Oxford University Press.

Brandenburger, A., & Li, X. (2015). *Thinking about thinking and its cognitive limits*. Working paper, New York University.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using an SPM toolbox. *In 8th international conference on functional mapping of the human brain*, 16(2), 497.

Camerer, C. F. (1991). Does strategy research need game theory? *Strategic Management Journal*, 12(S2), 137–152.

Camerer, C. F. (2009). Behavioral game theory and the neural basis of strategic choice. In *Neuroeconomics* (pp. 193–206). London, England: Academic Press.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice*, 11(2), 73–92.

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping*, 30(8), 2313–2335.

Collette, S., Pauli, W. M., Bossaerts, P., & O'Doherty, J. (2017). Neural computations underlying inverse reinforcement learning in the human brain. *eLife*, 6, e29718.

Conway, J. R., Catmur, C., & Bird, G. (2019). Understanding individual differences in theory of mind via representation of minds, not mental states. *Psychonomic Bulletin & Review*, 26(3), 798–812.

Coricelli, G., & Nagel, R. (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, 106(23), 9163–9168.

Daniel, R., & Pollmann, S. (2014). A universal role of the ventral striatum in reward-based learning: Evidence from human studies. *Neurobiology of Learning and Memory*, 114, 90–100.

Devaine, M., Hollard, G., & Daunizeau, J. (2014). The social Bayesian brain: Does mentalizing make a difference when we learn? *PLoS Computational Biology*, 10(12), e1003992. https://doi.org/10.1371/journal.pcbi.1003992

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, 55(2), 705–712.

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, 18(9), 547–558.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614), 1898–1902.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of" theory of mind" in story comprehension. *Cognition*, 57(2), 109–128.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, 318(5854), 1305–1308.

Friston, K., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6(3), 218–229.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189–210.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4), 531–534.

Fu, G., Xiao, W. S., Killen, M., & Lee, K. (2014). Moral judgment and its relation to second-order theory of mind. *Developmental Psychology*, 50(8), 2085–2092. https://doi.org/10.1037/a0037077

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, 7(2), 77–83.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind'in verbal and nonverbal tasks. *Neuropsychologia*, 38(1), 11–21.

Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2006). An FMRI analysis of the human hippocampus: Inference, context, and task awareness. *Journal of Cognitive Neuroscience*, 18(7), 1156–1173. https://doi.org/10.1162/jocn.2006.18.7.1156

Grotheer, M., Jeska, B., & Grill-Spector, K. (2018). A preference for mathematical processing outweighs the selectivity for Arabic numbers in the inferior temporal gyrus. *NeuroImage*, 175, 188–200.

Gruber, O., Indefrey, P., Steinmetz, H., & Kleinschmidt, A. (2001). Dissociating neural correlates of cognitive components in mental calculation. *Cerebral Cortex*, 11(4), 350–359.

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*, 105(18), 6741–6746.

Hedden, T., & Zhang, J. (2002). What do you think I think you think?: Strategic reasoning in matrix games. *Cognition*, 85(1), 1–36.

Hillebrandt, H., Dumontheil, I., Blakemore, S.-J., & Roiser, J. P. (2013). Dynamic causal modelling of effective connectivity during perspective taking in a communicative task. *NeuroImage*, 76, 116–124.

Hu, C., Di, X., Eickhoff, S. B., Zhang, M., Peng, K., Guo, H., & Sui, J. (2016). Distinct and common aspects of physical and psychological self-representation in the brain: A meta-analysis of self-bias in facial and self-referential judgements. *Neuroscience & Biobehavioral Reviews*, 61, 197–207.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, 6(6), 435–448.

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. (2017). Social learning through prediction error in the brain. *NPJ Science of Learning*, 2(1), 1–9.

Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *NeuroImage*, 56(2), 709–715.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41.

Khalvati, K., Park, S. A., Mirbagheri, S., Philippe, R., Sestito, M., Dreher, J.-C., & Rao, R. P. (2019). Modeling other minds: Bayesian inference explains human choices in group decision-making. *Science Advances*, 5(11), eaax8783.

Koster-Hale, J., Richardson, H., Velez, N., Asaba, M., Young, L., & Saxe, R. (2017). Mentalizing regions represent distributed, continuous, and abstract dimensions of others' beliefs. *NeuroImage*, 161, 9–18.

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

Kuo, W.-J., Sjöström, T., Chen, Y.-P., Wang, Y.-H., & Huang, C.-Y. (2009). Intuition and deliberation: Two systems for strategizing in the brain. *Science*, 324(5926), 519–522.

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, *61*(4), 1277–1286.

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2012). What eye movements can tell about theory of mind in a strategic game. *PLoS One*, *7*(9), e45961.

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, *65*, 276–291. https://doi.org/10.1016/j.neubiorev.2016.03.020

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, *4*(2), 61–64.

Mumford, J. A., Poline, J.-B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. *PLoS One*, *10*(4), e0126255.

Nagel, R., Brovelli, A., Heinemann, F., & Coricelli, G. (2018). Neural mechanisms mediating degrees of strategic uncertainty. *Social Cognitive and Affective Neuroscience*, *13*(1), 52–62.

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: Psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604–609.

Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. Cambridge, England: MIT press.

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.

Penny, W., & Holmes, A. (2004). Random effects analysis. In R. Frackowiak, J. Ashburner, W. Penny, S. Zeki, K. Friston, C. Frith, et al. (Eds.), *Human brain function* (pp. 843–850). San Diego: Elsevier.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*(3), 437–471.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293–305.

Prado, J., Chadha, A., & Booth, J. R. (2011). The brain network for deductive reasoning: A quantitative meta-analysis of 28 neuroimaging studies. *Journal of Cognitive Neuroscience*, *23*(11), 3483–3497.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Reber, T. P., Luechinger, R., Boesiger, P., & Henke, K. (2012). Unconscious relational inference recruits the hippocampus. *Journal of Neuroscience*, *32*(18), 6138–6148.

Rilling, J. K., King-Casas, B., & Sanfey, A. G. (2008). The neurobiology of social decision-making. *Current Opinion in Neurobiology*, *18*(2), 159–165.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*(4), 1694–1703.

Robalino, N., & Robson, A. (2012). The economic approach to 'theory of mind'. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*(1599), 2224–2233.

Robinson, J. L., Laird, A. R., Glahn, D. C., Blangero, J., Sanghera, M. K., Pessoa, L., ... Young, K. A. (2012). The functional connectivity of the human caudate: An application of meta-analytic connectivity modeling with behavioral filtering. *NeuroImage*, *60*(1), 117–129.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage*, *19*(4), 1835–1842.

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*(11), 1435–1446.

Schiebener, J., & Brand, M. (2015). Self-reported strategies in decisions under risk: Role of feedback, reasoning abilities, executive functions, short-term-memory, and working memory. *Cognitive Processing*, *16*(4), 401–416.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34.

Seger, C. A., & Cincotta, C. M. (2006). Dynamics of frontal, striatal, and hippocampal systems during rule learning. *Cerebral Cortex*, *16*(11), 1546–1555.

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286–1291.

Spunt, R. P., & Adolphs, R. (2014). Validating the why/how contrast for functional MRI studies of theory of mind. *NeuroImage*, *99*, 301–311.

Spunt, R. P., Elison, J. T., Dufour, N., Hurlemann, R., Saxe, R., & Adolphs, R. (2015). Amygdala lesions do not compromise the cortical network for false-belief reasoning. *Proceedings of the National Academy of Sciences*, *112*(15), 4827–4832.

Spunt, R. P., & Lieberman, M. D. (2012). Dissociating modality-specific and supramodal neural systems for action understanding. *Journal of Neuroscience*, *32*(10), 3575–3583.

Spunt, R. P., Satpute, A. B., & Lieberman, M. D. (2011). Identifying the what, why, and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, *23*(1), 63–74.

Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences*, *19*(12), 719–728.

Symeonidou, I., Dumontheil, I., Chow, W.-Y., & Breheny, R. (2016). Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of Experimental Child Psychology*, *149*, 81–97.

Tager-Flusberg, H., & Sullivan, K. (1994). A second look at second-order belief attribution in autism. *Journal of Autism and Developmental Disorders*, *24*(5), 577–586.

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.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642–1645.

Toelch, U., & Dolan, R. J. (2015). Informational and normative influences in conformity from a Neurocomputational perspective. *Trends in Cognitive Sciences*, *19*(10), 579–589. https://doi.org/10.1016/j.tics.2015.07.007

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289.

Van der Meer, L., Groenewold, N. A., Nolen, W. A., Pijnenborg, M., & Aleman, A. (2011). Inhibit yourself and understand the other: Neural basis of distinct processes underlying theory of mind. *NeuroImage*, *56*(4), 2364–2374.

Wan, X., Nakatani, H., Ueno, K., Asamizuya, T., Cheng, K., & Tanaka, K. (2011). The neural basis of intuitive best next-move generation in board game experts. *Science*, *331*(6015), 341–346.

Wan, X., Takano, D., Asamizuya, T., Suzuki, C., Ueno, K., Cheng, K., ... Tanaka, K. (2012). Developing intuition: Neural correlates of cognitive-skill learning in caudate nucleus. *Journal of Neuroscience*, *32*(48), 17492–17501.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

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.

Zhu, L., Mathewson, K. E., & Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic

learning. *Proceedings of the National Academy of Sciences, 109*(5), 1419–1424.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

---

**How to cite this article:** Zhen S, Yu R. Neural correlates of recursive thinking during interpersonal strategic interactions. *Hum Brain Mapp.* 2021;42:2128–2146. https://doi.org/10.1002/hbm.25355