



HAL
open science

Neurocomputational mechanisms engaged in detecting cooperative and competitive intentions of others

Rémi Philippe, Rémi Janet, Koosha Khalvati, Rajesh P N Rao, Daeyeol Lee, Jean-Claude C Dreher

► **To cite this version:**

Rémi Philippe, Rémi Janet, Koosha Khalvati, Rajesh P N Rao, Daeyeol Lee, et al.. Neurocomputational mechanisms engaged in detecting cooperative and competitive intentions of others. 2022. hal-03855508

HAL Id: hal-03855508

<https://cnrs.hal.science/hal-03855508v1>

Preprint submitted on 16 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Neurocomputational mechanisms engaged in detecting cooperative and competitive intentions of others

Rémi Philippe

CNRS - Institut des Sciences Cognitives Marc Jeannerod <https://orcid.org/0000-0003-2894-7688>

Rémi Janet

CNRS, Institut des Sciences Cognitives Marc Jeannerod

Koosha Khalvati

Paul G. Allen School of Computer Science and Engineering, University of Washington

Rajesh Rao

University of Washington

Daeyeol Lee

Johns Hopkins University <https://orcid.org/0000-0003-3474-019X>

Jean-Claude Dreher (✉ dreher@isc.cnrs.fr)

CNRS, Institut des Sciences Cognitives Marc Jeannerod <https://orcid.org/0000-0002-2157-1529>

Article

Keywords:

Posted Date: January 18th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1160167/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1

2 **Neurocomputational mechanisms engaged in**
3 **detecting cooperative and competitive intentions of**
4 **others**

5

6

7

8 **R. Philippe¹, R. Janet¹, K Khalvati², R.P.N. Rao^{2,3}, D Lee⁴, JC. Dreher^{1*}**

9

10

11

12 ¹ CNRS-Institut des Sciences Cognitives Marc Jeannerod, UMR5229, Neuroeconomics, reward, and
13 decision making laboratory. 67 Bd Pinel, 69675 Lyon, FRANCE

14 ² Paul G. Allen School of Computer Science and Engineering, University of Washington, 185 Stevens
15 Way, Seattle, WA 98195

16 ³ Center for Neurotechnology, University of Washington, Seattle, WA 98195

17 ⁴ The Zanvyl Krieger Mind/Brain Institute, Kavli Neuroscience Discovery Institute, Department of
18 Neuroscience, Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N.
19 Charles St, Baltimore, MD 21218, USA

20

21 * Corresponding author

22

23

24

25

26

27

28

29 **Abstract**

30 Humans frequently interact with other agents whose intentions can fluctuate over time between
31 competitive and cooperative strategies. How does the brain decide whether the others' intentions are
32 to cooperate or compete when the nature of the interactions is not explicitly signaled? We used model-
33 based fMRI and a task in which participants thought they were playing with another player. In fact, this
34 agent was an algorithm alternating without signaling between cooperative and competitive strategies.
35 A neurocomputational mechanism underlying arbitration between competitive and cooperative experts
36 outperforms other learning models in predicting choice behavior. The ventral striatum and ventromedial
37 prefrontal cortex tracked the reliability of this arbitration process. When attributing competitive
38 intentions, these regions increased their coupling with a network that distinguish prediction error related
39 to competition *versus* cooperation. These findings provide a neurocomputational account of how the
40 brain dynamically arbitrates between cooperative and competitive intentions when making adaptive
41 social decisions.

42

43 Introduction

44 During social interactions, humans are often uncertain whether others intend to compete or
45 cooperate. The intentions of other agents can fluctuate over time, making it challenging to develop
46 behavioral strategies. A key question is to understand how the brain decides whether the other is
47 cooperating or competing during volatile situations in which the nature of the social interactions is not
48 explicitly determined, as when others interact to achieve a common goal while maximizing their own
49 benefits. This question is of importance since it lies at the heart of strategic social decision making¹⁻⁹.
50 In these types of situations, other agents can change behavior according to cooperative or competitive
51 intentions. Cooperation is generally defined as involving a group of individuals working together to attain
52 a common goal^{10,11}. In contrast, competition involves one person attempting to outperform another in
53 a zero-sum situation¹². A number of theoretical accounts and experimental results demonstrate that
54 the ability to mentalize, i.e. to simulate the other's belief about one's next course of action, is crucial for
55 strategically sophisticated agents^{6,7,13,14}. The neurocomputational mechanisms engaged in attributing
56 intentions to others has been studied in situations in which participants are explicitly informed about the
57 nature of the interactions, either in a collaborative context alone¹⁵⁻¹⁷ or in a competitive context alone
58^{8,18-25}. For example, during a cooperative game such as the coordination game, one of the best
59 strategies is to try to choose one of two presented targets consistently. In contrast, in a competitive
60 game such as the matching pennies game^{19,25}, the optimal strategy is to choose between two targets
61 equally often and randomly across trials. If the identity of the game played is not known, the agent has
62 to adjust his/her strategy based on repeated interactions with others and to infer
63 cooperation/competition on the basis of observations. How the brain achieves such inference poses a
64 unique computational problem because it not only requires the recursive representation of reciprocal
65 beliefs about other's intentions, as in cooperative or competitive contexts alone, but it also requires one
66 to decide whether the other is competing or cooperating to deploy an appropriate behavioral strategy.

67 Here, we sought to determine the neurocomputational mechanisms that underlie the inferences
68 of whether a person is competing or cooperating during volatile situations in which the nature of the
69 interactions is not explicitly signaled. A recent computational account proposed that arbitration between
70 strategies is determined by their predictive reliability, such that control over behavior is adaptively
71 weighted toward the strategy with the most reliable prediction²⁶. This approach has been tested
72 successfully in the domains of instrumental or Pavlovian action selection²⁷, model-based and model-
73 free learning²⁸ and learning by imitation or emulation²⁹. Extending this concept of a mixture of experts
74 to social interactions, we investigated whether the brain relies on distinct experts to compute the best
75 choice between two possible intentions attributed to others (cooperation or competition) and then
76 weights them by their relative reliability. We tested and compared these mixtures of models, that
77 attribute intentions to others dynamically, with different classes of learning models: non-Bayesian vs
78 Bayesian and non-mentalizing vs mentalizing (**see table 1**). This allowed us to identify the algorithms

79 and brain mechanisms engaged with a key component of the estimation of other's intentions, i.e.
80 whether the social partner was cooperating or competing.

81 The majority of theoretical frameworks used to model feedback-dependent changes in decision
82 making strategies, such as choice reinforcement and related Markov Decision Process (MDP) models,
83 assume that optimal decisions can be determined from the observable events and variables by the
84 decision makers. Clearly, these assumptions do not capture the reality and complexity of human social
85 interactions because observable behaviors of other individuals provide only very partial information
86 about their likely future behaviors. Moreover, model-free RL algorithms assume that values (utility or
87 desirability of states and actions), change incrementally across trials. This assumption is incorrect when
88 option values change abruptly, such as when the intention of the other shifts between cooperation and
89 competition. These limitations explain why agents basing their behavior only on standard RL models
90 can be exploited by opponents using more sophisticated algorithms ^{6,30}.

91 A more accurate account of strategic learning is based on a family of RL models which adds a
92 mathematical term to the classical Temporal Difference (TD) algorithm to consider the other as an agent
93 having their own policy, which can be influenced by oneself ^{6,30,31}. For example, fictitious play learning
94 proposes a basic form of mentalizing by having a representation of the other's strategy. Influence
95 models also consider that RL can be supplemented by a mentalizing term that represents how our
96 actions influence those of others, updated through a belief prediction error ^{2,6,19,21,30,32-34}. Such influence
97 models formalize not only how players react to others' past choices, (first-order beliefs in Theory of
98 Mind: ToM), but also how they anticipate the influence of their own choices on the others' behavior (i.e.,
99 mentalizing-related second-order beliefs). Another modeling approach of theory of mind used Bayesian
100 algorithms to model inferences about the future actions of another, attempting to take their point of view
101 and to simulate their decision ^{13,17,35}. This strategy can be performed recursively so that participants
102 make inferences concerning the others' inferences and so on. Such a sophisticated approach could be
103 grounded in the theoretical framework of Partially Observable Markov Decision Processes (POMDPs)
104 ³⁶. POMDPs provide a probabilistic framework for solving tasks involving action selection and decision
105 making under uncertainty ^{37,38}. Notably, this approach has recently been applied to strategic cooperation
106 in groups ^{36,39,40}. These models, however, have mainly been limited to signaled cooperative or
107 competitive tasks where the intentions of players do not change over a given period^{13,35,41,42}.

108 Here, we tested the predictions of these different families of learning models against one
109 another, investigating not only non-Bayesian vs Bayesian models and non-mentalizing vs mentalizing
110 models, but also a mixture of models deploying an arbitration process whereby the influence of
111 attributing intentions to others is dynamically modulated depending on which type of intention (i.e.
112 cooperative vs competitive) is most suitable to guide behavior at a given time. We did so by using a
113 novel model-based fMRI design (**Fig. 1**) consisting of an iterative dyadic game in which participants
114 were told that they would interact with another person via a computer. Unbeknownst to them, the other

115 player was an artificial agent that switched between blocks of cooperative trials and blocks of
116 competitive trials when playing a card matching game. Thus, the algorithm's goals were the same as
117 those of participants in the Cooperative blocks but were orthogonal in Competitive blocks. Participants
118 remained uncertain with respect to the goals of their "partner" or "opponent", which alternated, without
119 being signaled. This task allowed us to investigate the algorithms used by the brain to recognize the
120 "intentions" of others and to adopt appropriate strategies when the modes of interaction (cooperation
121 vs competition) are not indicated.

122 We found that the model accounting best for behavior was a mixture of influence models,
123 referred to as Mixed-Intention Influence Model. Two expert systems work together to make strategic
124 decisions, one assessing competitive intentions and the other assessing cooperative intentions, a
125 controller weighting between these experts according to their relative reliabilities. Each expert system
126 uses a classic RL algorithm complemented with a mentalizing term to infer the other's actions. This
127 Mixed-Intention Influence model accounts for behavior observed in naturalistic environments in which
128 the other's goal is often only partially congruent with one's own, allowing for a continuous range of
129 behavior between pure cooperation and pure competition. A brain network including the ventromedial
130 prefrontal cortex (vmPFC) and the ventral striatum tracked the reliability signal from the controller. This
131 finding indicates that the Mixed-Intention Influence model captures the higher-order structure of the
132 mixed-intentions task (i.e., alternation between cooperation and competition). When comparing trials
133 classified as competitive versus cooperative by the controller, we also identified a brain system engaged
134 with an updating signal used for learning. Finally, when participants expected higher utility for choosing
135 according to a competitive rather than cooperative strategies, the vmPFC and the ventral striatum,
136 tracked the intentions of others and showed changes in functional connectivity with a brain system
137 including the right temporo-parietal junction (rTPJ), dorsolateral prefrontal cortex (dlPFC) and the intra-
138 parietal sulcus (IPS), which discriminates reward prediction error (PE) between believed modes of
139 interaction. Together, these results provide a model-based account of the neurocomputational
140 mechanisms guiding human strategic decisions during games in which the intentions of others fluctuate
141 between cooperation and competition.

142

143 Results

144 Behavioral signature of tracking intentions

145 We assessed how participants used the history of previous interactions to make their choices.
146 We used logistic regression to examine whether participants selected the same target as that from the
147 previous trial (“Stay”) or chose the other target (“Switch”), depending on whether the previous three
148 trials (at t-1, t-2 and t-3) had been won or lost, whether the previous decisions had been to Stay or
149 Switch, and whether the previous interactions from those trials indicated cooperation (see below). We
150 also added sex, age and the number of trials as control variables. All trials except the 5 initialization
151 trials were included in this analysis. Cooperation was defined by a binomial variable, representing the
152 interaction between the last action of the Artificial Agent (AA) and the participant’s own previous
153 outcome (“Cooperativity signature”). This variable was set to 1 if either the participant had won on the
154 previous trial and the AA stayed on the same target for the next trial, or if the participant lost on the
155 previous trial and the AA switched to the other target the trial just after. Otherwise, the variable was set
156 to 0. Indeed, from the perspective of the participant, if the AA is a cooperative partner, both players win
157 at the same time and should then choose to keep the same target to be more predictable.

158 We found that the “Cooperativity signature” predicted an increase in the “stay” probability of
159 participants at t-1 and t-2 (Cooperativity signature t_{-1} : *estimate* = 0.05, $p=0.021$;
160 Cooperativity signature t_{-2} : *estimate* = 0.05, $p=0.002$, χ^2 test **Fig. 2a**). This suggests the participants
161 tracked whether the other agent was cooperating during the two previous trials (but not before).
162 Participants used the outcome of the latest trial to make the next decision (staying or switching target)
163 according to a win/stay, lose/switch strategy (winning $_{t-1}$: *estimate* = 0.16, $p<0.0005$, χ^2 test **Fig. 2a**).
164

165 Computational models tracking intentions of the other agent

166 To elucidate the computations underlying strategic decision making, we compared the results
167 of different computational models. These models were split into five classes (**see SI**). The first class of
168 models, based on heuristics, included Win-Stay/Lose-Switch and Random Bias models. The other four
169 classes of algorithms can be classified into non-Bayesian *versus* Bayesian model families along one
170 dimension and mentalizing *versus* non-mentalizing model families along the other dimension. Thus, the
171 second class of models includes non-Bayesian, non-mentalizing models represented by reinforcement
172 learning (RL) models. The third class represents non-Bayesian mentalizing models, namely the
173 “influence models” which are RL models with an additional term representing how the actions of one
174 player influence those of the other player. The fourth class corresponds to Bayesian non-mentalizing
175 models, including (1) a Hierarchical Gaussian Filter (HGF) which tracks the volatility of outcome⁴³, (2)
176 the k -Bayesian Sequence Learner which tracks the probability that one target will be selected by the
177 AA after a history of specific length k and (3) the *active inference* model which minimizes the expected

178 free energy⁴⁴. The fifth class of models contains Bayesian mentalizing models, which are the k -ToM
179 models using recursive Bayesian inferences of depth k to predict the future choice of the AA. Each
180 mentalizing model was tested using 3 versions: a competitive, a cooperative and a ‘mixed intentions’
181 version. The ‘mixed intentions’ version computes one decision value according to a competitive expert
182 and another according to a cooperative expert and arbitrates between the two, based on the difference
183 in their respective reliability (see SI, Fig. 3a). We defined reliability as the difference in unsigned value
184 functions for two choices given by particular learning algorithms.

185 Next, we performed a group-level random-effect Bayesian model selection on the models’
186 computed free energy, taking into account potential outliers and the number of free parameters^{45,46}.
187 We found that the ‘Mixed Intentions Influence Learning’ Model was the most frequent best fit across the
188 population (Fig. 2b), demonstrating that subjects employed mentalizing-related computations in our
189 mixed intentions task. This finding also indicates that arbitration between a cooperative and a
190 competitive expert best explains most participants’ behavior, rather than either expert taken individually.
191 Additionally, only the Mixed-Intention Influence model (and not the cooperative or the competitive one)
192 succeeded in producing behavior, similar to participants, with respect to the effect of the Cooperativity
193 signature on the probability to stay (Fig. 3b and Extended Fig.1, see SI). We conducted a logistic
194 regression to understand how the Mixed-Intention Influence model explained differences in behavioral
195 strategy to stay or switch target. This analysis included the reward prediction error at $t-1$, the valence
196 of the arbitration between cooperative and competitive intention at time t ($\text{sign}(\Delta)$; 1 for cooperative and
197 -1 for competitive), and the interaction between these two variables. This analysis revealed a main
198 effect of the valence of the arbitration (*valence of the arbitration* _{t} : *estimate* = 0.24, $p < 0.0005$, χ^2
199 test Fig. 3c) indicating that participant tend to stay more on the same target when they attributed
200 cooperative intention to the other. Moreover, we found an interaction effect, i.e. participants did not
201 integrate the prediction error in their strategy in the same way given the attributed intention
202 (*valence of the arbitration* _{t} * *rPE* _{$t-1$} : *estimate* = 0.20, $p = 0.0227$, χ^2 test Fig. 3c). That is, higher
203 negative prediction errors increased the probability that the participant would stay on the same target
204 when the controller attributed cooperative intentions compared to when it attributed competitive
205 intentions. In addition, we also performed another logistic regression analysis using the same variables
206 and the actual mode of interaction (i.e. competitive block trials *versus* cooperative block trials), rather
207 than the classified mode of interaction made by the controller. We did not find the same interaction
208 effect when we compared actual competitive and cooperative block trials (*Block type* _{t} *
209 *rPE* _{$t-1$} : *estimate* = 0.003, $p = 0.56$, *Block type* _{t} : *estimate* = 0.03, $p = 0.229$, and *PE* _{$t-1$} : *estimate* =
210 0.14, $p < 0.0005$, χ^2 test; Extended Fig.2, see SI), showing that the classified intentions, but not
211 competitive or cooperative blocks, affected the use of prediction error.

212 We reasoned that when facing an individual who can change his/her intentions to compete or
213 cooperate over time, the brain may rely on distinct experts to compute the best choice based on these

214 two possible intentions (i.e. cooperative or competitive), weighted by their relative reliabilities. We
215 therefore built such an ‘arbitrator’ computation as a sigmoid function of the difference in reliability
216 between the Cooperative and Competitive interactions (Δ), added to a bias (δ) that characterized each
217 individual’s tendency to attribute competitive ($\delta > 0$) or cooperative ($\delta < 0$) intentions to others. To assess
218 intentions of the other, participants only have access to the outcomes of previous interactions, the
219 choice (to stay or switch) of the artificial agent on previous trials, and the interaction between these two
220 types of information.

221 We hypothesized that repeated social victories should favor the attribution of cooperative
222 intentions because a series of victories suggests that both players are satisfied with the outcome. In
223 such situations the other player (i.e. AA) has become more predictable, which is an important feature
224 to build cooperation⁴⁷. Moreover, the interaction between outcome and AA’s choice (i.e. the tendency
225 of the AA to “stay” after a participant wins or “switch” after a participant loses) should drive the arbitrator
226 to favor the cooperative mode, because playing the same winning target for both players corresponds
227 to the optimal Nash equilibrium of the cooperative game. To test this hypothesis, we regressed the
228 signed difference in reliability on (1) the participant’s last outcome, (2) AA’s choice to “stay” or to “switch”
229 and (3) the interaction between the participant’s outcome and the AA’s choice to stay or switch
230 (Cooperativity signature) over up to three retrospective trials. We found that the past two interactions
231 between participant’s outcome and AA’s action (Cooperativity signature), the last outcome and switches
232 by the AA at trial $t-2$ and $t-3$ explained the difference in reliability
233 (*Cooperativity signature* $t-1$: *estimate* =0.59, $p < 0.0005$;
234 *Cooperativity signature* $t-2$: *estimate* =0.06, $p = 0.037$; *Victory*: *estimate* =1.98,
235 $p < 0.0005$; *switch* $t-2$: *estimate* =-0.18, $p < 0.0005$; *switch* $t-3$: *estimate* =-0.07, $p < 0.0005$), χ^2 test,
236 **Fig. 4a**).

237 Together, these analyses show that participants’ behavior, when alternating between
238 unsigned cooperative and competitive blocks, is best explained by the Mixed-Intention Influence
239 model. According to these findings, people use mentalization to update their beliefs about future chosen
240 targets, and dynamically arbitrate between the predicted intentions of the other agent to compete or
241 cooperate (**Fig. 3a**).

242

243 **Model-based fMRI analyses**

244 First, we constructed a GLM (GLM1) to identify brain regions tracking the arbitration process
245 (i.e. Δ : signed reliability difference, reliability for cooperation minus that of competition) between the two
246 experts (one for cooperation, the other for competition). We added the reliability difference Δ as
247 parametric regressor at the decision stage, and the expected reward for staying on the same target, as
248 non-orthogonalized parametric regressors to allow them to compete for the variance. We added the

249 reward prediction error as a parametric regressor at the outcome time and we controlled for the other's
 250 intention effect by adding Δ as a non-orthogonalized regressor. The bilateral ventral striatum
 251 ($x,y,z=9,12,0$ and $x,y,z=-12,9,-6$), vmPFC ($x,y,z=6,45,-8$), postcentral gyrus ($x,y,z=-20,-44,48$), and
 252 middle cingulate cortex (MCC; $x,y,z=11,-15,57$, $p<0.05$ whole-brain family-wise error (FWE), **Fig. 4b**
 253 **and 4c**) tracked the difference in reliability between experts (Δ) at the decision time. Bilateral dorsal
 254 striatum (DS; $x,y,z=17,6,-12$ and $-14,3,-11$), bilateral orbitofrontal cortex (OFC; $x,y,z=44,36,-14$ and -44
 255 $52,8$), posterior cingulate cortex (PCC; $x,y,z=2,-34,38$), and bilateral angular gyrus ($x,y,z=45,-30,46$ and
 256 $-54,-62,39$) ($p<0.05$ FWE, **Fig. 5a**) encoded the reward prediction error at the outcome time.

257 To investigate brain areas encoding the reward prediction error that were more engaged when
 258 the controller classified a trial as competitive vs cooperative, we tested another GLM (GLM2). Trial
 259 onsets were separated according to whether the value of the signed reliability difference Δ added to the
 260 bias was positive or negative. If this value was ≥ 0 , the trial was classified as Cooperative, and
 261 Competitive otherwise. The computed expected reward for staying on the same target was used as a
 262 parametric regressor at the time of choice. We found that the right dlPFC ($x,y,z=35,11,36$), the IPS
 263 region ($x,y,z=50,-50,32$) and the right temporoparietal junction (rTPJ; $x,y,z = 51,-50,33$, $p<0.05$ FWE)
 264 were more engaged in encoding reward prediction error in trials classified as Competitive versus
 265 Cooperative ($p<0.05$, FWE **Fig. 5b and 5c**). This effect could not be explained by less variance in the
 266 PE regressor in trials classified as Competitive trials compared to those classified as Cooperative,
 267 because we observed no difference in regressor variance on these two types of trials ($p=0.57$, Levene's
 268 test). No region was more engaged in trials classified as Cooperative compared to those classified as
 269 Competitive.

270 To further investigate the relationship between the behavior to stay after a trial classified as
 271 Competitive *versus* classified as Cooperative and the BOLD signal, we conducted a logistic regression.
 272 Explanatory variables were the average of the weighted time series in the dlPFC and rTPJ/IPS region
 273 observed in GLM2, the valence of the controller (Δ) and their interactions. This signal was extracted at
 274 the time of the outcome and convolved with the hemodynamic function. Because there was an average
 275 of 6.6 scans per trial, we took as a predictor variable the mean of this weighted signal through the 7
 276 scans following the outcome presentation. We found an interaction between the weighted time series
 277 and the valence of the arbitration (*valence of the arbitration_t * weighted time series_{t-1} : estimate =*
 278 *-14.4, p = 0.0312, χ^2 test*). *Post hoc* tests further revealed that this effect was driven by trials classified
 279 as Competitive (*weighted time series_{t-1} of trial classified as competitive_{t-1} : estimate =*
 280 *-37.3, p = 0.040, χ^2 test*). This result indicates that activation of dlPFC and rTPJ/IPS increases the
 281 probability of switching following a trial classified as Competitive.

$$282 \quad \text{logit}(P(\text{stay})) = \beta_0 + \beta_1 * \text{sign}(\Delta) + \beta_2 * \text{Time Series} + \beta_3 * \text{sign}(\Delta) * \text{Time Series}$$

283

284 **Connectivity analysis**

285 Finally, we performed a generalized psycho-physiological interaction (gPPI) seed to voxels
286 connectivity analysis to understand the interactions between brain regions tracking the arbitration
287 process (i.e. Δ : reliability difference) for the cooperative and competitive experts and those more
288 engaged with the PE when the controller attributes more competitive than cooperative intentions to the
289 other (**see Online Methods**). We used the ventral striatum and vmPFC, which encoded the controller,
290 as seed regions (ROI extracted from the GLM1 striatal and vmPFC activity) for trials classified as
291 competitive compared to those classified as cooperative (i.e. trials for $\Delta < 0$ or $\Delta > 0$) at the decision time.
292 We found stronger functional connectivity between regions encoding the difference in reliability and the
293 right dlPFC ($x,y,z=38,34,34$), the left IPS region ($x,y,z=-48,-44,58$) and the left TPJ ($x,y,z=-42,-40,50$,
294 $p < 0.05$ FWE; **Fig. 6a**) at the decision time for trials classified as Competitive compared to those
295 classified as Cooperative. This result indicates that the dlPFC, IPS and left TPJ receive an input signal
296 at the time of choice according to the difference in reliability with respect to the Competitive *versus*
297 Cooperative believed intention of others. This brain network largely overlaps with the brain network
298 observed with the PE differentiating between trials classified as Competitive *versus* Cooperative at the
299 outcome (**Fig. 5b** and **Extended Fig. 4**).

300

301 Discussion

302 To make a strategic decision when facing an individual with unknown and fluctuating intentions,
303 it is necessary to make inferences as to whether we are in a competitive or cooperative situation. In the
304 context of minimal information, for example when only the choices of the other, but not their outcomes,
305 are available, such inferences are much more difficult than when one is in a specific known setting (e.g.,
306 in a competitive game) ¹⁹. Here, we provide evidence that the brain engages in dynamic tracking of
307 another individual's cooperative/competitive intentions, despite having no explicit information regarding
308 whether the situation is cooperative or competitive. We found that strategies of participants were mostly
309 affected by the outcomes of previous interactions and by a "signature" of the other's cooperativity, i.e.
310 the tendency of the other (here the Artificial Agent or AA) to stay on the same target after the
311 participant's winning. Comparison between computational models demonstrated that such behavior is
312 best explained by a model in which choice is driven by a controller tracking the reliability difference
313 between cooperative and competitive intentions. The fMRI results show that the neural computations
314 of this controller are implemented in the ventral striatum and in the vmPFC. Thus, both behavior and
315 brain imaging results can be accounted for by a model that includes a controller that allocates resources
316 according to different experts' predictions. At the time of outcome, a common brain network, including
317 the rostral anterior cingulate cortex (rACC), ventral striatum and lateral OFC encoded prediction error
318 in trials classified as Competitive or as Cooperative. However, prediction error signals in some brain
319 areas also depended on the classification of the current trial as Cooperative or Competitive as classified
320 by the controller. That is, a distinct brain network, composed of the bilateral dlPFC, bilateral IPS regions
321 and the rTPJ was more engaged for trials classified as Competitive compared to those classified as
322 Cooperative. This latter brain network reflects a differential use of the outcome of the social interaction
323 as a function of whether it is classified as Competitive or Cooperative (**Fig. 2d**).

324 Mentalizing processes are essential to correctly infer the strategy of others. This is true in the
325 cooperative context, in which participants performed above chance, reflecting their ability to effectively
326 infer the other's (i.e. AA's) behavior. In the competitive context, participants performed below chance
327 level, showing that the AA was able to predict their behavior and to exploit their previous
328 choices/outcomes. The Mixed-Intention Influence model had the best ability to predict data and to
329 generate very similar behavior to the participants. Each expert model is an expanded RL model, with a
330 term accounting for one's previous choice influencing the choice of the other. Although only the
331 influence term differed between the Competitive and Cooperative models, the Mixed-Intention Influence
332 model tracked intentions based on this second order mentalizing term by weighting the contribution of
333 a cooperative and of a competitive expert. One key aspect of this Mixed-Intention Influence model is
334 that it captures higher order structures (fluctuations between cooperation and competition) during social
335 interactions. In contrast, one important limitation of the classical RL model is that it does not exploit
336 higher-order structures such as interdependencies between different stimuli, actions, and subsequent

337 rewards. Previous studies demonstrated that models incorporating such structures can account for
338 individual decision making in different situations ^{48–51}. Here, we demonstrate that the representation of
339 abstract states, such as whether the other is cooperating or competing, can be extended to social
340 decisions and underlies the ability to build strategies. To confirm that the Mixed-Intention Influence
341 model accounted more for neural activity in brain areas involved in social interactions, we formally
342 compared the brain regions covarying more with the expected reward for staying on the same target,
343 as computed by the winning model, compared to the expected reward for staying on the same target,
344 computed by a simple RL model (**Extended Fig. 3**). One crucial difference between a simple RL model
345 and the Mixed-Intention Influence model is that in the former, only the value of the chosen option is
346 updated and the valuation of the option that was not chosen does not change. In the latter, both the
347 values of the chosen and unchosen options are updated to incorporate the knowledge that the current
348 state has a given reliability to be cooperative or competitive. The controller weights the valuation
349 produced according to the competitive or cooperative hypothesis which is computed as a sigmoid of
350 the difference in reliability between the two experts.

351 Activity in the ventral striatum and vmPFC increased as the cooperative prediction from the
352 controller became more reliable than the competitive prediction, as a result of the outcomes of the
353 previous interactions (reliability difference modulated by last outcome) and the other's "Cooperativity
354 signature" over the last trials. Thus, these brain regions dynamically track the reliability difference
355 between intentions classified as cooperative and competitive in a situation where the nature of the social
356 interactions is implicit. Previous reports demonstrated a role of the ventral striatum when making
357 cooperative choices alone, in response to a partner's cooperative choice in an explicit cooperation task
358 ⁵² and also in the attribution of intentions in a competitive context ³⁰. Our findings show that strategic
359 social behavior can be explained by a Controller Theory according to which cooperative/competitive
360 social behavior results from the interaction of multiple systems, each proposing possible strategies for
361 action ^{26,28,29}.

362 One strength of our computational approach was to assess and compare a large variety of
363 competing models, (active inference, recursive learning models k-TOM and a mixture of experts using
364 k-TOM, influence models for only cooperative strategies or competitive strategies, a mixture of experts
365 using influence models, fictitious learner, Bayesian Sequence learner, Hierarchical Gaussian Filter,
366 Reinforcement Learning and Heuristic models). Many have never previously been directly tested
367 against each other. Our results agree with studies concluding that social learning may be driven by non-
368 specific reinforcement processes that include a mentalizing term ^{3,6,8,30,53}. We demonstrate that when a
369 task is not explicitly signaled as cooperative or competitive, this evokes the arbitration between
370 strategies determined by predictive reliability. Behavior is hence controlled by giving a higher weight to
371 the strategy with the most reliable prediction ²⁹. At first glance, it may be surprising to observe that the
372 mixture of expert influence models performs better than mathematically more sophisticated models,

373 such as POMDP models (HGF, active inference, k-ToMs) and models mimicking different levels of
374 mentalizing (k-TOM). However, this is likely because in our setting the only information that can be
375 integrated by participants is their own choices, rewards and the history of the choices made by the other
376 (i.e. AA). The nature of the social interaction is never explicitly signaled (participants are not told whether
377 the other is cooperative or competitive), and the rewards of the other are not observed. This uncertainty
378 could therefore result in the failure of POMDP models to reproduce human behavior, particularly when
379 sudden flips occur between the AA strategies. This contrasts with previous neuroimaging studies that
380 investigated learning of social interactions in either competitive or cooperative situations alone
381 (matching-pennies or rock-paper-scissors games against computerized opponents) ^{25,54}. Our findings
382 also broadly agree with a cognitive hierarchy of strategic learning mechanisms, proposing that distinct
383 levels of strategic thinking correspond to different levels of sophistication of learning mechanisms (in
384 increasing order of complexity: reinforcement learning, fictitious play learning and influence learning)
385 ⁵⁵. However, we propose a more general model based on a mixture of influence learning experts that
386 function in parallel and are then compared with respect to their relative reliability.

387 Competitive social interactions often emerge in situations where an agent's outcome depends
388 on the choices of others, which requires the ability to infer the intentions of others ⁶. In the context of
389 our mixed intentions task, when participants attributed competitive, as opposed to cooperative
390 intentions to others, the dlPFC and rTPJ/IPS specifically encoded a relative PE. Although this network
391 has previously been reported when inferring the intentions of others ^{8,39,56}, the strength of our
392 computational account of theory of mind processes is to specify that this brain network computes a PE
393 difference between trials that the controller classified as competitive *versus* cooperative. This PE
394 difference reflects a differentiation in the implementation and use of the outcome of the social interaction
395 as a function of the classified interaction (**Fig. 3c**). Note that PE was not more volatile in trials when the
396 competitive expert is more reliable than the cooperative expert. This rules out the possibility that the
397 observed PE difference reflects higher PE volatility in competitive contexts. When comparing intentions
398 classified as cooperative compared to competitive, participants tended to be more predictable, staying
399 more on the same target after experiencing an unexpected social defeat (i.e. after higher negative PE)
400 (**Fig. 3c**). This behavior likely reflects a signal sent to the other to indicate one's willingness to stay on
401 the same target, despite bearing the cost of staying on this target ^{6,8,57}. This is a key feature of successful
402 coordination ⁴⁷ in which agents who want to trigger reciprocity ⁵² are willing to incur a cost to promote
403 cooperation from the other.

404 Finally, we found higher functional connectivity between seed regions that encode the reliability
405 difference of the controller (vmPFC and striatum) and brain regions more engaged in PE for trials
406 classified as Competitive *versus* Cooperative (dlPFC, TPJ) (**Fig. 6**). This indicates that brain regions
407 engaged in the input of the arbitration process at the time of choice are more strongly coupled (resp.
408 decoupled) with brain regions encoding PE for intentions classified as Competitive (vs Cooperative).

409 This reflects a differential use of the outcome of the social interaction as a function of whether it is
410 classified as Competitive or Cooperative. Thus, according to the attributed intention of other by the
411 controller, PE signals differed and the strength of functional coupling increased between regions
412 encoding the reliability difference of the controller and the dlPFC-TPJ network. When one expert is
413 more reliable than the other, the Mixed-Intention Influence model predicts that the reward PE is driven
414 by the valuation of the more reliable expert. Since the only difference between experts is the sign of the
415 second-order mentalizing term, this suggests that the implication of the dlPFC-TPJ network at the
416 outcome is engaged when there is a need to mentalize intentions of other agents with opposing goals
417 (i.e. intentions classified as Competitive). Moreover, engagement of the dlPFC-TPJ network increases
418 the probability of switching following a trial classified as competitive by the controller, allowing
419 behavioral adaptation by virtue of the reliability difference signal.

420 Together, our work provides evidence that the mixture of experts model explains behavior in
421 socially volatile situations differing only by the reward function of other agents. These two experts only
422 differ by their priors on how their reward function takes into account another agent's reward (i.e. the
423 second order mentalizing term), and were sufficient to discriminate the others' intentions. These findings
424 provide a mechanistic framework explaining the neurocomputations underlying learning in strategic
425 social interactions. We extend to theory of mind processes (i.e inferring cooperative vs competitive
426 intentions) a computational account similar to mixture of experts proposed to arbitrate between
427 strategies in other domains, such as exploitation vs exploration ^{58,59}, "model-based" vs "model-free"
428 systems ^{26,28,60} and learning by imitation vs emulation ²⁹. Finally, our Mixed-Intention Influence model
429 may be useful in the field of computational neuropsychiatry to identify the specific computational
430 components that are modified in theory of mind alterations, a key feature of autism spectrum disorder
431 ⁶¹.

432

433 **Author contributions**

434 J-C.D., R.P, R. P.N.R, and D.L. developed the general concept, experiment and models. R.P.
435 programmed the task and ran the experiment under the supervision of J.-C.D. R.P. developed the
436 models and implemented the algorithms under the supervision of J.-C.D., and analyzed the data in
437 collaboration with J.-C.D. R.P. and J.-C.D. wrote the manuscript in collaboration with K.K., D.L and
438 R.P.N.R.

439

440 **Acknowledgements**

441 This research has benefited from the financial support of IDEXLYON from Université de Lyon
442 (project INDEPTH) within the Programme Investissements d’Avenir (ANR-16-IDEX-0005) and of the
443 LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements
444 d’Avenir (ANR-11-IDEX-007) operated by the French National Research Agency. This work was also
445 supported by grants from the Agence Nationale pour la Recherche and NSF in the CRCNS program to
446 JCD (ANR n°16-NEUC-0003-01), and by CRCNS NIMH grant no. 5R01MH112166-03, NSF grant no.
447 EEC-1028725, and a Templeton World Charity Foundation grant to RPNR.. We thank the CERMEP
448 Staff for help during scanning and Pr Ed Derrington for proof reading the draft of the manuscript.

449

450 **Competing interests**

451 The authors declare that they have no competing interests.

452

453 **Resources availability**

454 **Lead contact**

455 Further information and requests for resources and inquiries should be directed to and will be
456 fulfilled by the lead contact, Jean-Claude DREHER (dreher@isc.cnrs.fr).

457 This study did not generate new unique reagents.

458 **Materials availability**

459 This study did not generate new unique reagents.

460 **Data and code availability**

461 Code supporting main results are available at GitHub
462 (https://github.com/remiphilipp/Mixture_intention.git).

463 **Method**

464 **Participants**

465 32 participants (aged 20-40, $M = 27$, $SD = 5.1$ - 17 women) were recruited via a daily local
466 newspaper and the University of Lyon 1 mailing list. All participants were screened to exclude those
467 with medical conditions including psychological or physical illnesses or a history of head injury to
468 prevent having confounding variables. They all provided informed consent and were paid a fixed
469 amount. However, they were financially motivated in being told that they would be paid as a function of
470 their decisions.

471 **Mixed intentions task**

472 Participants performed a novel task comprising 163 trials in an MRI scanner. They were led to
473 believe that they were interacting with another person via a computer interface, while in fact they were
474 playing against an artificial agent (AA) managed by a computer program. Such simulated social
475 interactions allowed us to investigate the dynamics and neural mechanisms arbitrating between multiple
476 learning algorithms. Participants were faced with a screen containing four cards, two face down (the
477 other player's cards) and two face up (their own cards). Participants were informed that to win, they had
478 to choose the card of the same color as the one the other person was going to choose. Experimenters
479 were careful not to specify whether the other was an adversary or a partner. Participants were told that
480 they and the other player had to make their choices in four seconds (**Fig. 1a**). If the Artificial Agent (AA)
481 played before the participant, one of the two face down cards was removed from the playing field. If the
482 participant chose first, only the selected card remained on the playing field. Then, when both had
483 chosen, the chosen cards were revealed and the participant received a reward if the card colors
484 matched, otherwise they received nothing. Participants were led to believe that their final payoff would
485 be increased by 10 ct (euro) for each winning interaction. No information about the other's payoff was
486 given to the participants, they only knew that after an interaction, the other 'participant' would see the
487 same screen but with their outcomes which could be different from the participant.

488 Importantly, unbeknownst to the participants, the artificial agent alternated between Competitive
489 and Cooperative trial blocks. During this mixed intentions task, the AA's strategy was determined by
490 alternating 13 trials of a hide and seek (HS) task (Competitive blocks), and 10 trials of a coordination
491 game (Cooperative blocks). The artificial agent algorithm was designed to predict the color that would
492 be chosen by the participant on the basis of a probabilistic analysis of the two previous choices and
493 outcomes (**see SI** for the algorithm). Here we defined a competitive choice, made by the AA, as
494 choosing the card of the color the participant was expected not to play and a cooperative choice as
495 choosing the card with the same color. Thus, the artificial agent exploited the bias of the participants in
496 a stochastic way, i.e. the more predictable the participant was, the more the algorithm made correct
497 competitive or cooperative choices (see SI). Participants were not informed of the switches between

498 the two blocks (Cooperative vs Competitive), however their goal was always to choose the same color
499 as that chosen by the other player (i.e. the AA).

500 The HS task is competitive, and the computer uses the record of the participant's choice and
501 reward history to minimize the participant's payoff. Therefore, in this case the subject's optimal strategy
502 during the HS task is to choose the two targets equally often and randomly across trials. During the
503 coordination game, the AA tried to maximize the subject's payoff and in this case the subjects should
504 try to choose one of the two targets consistently so that the computer can choose the same target as
505 them. Since the participant is not informed of either the goals of the AA or the switches between blocks,
506 they must adjust their strategy based on recent experience and infer cooperation/competition on the
507 basis of their observations.

508 This task was designed to identify key components of the estimation of intentions regarding
509 whether others are cooperating or competing. We took advantage of the fact that an individual's
510 estimates as to whether they are engaged in a cooperative or competitive interaction can be assessed
511 even when the individual is interacting with a computer program rather than another person. Transitions
512 between the competitive and cooperative blocks were unsignaled, therefore subjects had to discover
513 by trial and error the most successful strategy over consecutive blocks. This alternation between the
514 two interaction modes functioned well because the participant's winning rate was significantly higher in
515 cooperative (mean 60% std 1%) than in competitive (mean 44% std 1%) trials (paired t-test $p < 10^{-4}$).

516

517 **Artificial agent**

518 The AA calculated the probability p for the participant to select a particular target color based on
519 the history of the two previous choices and their outcomes. Then to make the artificial agent behave
520 more like a real person, this prediction was exploited in a probabilistic fashion (**see SI**). In the
521 cooperative mode the AA chose the color card it predicted with probability p . In the competitive mode
522 this color was chosen with probability $1-p$.

523

524 **Behavioral analysis**

525 For the logistic regressions, we reported significant marginal effect of a given variable under the
526 name "*estimate*" (for example: *Cooperativity signature* t_{-1} : *estimate*).

527 Logistic regression : $\ln\left(\frac{P}{1-P}\right) = x_0 + x_1X_1 + x_2X_2 + \dots$

528 X_i represents independent variables and x_i represents the associated weights in the logistic
529 regression. P represent the probability of a given event. The marginal effect of the variable X_1 is defined
530 as:

531
$$\hat{y}_1 = \text{mean}(\text{logit}^{-1}(x_1))$$

532 The mean is computed across all observed data. Thus, the marginal effect called “*estimate*” can
533 easily be interpreted as the discreet change of the dependent variable given a unitary change of an
534 independent variable.

535 For the linear regressions, reported “*estimate*” represents x_i i.e. the regression coefficient.
536 Indeed, in a linear regression, marginal effect of a variable is equal to the estimated coefficient.

537 **Models**

538 To test for a dynamic tracking of implicit intention we compared 14 models, 9 involved theory of
539 mind (*Inf,k-ToM*), the others were to control for other possible strategies. The influence models (*Inf*) rely
540 on aaylor expanded reinforcement learning⁶² to take into account the influence of one’s own strategy on
541 the strategy of the other. *k-ToM* models also take into account the influence of one’s own strategy on
542 the other but in a Bayesian fashion^{13,35}. These two models were adapted in their cooperative and
543 competitive versions. Moreover, we constructed an adaptation of these two models (*Inf,k-ToM*) in which
544 an arbitrator weights the cooperative and competitive versions according to their reliability before
545 making the decision. Finally, because k-ToM is a recursive model (“I think that you think that...”), we
546 included k-ToM of depth one and two for each version.

547 To control for strategies that did not include theory of mind we added 5 other models including
548 two Bayesian inference types (*HGF* and *BSL*). The Hierarchical Gaussian Filter (*HGF*)^{43,63} basically
549 tracks the external volatility of the artificial agent choices in a Bayesian hierarchical way. The Bayesian
550 Sequences Learner (*BSL*) strategy relies on Bayesian inference given past sequences of choices. In a
551 model free analysis, we found that participants tended to use the past 2 choices to make their next
552 choice, so we used sequences of depths 2 and 3. Finally, we added two non-Mentalizing non Bayesian
553 models, a reinforcement learning model (RL) and a model based on the heuristic Win/stay –
554 Lose/Switch that we observed in the model free analysis.

555 The Bayesian Model Selection (BMS) was performed using the VBA toolbox (Variational
556 Bayesian Analysis) in a random effect analysis relying on the free energy as the lower bound of model
557 evidence. We use protected Exceedance Probability measurements (pEP)⁴⁵ to select the model which
558 is used most frequently in our population.

559

560

<i>Model</i>	<i>Mentalizing</i>	<i>Bayesian</i>	<i>Mixed intentions</i>
<i>Influence</i>	+	-	-
	(coop and comp)		
<i>Fictitious</i>	+	-	-
<i>Influence mixed intentions</i>	+	-	+
<i>k-ToM</i>	+	+	-
	(coop and comp)	(depth 1 and 2)	
<i>k-ToM mixed intentions</i>	+	+	+
<i>Active Inferences</i>	-	+	-
<i>HGF</i>	-	+	-
<i>BSL</i>	-	+	-
		(depth 2 and 3)	
<i>RL</i>	-	-	-
<i>Wst/Lsw</i>	-	-	-

562 **Table 1.** Classification of models according to 3 categories. The first depends on the ability of the model
563 to mentalize, the second depends on whether the model is a Bayesian model, and the third concerns
564 models that could be used with a mixture of experts.

565

566 **fMRI data acquisition**

567 MRI acquisitions were performed on a 3 Tesla scanner using EPI BOLD sequences and T1
568 sequences at high resolution. Scans were performed in a Siemens Magnetom Prisma scanner
569 HealthCare at CERMEP Bron (single-shot EPI, TR / TE = 1600/30, flip angle 75°, multiband acquisition
570 (accelerator factor of 2), in an ascending interleaved manner with slices interlaced 2.40 mm thickness,

571 FOV = 210 mm. We also use the iPAT mode with an accelerator factor of 2 and the GRAPPA method
572 reconstruction. The number of volumes acquired varied given the time the participant took to make their
573 decisions. The first acquisition was made after stabilization of the signal (3 TR). Whole-brain high-
574 resolution T1-weighted structural scans (0.8 x 0.8 x 0.8 mm) were acquired for each subject, co-
575 registered with their mean EPI images and averaged across subjects to permit anatomical localization
576 of functional activations at the group level. Field map scans were acquired to obtain magnetization
577 values that were used to correct for field inhomogeneity.

578

579 **fMRI data analysis**

580 Image analysis was performed using SPM12 (Wellcome Department of Imaging Neuroscience,
581 Institute of Neurology, London, UK, fil.ion.ucl.ac.uk/spm/software/spm12/). Time-series images were
582 registered in a 3D space to minimize any effect that could result from participant head-motion. Once
583 DICOMs were imported, functional scans were realigned to the first volume, corrected for slice timing
584 and unwarped to correct for geometric distortions. Inhomogeneous distortions-related correction maps
585 were created using the phase of non-EPI gradient echo images measured at two echo times (5.20 ms
586 for the first echo and 7.66 ms for the second). Finally, in order to perform group and individual
587 comparisons, they were co-registered with structural maps and spatially normalized into the standard
588 Montreal Neurological Institute (MNI) atlas space using the DARTEL method. Then we ran ARTrepair
589 to deweight scans that could include movement artefacts ⁶⁴.

590 We ran general linear models (GLMs) analyses to identify which brain regions encoded: (a)
591 one's belief that one is interacting in a cooperative or in a competitive situation (Δ); (b) the reward
592 prediction error (PE) after interactions classified as cooperative or competitive; (c) the PE difference
593 between the trials classified as cooperative vs competitive. In every GLM, an event was defined as a
594 stick function. The participant's button press and the AA's selection of target were defined as onset of
595 no interest in all GLMs. For all GLMs, missing trials were modeled with four events (cue, participant's
596 button press, AA's choice and outcome) as separate onsets without additional parametric regressors.
597 Head movement parameters were added as parametric regressors of no interest to account for motion-
598 related noise. Because the behavioral analysis showed that the bias towards competitive interaction
599 affects the strategy of participants, we added the competitive bias (δ) as a covariate at the second level
600 analysis in all GLMs.

601 Specifically, in GLM1, there were 4 onsets, including the time of the cue presentation (cards on
602 screen), participant's button press, AA's choice and the feedback time. Parametric regressors were the
603 difference in reliability Δ , the expected reward for staying at the time of the cue onset and the reward
604 prediction error (PE) at the feedback time, as well as Δ , to control for the effect of the believed intention

605 of the other on the PE brain encoding. For each GLM, we turned off the serial orthogonalization function
606 of regressors to allow it to compete for the variance.

607 In a second GLM (GLM2), we separated trials given the sign of $\Delta - \delta$ (positive or negative) to
608 identify brain regions specifically engaged in cooperative or competitive mental states (δ is a free
609 parameter capturing the participant's bias toward competitive intent). Δ refers to the difference in
610 reliability of cooperative and competitive prediction and δ is the competitive bias. For this GLM, there
611 were 6 onsets, including the cue for trials classified as cooperative or competitive, participant's button
612 press, AA's choice and the feedback time for trials classified as cooperative or competitive. Trials were
613 classified as either cooperative or competitive and parametric modulators were: the difference in
614 reliability Δ and the expected reward for staying on the same target at the time of the cue and the PE
615 and Δ at the time of feedback. Three participants who always attributed the same intention to the AA
616 were not included in GLM2.

617 To test the additional hypothesis that brain activation observed for believed other's intentions (in
618 **Fig. 4b**) is also present in competitive vs cooperative blocks, we conducted two more GLMs. The first,
619 GLM3 is similar to GLM2, i.e., we separated trials into two categories (cooperative and competitive),
620 but the differentiation was made using the real mode of interaction of the AA rather than the
621 classification made by the controller. Other onsets and parametric regressors were left unchanged.

622 Finally, a last GLM was applied to check that the results observed in GLM2 were not simply due
623 to the effect of volatility of the rewarded target. This GLM (GLM4) is similar to GLM2, i.e. trials were
624 classified according to the sign of $\Delta - \delta$. The only difference was that we added the actual probability
625 that the AA would choose the same target as the previous trial as a parametric regressor at both the
626 time of the cue and at the outcome.

627 We computed one sample t-tests with contrasts for main effect of Δ in GLM1 and effect of PE at
628 the outcome time. Then we computed the contrast between competitive and cooperative PE regressors
629 in GLM2, GLM3 and GLM4. Finally, we computed a paired t-test between this contrast, derived from
630 GLM2 and GLM3, to formally show that activation coming from the difference between classified trials
631 was significantly higher than those coming from the difference between the actual modes of interaction
632 as determined by the trial block.

633 Reported brain areas show a significant activity at the threshold of $p < 0.05$, whole brain family-
634 wise error (FWE), corrected for multiple comparisons at the cluster level (threshold at $P < 0.001$
635 uncorrected).

636

637 **Psychophysiological interaction (PPI) analysis**

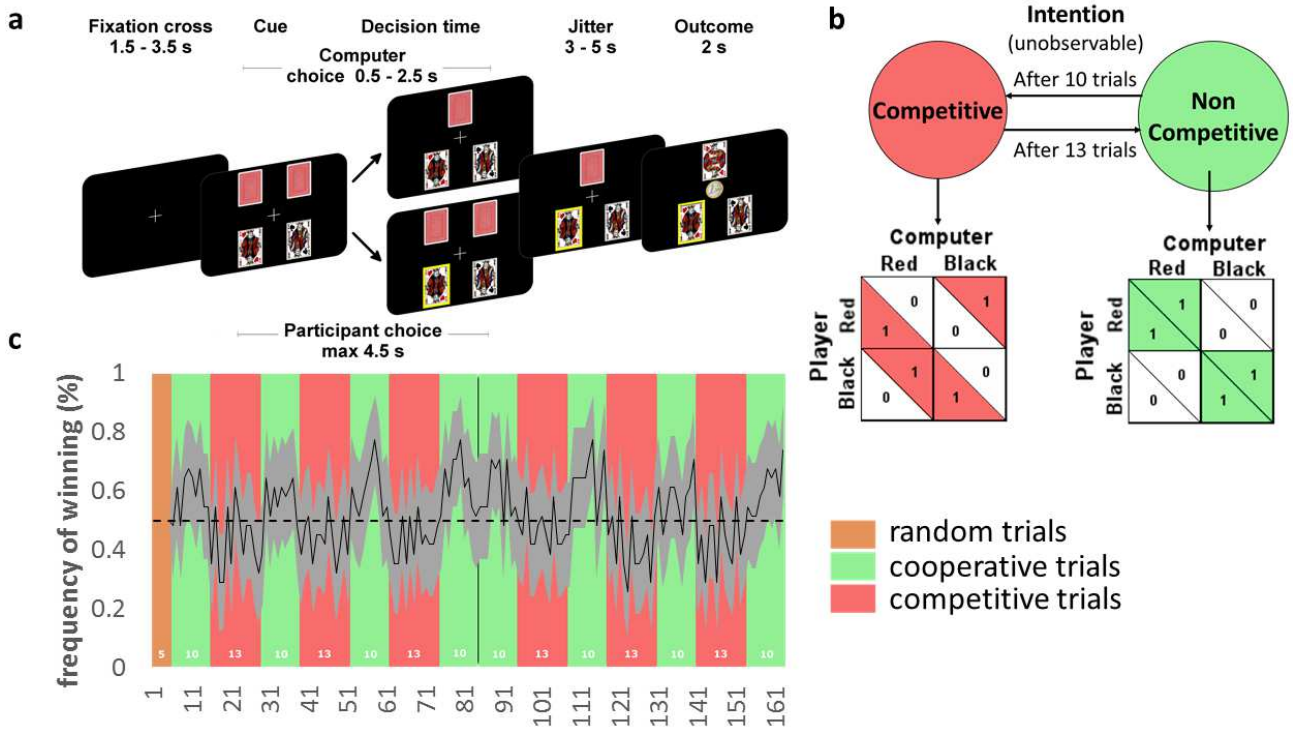
638 We defined the attribution of cooperative or competitive intentions at the time of decision making
639 as the psychological factor. Thus, we were able to investigate the difference in functional connectivity
640 when making a decision under cooperative or competitive intent. For this PPI analysis, we focused on
641 decision time and on functional connectivity between regions encoding the others' intentions and all
642 other voxels. Thus, for the physiological factor we took the BOLD signal of the striatal region elicited in
643 GLM1 as encoding the intention of others. Otherwise, we used same regressor parameters and onsets
644 as GLM2.

645 Reported brain areas show a significant activity at the threshold of $p < 0.05$, whole brain family-
646 wise error (FWE) corrected for multiple comparisons at the cluster level (threshold at $P < 0.001$
647 uncorrected).

648

649 **Figures**

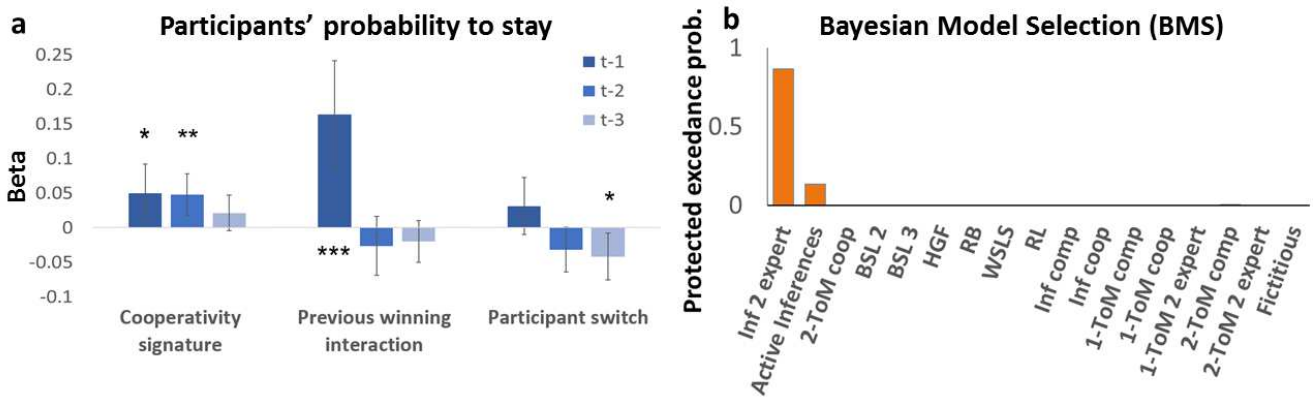
650



651

652 **Figure 1.** fMRI experiment. **a.** After a fixation cross, four cards were presented on the screen. The two
 653 cards shown on top of the screen represent the cards presented to the opponent/partner (i.e. Artificial
 654 agent), and not seen by the participant while the two kings (one black and one red) are the cards
 655 presented to the participant (shown in the bottom of the screen). The participants had to choose
 656 between these two cards. At the time of decision, the upper screen represents the display if the AA
 657 makes it choice first, while the lower screen shows how one card is highlighted with yellow border if
 658 the participant makes his choice first. Then a screen presents the participant's and Artificial Agent's choices
 659 together. Finally, at the time of outcome the participant wins if both he/she chooses the same card as
 660 the AA (here red king). **b.** Payoff matrix of the two types of block. **c.** Frequency of winning (black line)
 661 during competitive (red background) and cooperative (green background) blocks. The grey area
 662 represents the 95% confidence interval. The orange background represents 5 initial trials in which the
 663 AA played randomly for initialization purpose.

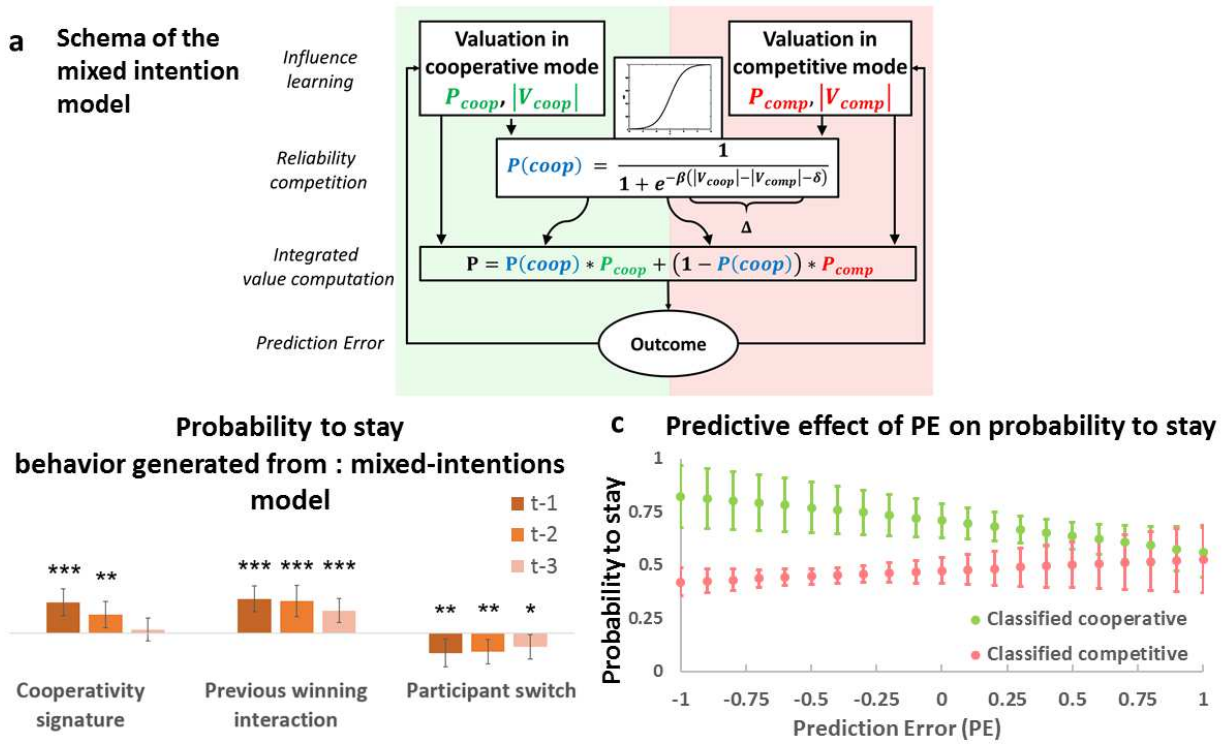
664



666

667 **Figure 2. a.** Model-free analysis. Random-effect logistic regression of the decision to stay after
 668 selecting a specific target with respect to the action of the artificial agent “Cooperativity signature” (i.e.
 669 participant wins then AA stays or participant loses then AA switches), the previous winning interaction
 670 (i.e. success or failure of past trials) and the choice to switch or stay, over the previous three trials. Error
 671 bars are the 95% confidence interval. **b.** Model comparisons based on Bayesian model selection. The
 672 protected exceedance probabilities indicate that the Mixed-Intention Influence model (Inf 2 expert)
 673 explains decisions in the mixed intention task better than others: Active inference; k-ToM; Bayesian
 674 Sequence Learner (BSL); Hierarchical Gaussian Filter (HGF); Reinforcement Learning (RL); Heuristic
 675 models: Random Bias (RB); Win/Stay-Lose/Switch (WSLS).

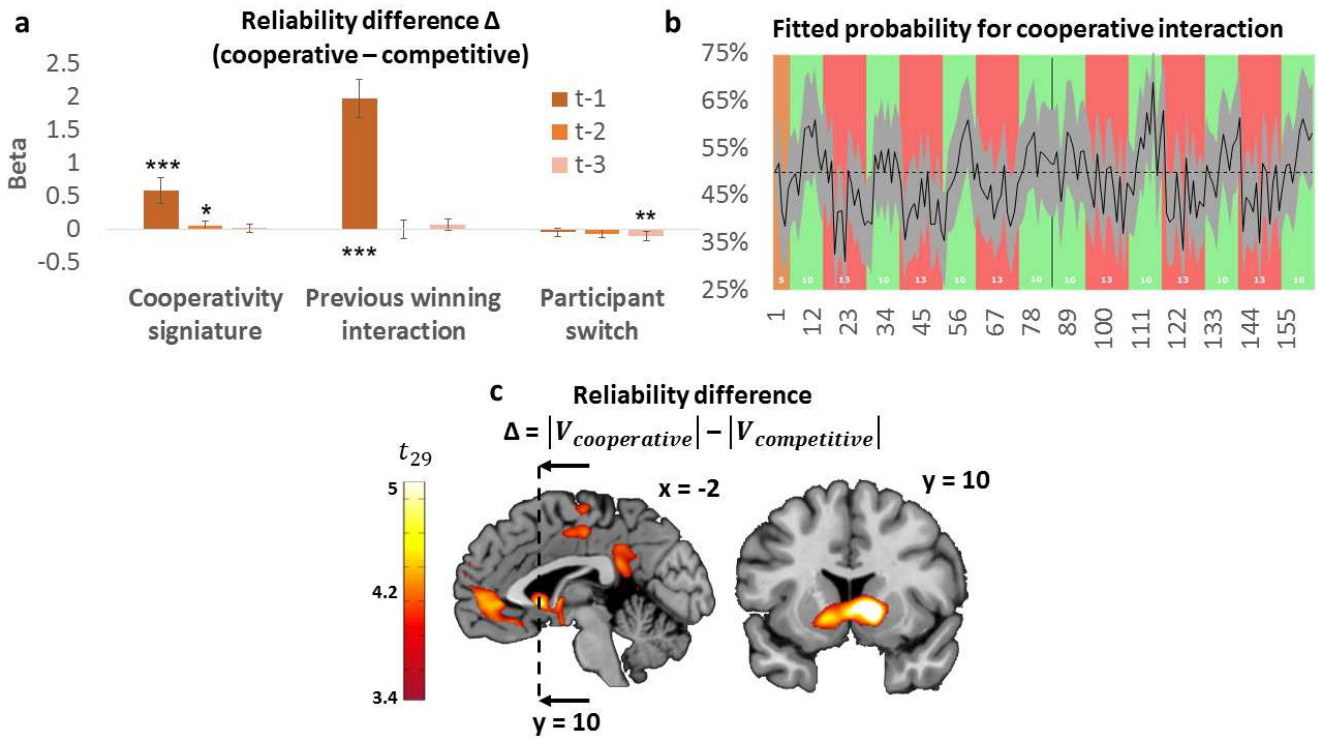
676



678

679 **Figure 3. a.** Scheme of the Mixed-Intention Influence model. Two influence models (one cooperative
 680 and the other competitive) compute a value for choosing one specific target (the black one). A controller
 681 uses the difference between the absolute value of the value of each expert (called reliability) to compute
 682 a probability that the other is cooperating. Then, the model weights the value of each expert according
 683 to the probability of being in cooperative and in competitive modes to produce a final decision value.
 684 Then it compares its predictions to the actual reward and compute again a new value for each expert.
 685 **b.** Model-free generative analysis. We generated one hundred sets of data using a free parameter from
 686 a normal distribution with mean and standard deviation calculated from the models fitted to the
 687 population, against the same artificial agent that participants played. We regressed the behavioral
 688 decision to stay after selection of a specific target on the previous trial depending on the interaction of
 689 the previous outcome and the action of the artificial agent (“Cooperativity signature”), the success or
 690 failure of up to three previous trials, and the action to switch or stay of the participant. Error bars are
 691 the 95% confidence interval (random-effect logistic regression). **c.** Marginal effect of the prediction error
 692 on the probability to stay on the same target in trials classified as Cooperative (green) and trials
 693 classified as Competitive (red). Error bars are the 95% confidence interval. *p < 0.05, **p < 0.01, ***p <
 694 0.001 (random-effect logistic regression).

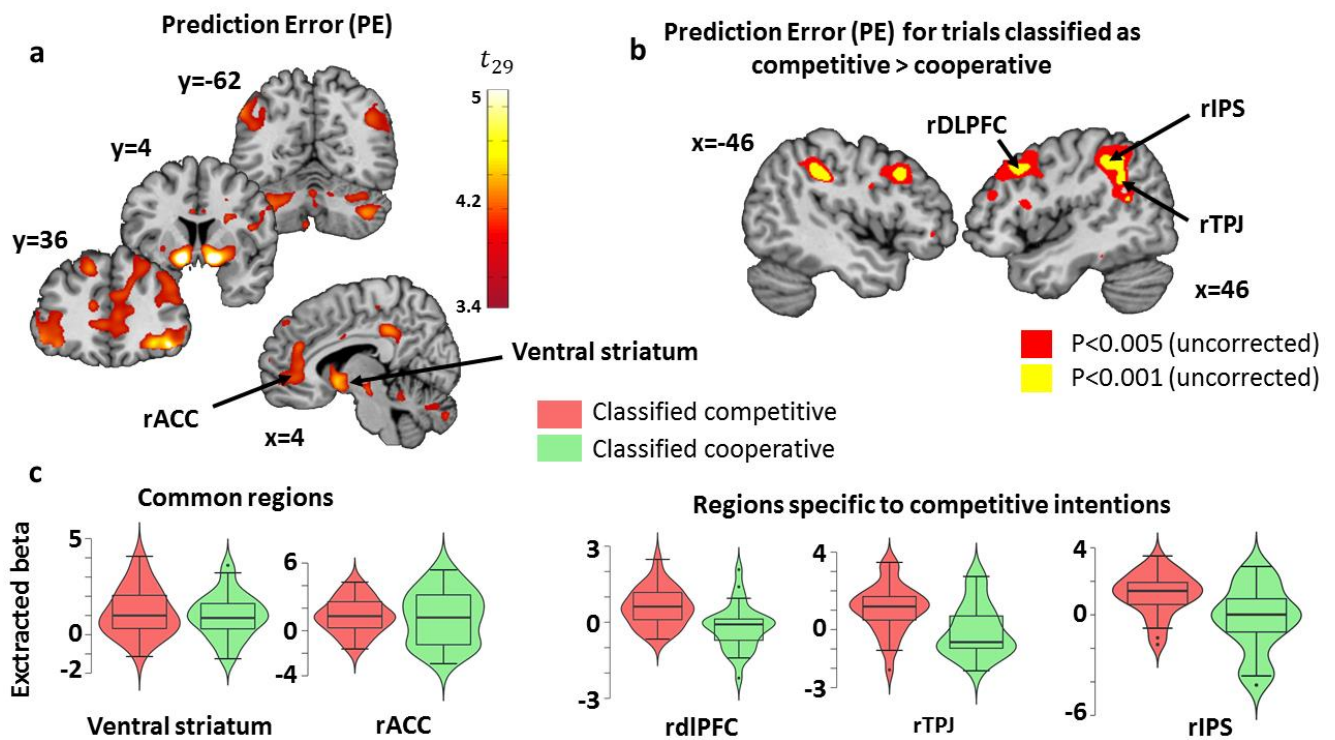
695



696

697 **Figure 4. a.** Difference in reliability is influenced by the Cooperativity signature of the Artificial Agent
 698 (AA), specifically the interaction of the previous subject's outcome followed by the action of the artificial
 699 agent (Participant wins then AA stays and Participant loses then AA switches), the latest outcome and
 700 the computer's switch at trial t-2 and t-3. Error bars are the 95% confidence interval. **b.** Mean probability
 701 of the participant attempting to cooperate across all participants (black line) for the 163 trials. The initial
 702 orange area is the 5 random initializing trials, green areas are the Cooperative blocks and red areas
 703 the Competitive blocks. The grey area is the 95% confidence interval. **c.** BOLD signal in ventral striatum,
 704 mPFC and posterior cingulate cortex (PCC) ($p < 0.05$ whole-brain family-wise error) are correlated with
 705 the difference in reliability, Δ , of estimated competitive and cooperative intentions.

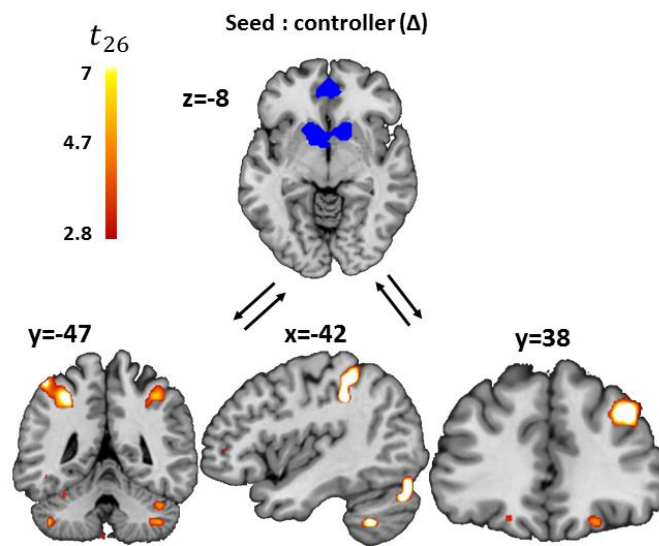
706



707

708 **Figure 5.** Correlations between BOLD activity and prediction error. **a.** Brain regions in which BOLD
 709 signal correlates with prediction errors for trials classified by the controller to be either competitive or
 710 cooperative. **b.** Brain regions in which BOLD activity correlates more with PE on trials estimated to be
 711 competitive compared to trials estimated to be cooperative. This network comprised dlPFC ($x,y,z =$
 712 $30,9,42$), IPS ($x,y,z = 42,-47,42$) and the rTPJ ($x,y,z = 51,-50,33$, $p < 0.05$ whole-brain family-wise error).
 713 **c.** Beta value extracted for trials estimated to be either competitive or cooperative. Left: regions in the
 714 ventral striatum (left $x,y,z = -14,3,-11$ + right $x,y,z = 17,6,-12$) and rACC ($x,y,z = 6,42,-3$) with increased
 715 activation in trials estimated to be either competitive or cooperative. Right: specific brain regions
 716 activated only when trials were classified as Competitive: dlPFC ($x,y,z = 30,9,42$), IPS ($x,y,z = 42,-$
 717 $47,42$) and rTPJ ($x,y,z = 51,-50,33$) from 8 mm spheres centered on peak activation.

718



720

721 **Figure 6.** Neural mechanisms of arbitration between the attributions of competitive and cooperative
 722 intentions to the AA. Connectivity analysis: the BOLD signal was extracted from seeds regions (mPFC
 723 and ventral striatum using GLM1) computing the reliability difference between cooperative and
 724 competitive intentions of others (in Blue). The psychophysiological interaction effect shows higher
 725 functional coupling (voxels in red) with the left TPJ ($x, y, z = -42, -40, 50$), left IPS ($x, y, z = -32, -48, 50$) and
 726 right dlPFC ($x, y, z = 38, 34, 34$, $p < 0.05$ FWE threshold at $p < 0.001$) in trials classified as competitive as
 727 compared to those classified as cooperative.

728

729 **Extended data**

730 **Extended table**

731

732 Extended results table of GLM1, related to **Figure 4c**.

Table 1. Brain region that covaries with difference in reliability of interaction mode

	MNI peak cluster coordinates			k	Z score
	x	y	z		
Difference in reliability ($\Delta=R_{coop} - R_{comp}$)					
Ventral Striatum	14	9	-2	3375	4,99
Left hippocampus	-35	-53	3	660	4,3
mPFC	6	45	-8	1580	4,24
Right hippocampus	32	-36	12	3836	4,11
Middle cingulate cortex	11	-15	57	1732	3,95
Left Angular	-32	-53	27	857	3,86
Left postcentral gyrus	-20	-44	48	661	3,85
Rcomp – Rcoop					
No brain region					

** cluster reported at $p < 0.05$ FWE whole brain cluster corrected (initial cluster-forming threshold of $p < 0.001$ uncorrected)

733

734

735 Extended results table of GLM2, related to **Figure 5a**.

736

Table 2. Brain region encoding both competitive and cooperative prediction error.

	MNI peak cluster coordinates			k	Z score
	x	y	z		
Positively					
Right Dorsal Striatum	17	6	-12	21895	8,79
Left Dorsal Striatum	-14	3	-11	3488	8,78
Left Cerebellum	-44	-74	-45	6381	5,49
Right Cerebellum	29	-71	-29	5299	5,33
PCC	2	-35	38	1285	5,3
Right Angular/TPJ	45	-30	47	3581	4,71
Left Angular/TPJ	-54	-62	39	2044	4,34
Negatively					
No brain region					

** cluster reported at $p < 0.05$ FWE whole brain cluster corrected (initial cluster-forming threshold of $p < 0.001$ uncorrected)

737

738

739 Extended results table of GLM2, related to **Figure 5c**.

Table 3. Brain region that responded differently for the reward prediction error of a trial estimated to be competitive rather than cooperative

	MNI peak cluster coordinates			k	Z score
	x	y	z		
Competitive > Cooperative					
Right dlPFC	35	11	36	1732	4,41
Right Angular/TPJ	50	50	32	1758	4,32
Medial superior frontal gyrus	-2	30	47	505	3,82
Cooperative > Competitive					
No brain region					

** cluster reported at $p < 0.05$ FWE whole brain cluster corrected (initial cluster-forming threshold of $p < 0.001$ uncorrected)

740

741 Extended results table of connectivity analysis, related to **Figure 6**.

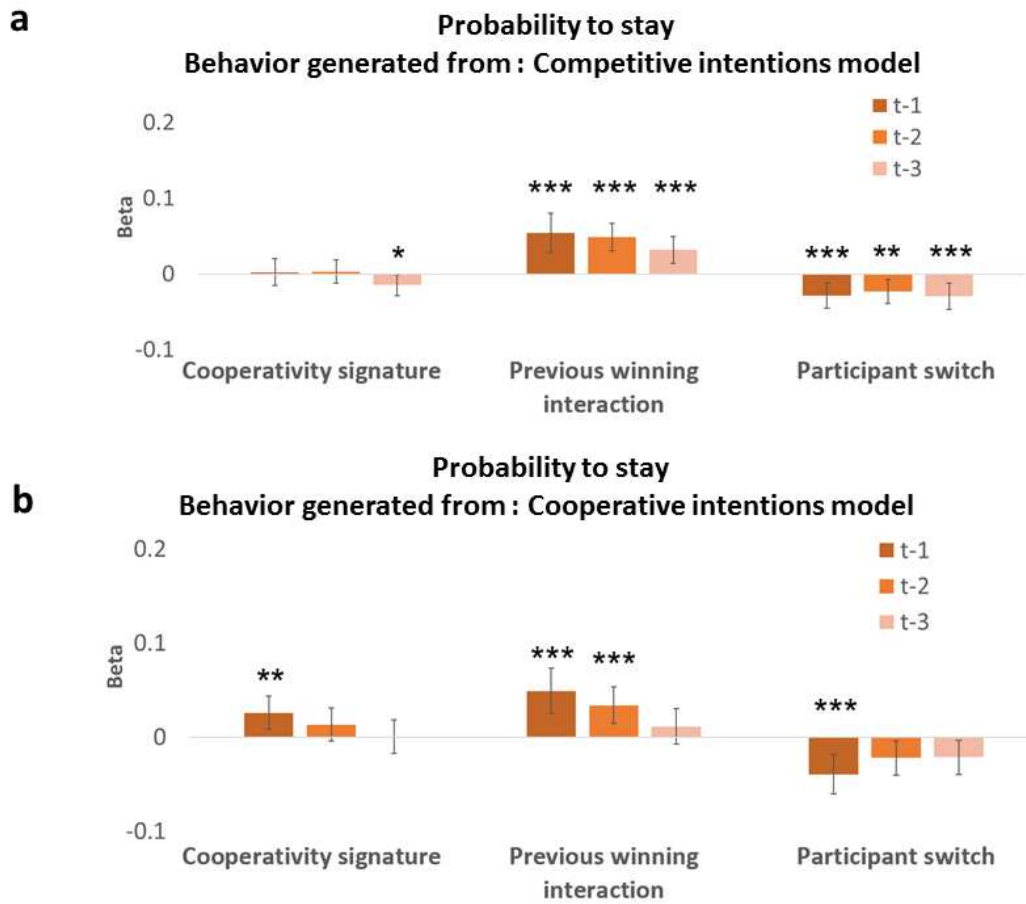
Table 4. Effect of the intention attribution at the decision time on the striatum/mPFC and dIPFC/TPJ connectivity strength.

gPPI Analyse	Seed	Peak regions	k	p-val unc.	p-val FWE	Z score
	<i>mPFC + striatum</i>	-48 -44 58	363	<0.0005	<0.0005	4.52
	<i>mPFC + striatum</i>	-2 -82 -36	153	<0.0005	0.037	4.34
Competitive > Cooperative	<i>mPFC + striatum</i>	38 34 34	150	<0.0005	0.040	4.33
	<i>mPFC + striatum</i>	-38 -78 -26	181	<0.0005	0.017	4.27
Cooperative > Competitive	mPFC + striatum	4 -46 16	744	<0.0005	<0.0005	4.48

742

743 **Extended figures**

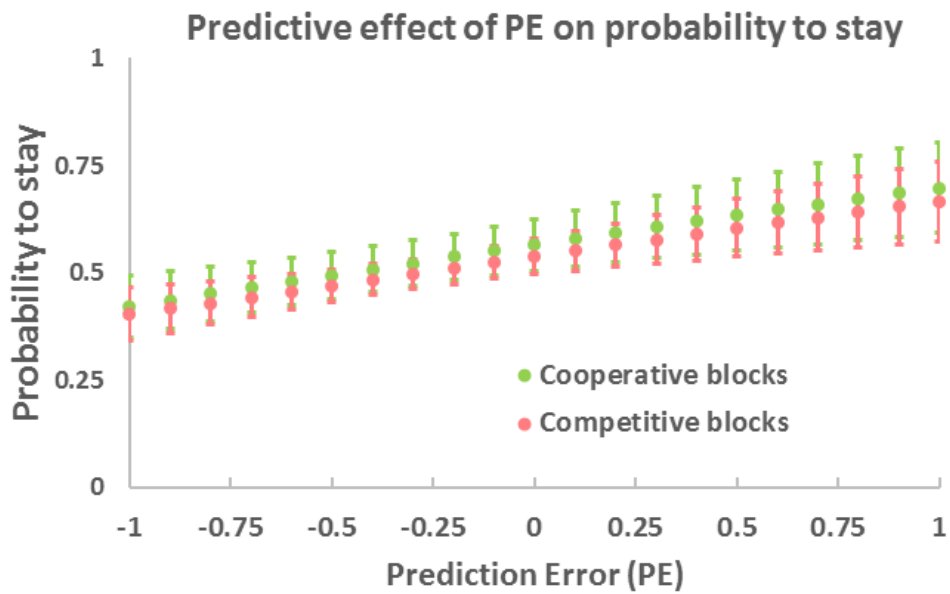
744



745

746 **Extended figure 1.** Model-based generative analysis. We generated one hundred sets of data using
747 free parameters from a normal distribution with mean and standard deviation calculated from the
748 "Influence models" in competitive (a) and cooperative (b) mode, fitted to the population. We generated
749 a data set against the same artificial agent that participants played. We regressed the interaction of the
750 previous outcome and action of the artificial agent (I win – AA stay / I lose – AA switch), the behavioral
751 decision to stay after selecting a specific target at the previous trial based on the success or failure of
752 the previous trial (Win) and the action to switch or stay of the artificial agent (Switch) in previous trials
753 up to three trials back. Error bars are the 95% confidence interval. *p < 0.05, **p < 0.01, ***p < 0.001
754 (random-effect logistic regression). Related to Computational models tracking intentions of the other
755 agent in the Result part.

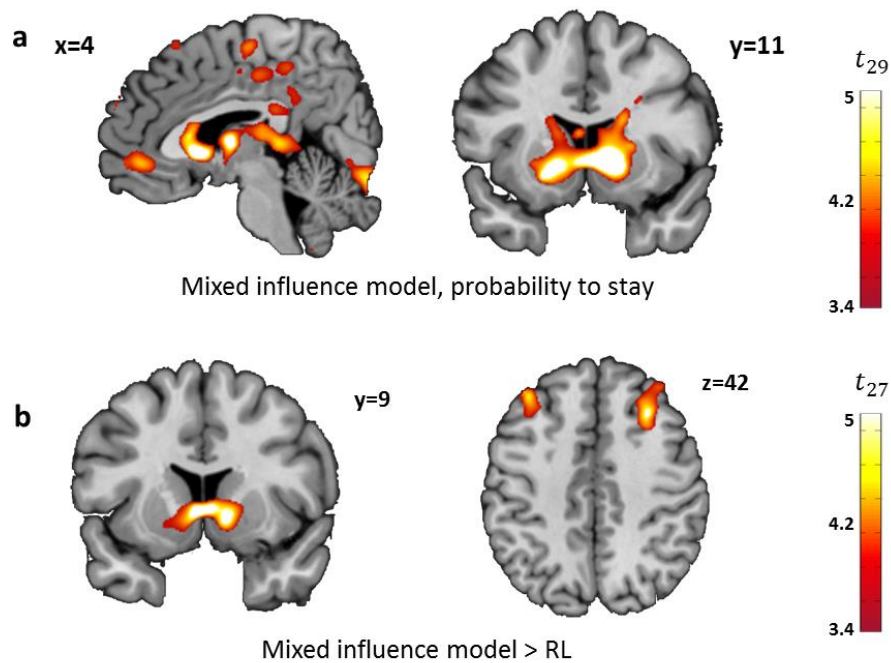
756



757

758 **Extended figure 2.** Marginal effect of the prediction error on the probability to stay on the same target
 759 in Cooperative blocks (green) and in Competitive blocks (red). Error bars are the 95% confidence
 760 interval. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (random-effect logistic regression). Related to Computational
 761 models tracking intentions of the other agent in the Result part.

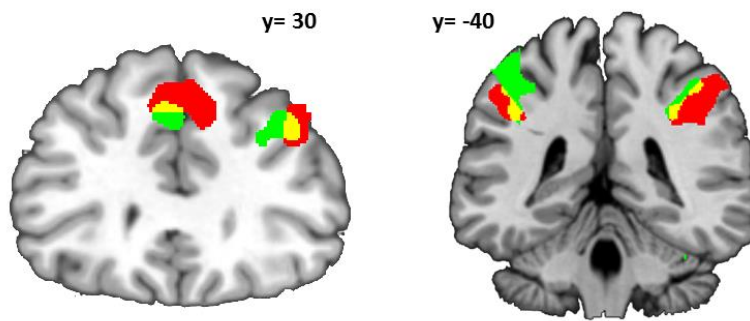
762



763

764 **Extended figure 3.a.** Neural correlates of the expected reward for staying on the same target as the
 765 previous trial, computed by the Mixed-Intention Influence model. (Significant ventral striatum correlation
 766 $x,y,z=14,11,-2$, $p<0.05$ FWE corrected threshold at $p<0.001$) **b.** Ventral Striatum ($x,y,z=6,12,0$), bilateral
 767 dlPFC ($x,y,z=-36, 33, 44$ and $x,y,z= 30,24,42$) and MTG ($x,y,z=65,-56,-8$, $p<0.05$ FWE corrected
 768 threshold at $p<0.001$) are best explained by the expected reward for staying of the Mixed-Intention
 769 Influence model rather than the expected reward for staying of a reinforcement learning model. Related
 770 to Model-based analyses in the result part.

771



772
773 **Extended figure 4.** Overlap (in yellow) of brain regions correlating more with PE when trials are
774 classified as competitive compared to those classified as cooperative (GLM2, in red), and of brain
775 regions more coupled with the controller region at the time of choice of trials classified as Competitive
776 compared to those classified as Cooperative (in green) ($p < 0.005$ for display purpose)

777

References

- 778
779
- 780 1. Carter, R. M. K., Bowling, D. L., Reeck, C. & Huettel, S. A. A distinct role of the temporal-parietal
781 junction in predicting socially guided decisions. *Science (80-.)*. **336**, 109–111 (2012).
 - 782 2. Bhatt, M. A., Lohrenz, T., Camerer, C. F. & Montague, P. R. Neural signatures of strategic types
783 in a two-person bargaining game. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19720–19725 (2010).
 - 784 3. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & F.S., M. Associative learning of social value.
785 *Nature* **456**, 245–249 (2008).
 - 786 4. Diaconescu, A. O. *et al.* Hierarchical prediction errors in midbrain and septum during social
787 learning. *Soc. Cogn. Affect. Neurosci.* **12**, 618–634 (2017).
 - 788 5. Suzuki, S., Adachi, R., Dunne, S., Bossaerts, P. & O’Doherty, J. P. Neural mechanisms
789 underlying human consensus decision-making. *Neuron* **86**, 591–602 (2015).
 - 790 6. Hampton, A. N., Bossaerts, P. & O’Doherty, J. P. Neural correlates of mentalizing-related
791 computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6741–
792 6746 (2008).
 - 793 7. Coricelli, G. & Nagel, R. Neural correlates of depth of strategic reasoning in medial prefrontal
794 cortex. *Proc. Natl. Acad. Sci.* **106**, 9163–9168 (2009).
 - 795 8. Hill, C. A. *et al.* A causal account of the brain network computations underlying strategic social
796 behavior. *Nat. Neurosci.* **20**, (2017).
 - 797 9. Ogawa, A. & Kameda, T. Dissociable roles of left and right temporoparietal junction in strategic
798 competitive interaction. *Soc. Cogn. Affect. Neurosci.* **14**, 1037–1048 (2019).
 - 799 10. Deutsch, M. A Theory of Competition and cCooperation. *Human Relations* **2**, 129–152 (1949).
 - 800 11. Deutsch, M. Cooperation and trust: Some theoretical notes. in *Nebraska Symposium on*
801 *Motivation, 1962*. 275–320 (Univer. Nebraska Press, 1962).
 - 802 12. Thibaut, J. W. & Kelley, H. H. *The Social Psychology of Groups*. (Wiley, 1969).
 - 803 13. Devaine, M., Hollard, G. & Daunizeau, J. The Social Bayesian Brain: Does Mentalizing Make a
804 Difference When We Learn? *PLoS Comput. Biol.* **10**, (2014).
 - 805 14. Camerer, C. F., Ho, T. H. & Chong, J. K. A cognitive hierarchy model of games. *Q. J. Econ.* **119**,
806 861–898 (2004).
 - 807 15. Mi, Q., Wang, C., Camerer, C. F. & Zhu, L. Reading between the lines: Listener’s vmPFC
808 simulates speaker cooperative choices in communication games. *Sci. Adv.* **7**, (2021).

- 809 16. Suzuki, S., Niki, K., Fujisaki, S. & Akiyama, E. Neural basis of conditional cooperation. *Soc.*
810 *Cogn. Affect. Neurosci.* **6**, 338–347 (2011).
- 811 17. Yoshida, W., Seymour, B., Friston, K. J. & Dolan, R. J. Neural mechanisms of belief inference
812 during cooperative games. *J. Neurosci.* **30**, 10744–10751 (2010).
- 813 18. Fareri, D. S. & Delgado, M. R. Differential reward responses during competition against in- and
814 out-of-network others. *Soc. Cogn. Affect. Neurosci.* **9**, 412–420 (2014).
- 815 19. Seo, H., Cai, X., Donahue, C. H. & Lee, D. Neural correlates of strategic reasoning during
816 competitive games. *Science (80-.)*. **346**, 340–343 (2014).
- 817 20. Lee, D., Conroy, M. L., McGreevy, B. P. & Barracough, D. J. Reinforcement learning and
818 decision making in monkeys during a competitive game. *Cogn. Brain Res.* **22**, 45–58 (2004).
- 819 21. Coricelli, G. & Nagel, R. Neural correlates of depth of strategic reasoning in medial prefrontal
820 cortex. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 9163–9168 (2009).
- 821 22. Lee, D., McGreevy, B. P. & Barracough, D. J. Learning and decision making in monkeys during
822 a rock-paper-scissors game. *Cogn. Brain Res.* **25**, 416–430 (2005).
- 823 23. Abe, H. & Lee, D. Distributed Coding of Actual and Hypothetical Outcomes in the Orbital and
824 Dorsolateral Prefrontal Cortex. *Neuron* **70**, 731–741 (2011).
- 825 24. Barracough, D. J., Conroy, M. L. & Lee, D. Prefrontal cortex and decision making in a mixed-
826 strategy game. *Nat. Neurosci.* **7**, 404–410 (2004).
- 827 25. Vickery, T. J., Chun, M. M. & Lee, D. Ubiquity and Specificity of Reinforcement Signals
828 throughout the Human Brain. *Neuron* **72**, 166–177 (2011).
- 829 26. O’Doherty, J. *et al.* Why and how the brain weights contributions from a mixture of experts. 1–18
830 (2021). doi:10.31234/osf.io/ns6kq
- 831 27. Dorfman, H. M. & Gershman, S. J. Controllability governs the balance between Pavlovian and
832 instrumental action selection. *Nat. Commun.* **10**, 1–8 (2019).
- 833 28. Wan Lee, S., Shimojo, S. & O’Doherty, J. P. Neural Computations Underlying Arbitration between
834 Model-Based and Model-free Learning. *Neuron* **81**, 687–699 (2014).
- 835 29. Charpentier, C. J., Iigaya, K. & O’Doherty, J. P. A Neuro-computational Account of Arbitration
836 between Choice Imitation and Goal Emulation during Human Observational Learning. *Neuron*
837 **106**, 687-699.e7 (2020).
- 838 30. Zhu, L., Mathewson, K. E. & Hsu, M. Dissociable neural representations of reinforcement and
839 belief prediction errors underlie strategic learning. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1419–1424
840 (2012).

- 841 31. Lee, D. & Seo, H. Neural Basis of Strategic Decision Making. *Trends Neurosci.* **39**, 40–48 (2016).
- 842 32. Seo, H. & Lee, D. Behavioral and Neural Changes after Gains and Losses of Conditioned
843 Reinforcers. *J. Neurosci.* **29**, 3627–3641 (2009).
- 844 33. Schwieren, C. & Weichselbaumer, D. Does competition enhance performance or cheating? A
845 laboratory experiment. *J. Econ. Psychol.* **31**, 241–253 (2010).
- 846 34. Bartolo, R. & Averbeck, B. B. Prefrontal Cortex Predicts State Switches during Reversal
847 Learning. *Neuron* **106**, 1044-1054.e4 (2020).
- 848 35. Devaine, M., Hollard, G. & Daunizeau, J. Theory of mind: Did evolution fool us? *PLoS One* **9**,
849 (2014).
- 850 36. Khalvati, K. *et al.* Modeling other minds: Bayesian inference explains human choices in group
851 decision-making. *Sci. Adv.* **5**, eaax8783 (2019).
- 852 37. Rao, R. P. N. Decision Making Under Uncertainty: A Neural Model Based on Partially Observable
853 Markov Decision Processes. *Front. Comput. Neurosci.* **4**, 1–18 (2010).
- 854 38. Khalvati, K., Kiani, R. & Rao, R. P. N. Bayesian inference with incomplete knowledge explains
855 perceptual confidence and its deviations from accuracy. *Nat. Commun.* **12**, 1–16 (2021).
- 856 39. Park, S. A., Sestito, M., Boorman, E. D. & Dreher, J. C. Neural computations underlying strategic
857 social decision-making in groups. *Nat. Commun.* **10**, 1–12 (2019).
- 858 40. Khalvati, K., Mirbagheri, S., Park, S. A., Dreher, J. C. & Rao, R. P. N. A Bayesian theory of
859 conformity in collective decision making. *Adv. Neural Inf. Process. Syst.* **32**, (2019).
- 860 41. Rusch, T. *et al.* A Neuro-Computational Characterization of Theory of Mind Processes during
861 Cooperative Interaction. 243–246 (2019). doi:10.32470/ccn.2019.1051-0
- 862 42. Rusch, T., Steixner-Kumar, S., Doshi, P., Spezio, M. & Gläscher, J. Theory of mind and decision
863 science: Towards a typology of tasks and computational models. *Neuropsychologia* **146**, 1–44
864 (2020).
- 865 43. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A Bayesian foundation for individual
866 learning under uncertainty. *Front. Hum. Neurosci.* **5**, 1–20 (2011).
- 867 44. Friston, K. *et al.* Active inference and epistemic value. *Cogn. Neurosci.* **6**, 187–224 (2015).
- 868 45. Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group
869 studies - Revisited. *Neuroimage* **84**, 971–985 (2013).
- 870 46. Rigoux, L. *et al.* Bayesian Model Selection for group studies. *Neuroimage* **46**, 1004–1017 (2009).
- 871 47. Glover, S. & Dixon, P. The role of predictability in cooperative and competitive joint action. *J.*

- 872 *Exp. Psychol. Hum. Percept. Perform.* **43**, 644–650 (2017).
- 873 48. Hampton, A. N., Bossaerts, P. & O’Doherty, J. P. The Role of the Ventromedial Prefrontal Cortex
874 in Abstract State-Based Inference during Decision Making in Humans. *J. Neurosci.* **26**, 8360–
875 8367 (2006).
- 876 49. Schuck, N. W., Cai, M. B., Wilson, R. C. & Niv, Y. Human Orbitofrontal Cortex Represents a
877 Cognitive Map of State Space. *Neuron* **91**, 1402–1412 (2016).
- 878 50. Niv, Y. Learning task-state representations. *Nat. Neurosci.* **22**, 1544–1553 (2019).
- 879 51. Baram, A. B., Muller, T. H., Nili, H., Garvert, M. M. & Behrens, T. E. J. Entorhinal and
880 ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning
881 problems. *Neuron* **109**, 713-723.e7 (2021).
- 882 52. Rilling, J. K., Gutman, D. A., Zeh, T. R. & Pagnoni, G. A Neural Basis for Social Cooperation.
883 *Neuron* **35**, 395–405 (2002).
- 884 53. Suzuki, S. *et al.* Learning to Simulate Others’ Decisions. *Neuron* **74**, 1125–1137 (2012).
- 885 54. Vickery, T. J., Kleinman, M. R., Chun, M. M. & Lee, D. Opponent identity influences value learning
886 in simple games. *J. Neurosci.* **35**, 11133–11143 (2015).
- 887 55. Griessinger, T. & Coricelli, G. The neuroeconomics of strategic interaction. *Curr. Opin. Behav.*
888 *Sci.* **3**, 73–79 (2015).
- 889 56. Boorman, E. D., O’Doherty, J. P., Adolphs, R. & Rangel, A. The behavioral and neural
890 mechanisms underlying the tracking of expertise. *Neuron* **80**, 1558–1571 (2013).
- 891 57. Behrens, T. E. J., Hunt, L. T. & Rushworth, M. F. S. The computation of social behavior. *Science*
892 *(80-)*. **324**, 1160–1164 (2009).
- 893 58. Domenech, P., Rheims, S. & Koechlin, E. Neural mechanisms resolving exploitation-exploration
894 dilemmas in the medial prefrontal cortex. *Science (80-)*. **369**, (2020).
- 895 59. Donoso, M., Collins, A. G. E. & Koechlin, E. Foundations of human reasoning in the prefrontal
896 cortex. *Science (80-)*. **344**, 1481–1486 (2014).
- 897 60. Kim, D., Park, G. Y., O’Doherty, J. P. & Lee, S. W. Task complexity interacts with state-space
898 uncertainty in the arbitration between model-based and model-free learning. *Nat. Commun.* **10**,
899 (2019).
- 900 61. Chung, Y. S., Barch, D. & Strube, M. A meta-analysis of mentalizing impairments in adults with
901 schizophrenia and autism spectrum disorder. *Schizophr. Bull.* **40**, 602–616 (2014).
- 902 62. Hampton, A. N., Bossaerts, P. & O’Doherty, J. P. Neural correlates of mentalizing-related

- 903 computations during strategic interactions in humans. *Proc. Natl. Acad. Sci.* **105**, 6741–6746
904 (2008).
- 905 63. Mathys, C. D. *et al.* Uncertainty in perception and the Hierarchical Gaussian Filter. *Front. Hum.*
906 *Neurosci.* **8**, 1–24 (2014).
- 907 64. Mazaika, P. K., Hoefft, F., Glover, G. H. & Reiss, A. L. Methods and Software for fMRI Analysis
908 of Clinical Subjects. *Neuroimage* **47**, S58 (2009).
- 909

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SIRPJCD.pdf](#)