

RESEARCH ARTICLE

Computational joint action: Dynamical models to understand the development of joint coordination

Cecilia De Vicariis^{1*}, Vinil T. Chackochan^{1,2}, Laura Bandini¹, Eleonora Ravaschio¹, Vittorio Sanguineti¹

1 Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Genoa, Italy, **2** Odstock Medical Ltd, Salisbury, United Kingdom

* cecilia.devicariis@unige.it



OPEN ACCESS

Citation: De Vicariis C, Chackochan VT, Bandini L, Ravaschio E, Sanguineti V (2024) Computational joint action: Dynamical models to understand the development of joint coordination. *PLoS Comput Biol* 20(10): e1011948. <https://doi.org/10.1371/journal.pcbi.1011948>

Editor: Lusha Zhu, Peking University, CHINA

Received: February 27, 2024

Accepted: September 24, 2024

Published: October 22, 2024

Copyright: © 2024 De Vicariis et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting information](#) files. All code written in support of this publication is publicly available at <https://github.com/SanguinetiLab/JAModel.git>.

Funding: This work was supported by the Italian Ministry of University and Research under the complementary actions to the NRRP (“Fit4MedRob - Fit for Medical Robotics” Grant # PNC0000007 to CDV) and under the Extended Partnership Neuroscience and Neuropharmacology (PE12) (MNESYS - A Multiscale integrated approach to the

Abstract

Coordinating with others is part of our everyday experience. Previous studies using sensorimotor coordination games suggest that human dyads develop coordination strategies that can be interpreted as Nash equilibria. However, if the players are uncertain about what their partner is doing, they develop coordination strategies which are robust to the actual partner’s actions. This has suggested that humans select their actions based on an explicit prediction of what the partner will be doing—a partner model—which is probabilistic by nature. However, the mechanisms underlying the development of a joint coordination over repeated trials remain unknown. Very much like sensorimotor adaptation of individuals to external perturbations (eg force fields or visual rotations), dynamical models may help to understand how joint coordination develops over repeated trials. Here we present a general computational model—based on game theory and Bayesian estimation—designed to understand the mechanisms underlying the development of a joint coordination over repeated trials. Joint tasks are modeled as quadratic games, where each participant’s task is expressed as a quadratic cost function. Each participant predicts their partner’s next move (partner model) by optimally combining predictions and sensory observations, and selects their actions through a stochastic optimization of its expected cost, given the partner model. The model parameters include perceptual uncertainty (sensory noise), partner representation (retention rate and internale noise), uncertainty in action selection and its rate of decay (which can be interpreted as the action’s learning rate). The model can be used in two ways: (i) to simulate interactive behaviors, thus helping to make specific predictions in the context of a given joint action scenario; and (ii) to analyze the action time series in actual experiments, thus providing quantitative metrics that describe individual behaviors during an actual joint action. We demonstrate the model in a variety of joint action scenarios. In a sensorimotor version of the Stag Hunt game, the model predicts that different representations of the partner lead to different Nash equilibria. In a joint two via-point (2-VP) reaching task, in which the actions consist of complex trajectories, the model captures well the observed temporal evolution of performance. For this task we also estimated the model parameters from experimental observations, which provided a comprehensive characterization of individual dyad participants. Computational models of joint action may help identifying the

study of the nervous system in health and disease to VS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

factors preventing or facilitating the development of coordination. They can be used in clinical settings, to interpret the observed behaviors in individuals with impaired interaction capabilities. They may also provide a theoretical basis to devise artificial agents that establish forms of coordination that facilitate neuromotor recovery.

Author summary

Acting together (joint action) is part of everyday experience. But, how do we learn to coordinate with others and collaborate? Using a combination of experiments and computational models we show that through multiple repetitions of the same joint task we select the action which represents the ‘best response’ to what we believe our opponent will do. Such a belief about our partner (partner model) is developed gradually, by optimally combining prior assumptions (how repeatable or how erratic our opponent behaves) with sensory information about our opponent’s past actions. Rooted in game theory and Bayesian estimation, the model accounts for the development of the mutual ‘trust’ among partners which is essential for establishing a mutually advantageous collaboration, and explains how we combine decisions and movements in complex coordination scenarios. The model can be used as a generative tool, to simulate the development of coordination in a specific joint action scenario, and as an analytic tool to characterize the individual traits or defects in the ability to establish collaborations.

Introduction

Many of our everyday activities take place in social settings and are coordinated with other people. Even seemingly simple interactions, like a pair of workers sawing timber with back and forth movements, a couple executing complex steps on a dance floor, two children playing badminton, or providing physical therapy to a patient, require that two individual minds are somehow connected and their bodies coordinated [1]. The characteristic feature of these interactions is that participants influence each others’ behavior through coupled sensorimotor exchanges within continuous action spaces, continuously in time, and possibly over repeated trials. Many of these interactions require coordination, which is manifested by physical and cognitive responses related to explicit or implicit knowledge of the interacting partner.

Most studies have been focusing on situations in which interacting agents have a shared goal and develop collective representations of such goal and the interactive setting [2–5]. However, ‘it is actually nonimitative or complementary responses that are especially critical to joint action’ [6]. In these situations, the process of understanding our partners and selecting our own strategy is more nuanced. In this perspective, this study focuses on situations in which participants have distinct goals, so that coordination is not set in advance but emerges gradually from the interaction.

When the two humans are physically coupled through a tool [7] or a pair of robot devices [8], haptics is a major source of information about their partner. These scenarios have been particularly useful for quantifying various aspects of joint action performance because dual haptic or ‘dyadic’ interfaces can be programmed to implement a variety of interaction modalities and allow to experimentally manipulate all aspects of the interaction [8–11].

During the last two decades, discrete-time dynamical models have been used to investigate sensorimotor learning and adaptation. These models have helped to clarify the underlying

mechanisms: the role of the sensory prediction error [12, 13]; the interplay of implicit and explicit adaptation mechanisms [14]; and the evidence of multiple concurrent adaptation processes across different time scales [15]. The same modeling framework has been used to assess how adaptation is altered in several neurological conditions, such as cerebellar degeneration [16] and multiple sclerosis [17], and to model neuromotor recovery through robot-assisted exercise [18].

A similar approach can be used to gain more insight on how we develop coordination strategies while we interact with our peers. Few studies have addressed the perceptual and control mechanisms of joint action [9–11, 19]. Dynamical models rely on optimality principles to account for perception and action selection within a single trial. Perception is believed to optimally combine prior beliefs and the available observations to minimize the prediction uncertainty [20, 21]. These same Bayesian mechanisms likely mediate the prediction of the actions of our partners [11, 19]. Partial information about the partner likely affects the development of a joint coordination. Action selection in joint coordination can be interpreted in terms of game theory. Several studies [9, 11, 22] have reported that the sensorimotor behavior of physically coupled subjects converges toward a Nash equilibrium—a situation in which none of the players can unilaterally improve their payoff [23]. However, the vast majority of previous studies addressing joint action within a game theoretic framework only focus on equilibrium situations [9, 22]. Very few studies [11, 24, 25] have addressed the way joint coordination is negotiated and learned in scenarios that involve movements. One simple learning strategy when the players play repeatedly the game is that at every round each player determines their best response based on their beliefs about how their opponents will play (fictitious play, FP)—see [26, 27]. FP has been implied as a candidate mechanism for the development of joint decision-making [24] or joint action in humans [11], at least in situations where the players use stationary strategies. Reinforcement learning mechanisms have been also implicated in situations where the players have knowledge about their task [25].

Here we present a general modeling framework to describe the development of coordination strategies in joint action. The model addresses a variety of interactive scenarios and can be used both as a simulation and an analytical tool. We first demonstrate the model with a toy problem—a sensorimotor version of a classical game with multiple Nash equilibria (Stag Hunt). We then focus on joint sensorimotor coordination games [11], in which participants are mechanically coupled, have a shared final target, and are instructed to cross a via-point (VP, different for each player) while minimizing their interaction force. We estimate the model parameters from the experimental data (action time series over trials). We show that the model correctly captures the different experimental conditions and correctly reproduces, for each condition, the temporal evolution of the coordination strategies.

Model

We specifically focus on situations in which the players know their own goal, but not their partner's, and they have imperfect information about their partner's action. We also assume that the agents develop a coordination over multiple rounds of the same game. Many joint action studies fall within this scenario. We will later discuss how these assumptions can be relaxed in order to deal with more general situations.

The overall model is depicted in Fig 1. At each trial t , the i -th player selects their own action, $u_i(t)$, on the basis of the partner's expected move (partner model), $x_i(t) = \hat{u}_{-i}(t)$ —we use $-i$ to indicate the player other than i . Once the actions have been carried out, through their sensory systems both players collect information $y_i(t)$ about their own and their partner's actions. This

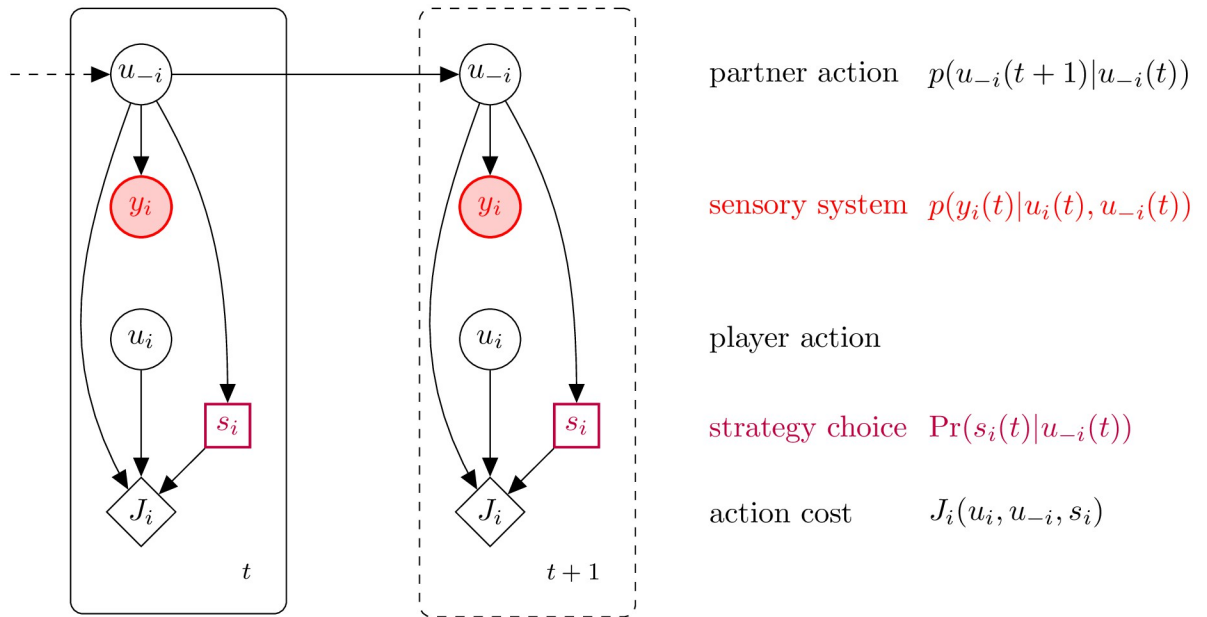


Fig 1. Probabilistic graphical model (influence diagram) of a single joint action player, that accounts for multiple strategy choices. At each time step, Player i receives information about partner action (u_{-i}) through their sensory system (y_i). Action selection involves the selection of a proper strategy (discrete variable s_i) and minimizing a cost function (J_i) which depends on both own (u_i) and partner's action (u_{-i}).

<https://doi.org/10.1371/journal.pcbi.1011948.g001>

information is then used to update their partner model, which will be accounted for when selecting the action on the next trial.

Joint tasks as mixtures of quadratic games

Individual motor tasks are often described in terms of a quadratic cost function which accounts for the trade-off between error and effort [28]. For the i -th player, the model specifies the task as a quadratic cost function, which depends on both players' actions u_1 and u_2 :

$$J_i(u_1, u_2) = \frac{1}{2} [u_1^T R_{11}^i u_1 + 2 u_1^T R_{12}^i u_2 + u_2^T R_{22}^i u_2] + r_1^{iT} \cdot u_1 + r_2^{iT} \cdot u_2 + z^i \quad (1)$$

Static games with infinite action space and quadratic costs are known as quadratic games [29]. The quadratic game formalism naturally extends individual motor control to joint action scenarios.

If all $R_{j,k}^i$ are symmetric and positive definite, a quadratic game has a single Nash equilibrium [29] which is calculated as:

$$u^* = -R^{-1} \cdot r \quad (2)$$

where $u^* = [u_1^T \ u_2^T]^T$, $r = [r_1^T \ r_2^T]^T$ and:

$$R = \begin{bmatrix} R_{11}^1 & R_{12}^1 \\ R_{21}^2 & R_{22}^2 \end{bmatrix} \quad (3)$$

Many joint action scenarios involve selecting an action among a set of alternatives—for instance, reaching one among several possible targets. To extend the quadratic game scenario

to these situations, we assume that the cost incurred by player i depends on a decision variable $s_i \in \{S_i^1 \dots S_i^M\}$, which specifies the decision between one of the M options that the i -th player has available. Therefore, Eq 1 becomes:

$$J_i(u_1, u_2, s_i) = \frac{1}{2} [u_1^T \cdot R_{11}^i(s_i) \cdot u_1 + 2 u_1^T \cdot R_{12}^i(s_i) \cdot u_2 + u_2^T \cdot R_{22}^i(s_i) \cdot u_2] + r_1^i(s_i)^T \cdot u_1 + r_2^i(s_i)^T \cdot u_2 + z^i(s_i) \tag{4}$$

Sensory system

At every trial t , the sensory system of each player provides an estimate of their partner’s action:

$$y_i(t) = H^i \cdot u_{-i} + v_i(t) \tag{5}$$

where $v_i(t) \sim N(0, \Sigma_y^i)$ is the sensory (measurement) noise of the i -th player, with covariance Σ_y^i . Parameter H^i reflects the structure of the sensory system and relates partner action to its sensory consequences.

Partner model

We assume that at every trial, t , each player predicts their partner’s next action by optimally combining predictions and actual observations—provided by the sensory system above—into a ‘partner model’. The i -th player’s prior belief about partner action u_{-i} can be summarized as:

$$u_{-i}(t + 1) = A^i \cdot u_{-i}(t) + w_i(t) \tag{6}$$

where $w_i(t) \sim N(0, \Sigma_x^i)$ is a process, or internal, noise and A^i is a retention (memory) parameter. Parameters A^i and Σ_x^i reflect the player’s prior belief on partner behavior—respectively, the extent to which partner behavior is stationary, or is erratic across trials. The covariance of the internal noise, Σ_x^i , can be interpreted as the player’s belief in the partner’s predictability. In particular, $A^i = 0$ denotes a memoryless partner behavior and $\Sigma_x^i = 0$ denotes a purely deterministic partner behavior.

Action selection and strategy choice

The optimal action of the i -th player at time t , $u_i(t)$ —and the optimal strategy $s_i(t)$ if multiple options exist—minimizes the expected cost $\bar{J}_i(u_i, s_i) = E\{J_i(u_i, u_{-i}, s_i)\}$, where the expectation is calculated over the partner’s action, u_{-i} :

$$E\{J_i(u_i(t), s_i(t), u_{-i}(t))\} = \int J_i(u_i, s_i, u_{-i}) \cdot p(u_{-i}) du_{-i} \tag{7}$$

However, the i -th player only knows their partner’s action through the history of sensory observations $D_i(t) = \{y_i(1) \dots y_i(t)\}$, so that the average is made over $p(u_{-i}|D_i(t))$.

The quantity $p(u_{-i}|D_i(t))$ results from the Kalman integration algorithm. At each time step t , the i -th player has a prior estimate (prediction) of partner action, defined by its mean $x_i(t)$ and covariance $P_i(t)$. These quantities are calculated by incorporating the sensory outcome of the actions performed at the previous iteration $y_i(t)$:

$$x_i(t + 1) = A^i \{x_i(t) + K_i(t) \cdot [y_i(t) - H^i \cdot x_i(t)]\} \tag{8}$$

$$P_i(t + 1) = A^i [I - K_i(t)H^i] P_i(t) A^{iT} + \Sigma_x^i$$

where the Kalman gain $K_i(t) = P_i(t)H^{iT}[H^iP_i(t)H^{iT} + \Sigma_y^i]^{-1}$ is calculated iteratively. The portion in $\{ \dots \}$ in Eq 8 optimally combines the action prediction at time t , $x_i(t)$ with the sensory information obtained after both players have performed their actions.

At each trial the Bayesian inference (Kalman filter algorithm) provides the expected value $x_i(t)$ of the partner’s action and its covariance matrix $P_i(t)$, so that the expected cost—see Eq 7 can be expressed in terms of these quantities and turns out to be a quadratic form in u_i :

$$\bar{J}_i(u_i, s_i) = E\{J_i(u_i, u_{-i}, s_i)\} = \frac{1}{2} [u_i - u_i^*(s_i, x_i)]^T \cdot R_{ii}^i(s_i) \cdot [u_i - u_i^*(s_i, x_i)] + q(s_i, x_i, P_i) \quad (9)$$

whose parameters depend on x_i and s_i —see the S1 File for complete derivation.

To account for variability in action selection—i.e. players not behaving in a deterministic way, or having an imperfect knowledge of their cost function—we assume that the selection of both strategy and action is in fact a stochastic process. The expected cost can be interpreted as an energy term so that the joint probability of u_i and s_i given the partner model can be expressed as a Boltzmann probability density function:

$$p(u_i, s_i | x_i) = p(u_i | s_i, x_i) \cdot \Pr(s_i | x_i) \propto e^{-\frac{1}{\lambda^i} \bar{J}_i(u_i, s_i)} \quad (10)$$

where λ^i reflects action selection ‘temperature’, with greater λ^i leading to greater randomness.

In the special case of $R_{ii}^i(s_i)$ being all positive definite, the action probability given the partner model x_i and the strategy $s_i = S_k^i - p(u_i | s_i = S_k^i, x_i)$ —has a multivariate Gaussian distribution, with mean $u_i^*(S_k^i, x_i)$ and covariance $\Sigma_u^i(S_k^i) = \lambda^i \cdot R_{ii}^i(S_k^i)^{-1}$. The probability of selecting strategy $s_i = S_k^i$ given the partner model x_i also depends on partner model and game structure:

$$\Pr(s_i = S_k^i | x_i) = \frac{|R_{ii}^i(S_k^i)|^{-1/2} \cdot e^{-\frac{1}{\lambda^i} q(S_k^i, x_i)}}{\sum_{j=1}^M |R_{ii}^i(S_j^i)|^{-1/2} \cdot e^{-\frac{1}{\lambda^i} q(S_j^i, x_i)}} \quad (11)$$

See the S1 File for complete derivation. As a consequence, when the $R_{ii}^i(s_i)$ are all positive definite, the action u_i given the partner model x_i has a Gaussian Mixture distribution.

The proposed action generation mechanism naturally satisfies the minimum intervention principle [30]: in a redundant task, the action cost does not depend on the task-irrelevant movement features and, conversely, is very sensitive to task-relevant features, so that action selection variability is modulated by task relevance. Further, the action generation model can be seen as an extension of the quantal response equilibria (QRE) model [31] to continuous games. As in the QRE, strategy choice depends on the task and is modulated by partner model uncertainty—the dependence of $q(s_i, x_i)$ on P_i . Likewise, probabilistic action selection reflects the task requirements and is modulated by the temperature parameter. In many learning scenarios, the selected actions are expected to gradually become more repeatable from trial to trial. To allow the ‘temperature’ parameter $\lambda^i(t)$ to decrease over trials, in the model we specify the initial temperature, $\lambda^i(1) = \lambda_1^i$, and a decay rate $a^i < 1$, so that

$$\lambda^i(t + 1) = a^i \cdot \lambda^i(t) \quad (12)$$

In the above derivation, we made the critical assumption that the $R_{ii}^i(s_i)$ are positive definite. In this case, the probability distribution of actions—is a Gaussian or a Gaussian Mixture. However, this condition may not be met in many types of scenarios, e.g. competitive games where the gain of one player corresponds to the loss of the other; in this situation the matrix $R_{ii}^i(s_i)$ may no longer be positive definite for some of the players. Nevertheless, if the action space is limited, Eq 10 still defines a valid action probability density function, though no longer a mixture of Gaussians.

At every round, each agent determines their best response to the empirical distribution of their partner's action. This process, known as fictitious play (FP) [27], is in many cases sufficient to learn a coordination strategy.

One player's behavior is specified in terms of seven parameters. The sensory noise covariance (Σ_y^i) specifies the reliability of the player's sensory system. A Σ_y^i close to 0 indicates that the agent has reliable information about partner actions. Greater values of Σ_y^i indicate unreliable sensory information and may possibly suggest diminished sensory capabilities.

The initial partner model is specified by the mean value (μ^i) and covariance (P_0^i) of a Gaussian probability distribution, which constitutes the prior belief about the partner's action. The retention rate (A^i) and internal noise covariance (Σ_x^i) specify the way the partner model evolves with time as the result of a Kalman estimation process. Therefore, greater internal noise (Σ_x^i) and low retention rate (A^i) indicate a prior belief that the partner behaves erratically. Conversely, low internal noise and greater retention denote a prior belief that the partner behaves in a stationary way.

Over trials, the belief regarding the partner action is optimally combined with the sensory information to update the partner model.

Finally, initial temperature (λ_1^i) and temperature decay rate (a^i) specify the degree of randomness in action selection and how it evolves with time. Low values of initial temperature (λ_1^i) and higher temperature decay (a^i) values indicate that the agent is more deterministic and converges more quickly to deterministic responses. Conversely, they keep a high level of exploration or randomness.

The model is general enough to simulate the trial-by-trial evolution of a large variety of joint action scenarios which involve perception and action. Further, the model is simple enough to be used to analyze joint action experiments. Given a time series of measured players' joint actions, it is possible to estimate the model parameters that best describe each player in the dyad. The supplementary notes in [S1 File](#) describe the maximum likelihood procedure to estimate the model parameters and report model fitting results.

Results

We demonstrate the model in a variety of interaction scenarios, in which two players engage in a sequence of interactions. We specifically focus on how convergence to a joint strategy is determined by the information available about their partner. In all scenarios, we show that the model generates plausible action sequences. When experimental results are available, they are compared with the model predictions.

Spatial Stag Hunt

We first demonstrate the model behavior in a toy (simulated) experiment, inspired by the classical Stag Hunt game. Two hunters have the option of jointly hunting a stag (S), or independently pursuing a rabbit (R). Hunting a stag results in a greater payoff, but can only be done if both partners cooperate. Catching a rabbit does not require partner cooperation, but results in a smaller payoff. This game has two Nash equilibria: both hunters catching a rabbit (R-R), and both hunters catching the stag (S-S). The R-R strategy has a lower payoff than the S-S strategy. The original version of the game assumes a discrete action space (R,S). Yoshida et al. (2008) [32] formulated a spatial version of the game, in which two players are required to move their hand within a one-dimensional workspace toward either a rabbit or a stag, located at fixed positions u_R and u_S . The participants are supposed to act simultaneously, without knowing the

decision of the partner. The payoff for catching the stag can only be achieved when both players are in u_S , whereas the rabbit payoff does not depend on the other agent's action.

Our modeling framework explicitly focuses on joint actions that involve movements and perception. This requires some extra assumptions about the physical arrangement of the simulated experiment. Specifically, we assumed that at each trial each player controls a cursor along a line representing the action space. The locations of both the rabbit and stag are fixed and known to both players. We additionally assumed that at the end of each trial, each player is provided with imperfect information on their partner's location. In a physical experiment, this could be done by a blurred display (random dots). The above assumptions are summarized in Fig 2 (left).

To account for the two Nash equilibria of this game, we introduce a discrete variable s_i which denotes whether player i is pursuing the 'rabbit' ($s_i = 1$) or the 'stag' ($s_i = 0$). Therefore, the cost function for player i is:

$$J_i(u_i, u_{-i}, s_i) = s_i \cdot \left[\frac{1}{2} w \cdot |u_i - u_R|^2 + z_R \right] + (1 - s_i) \cdot \left[\frac{1}{2} w \cdot (|u_i - u_S|^2 + |u_{-i} - u_S|^2) + z_S \right] \quad (13)$$

The model parameters w , z_R and z_S are calculated on the basis of the cost matrix of Fig 2 (right). Details on the parameters computation are provided in the Methods section. We additionally assume that at the end of each trial, both participants are displayed the cost incurred according to Eq 13. Both are instructed to keep their cost as low as possible.

We simulated the model by varying the reliability of sensory information ('low' or 'high' Σ_y^i) and the belief of partner predictability ('low' or 'high' Σ_x^i). The simulation results are summarized in Fig 2. Convergence to the S-S strategy critically requires that the partner is believed to be predictable (low Σ_x^i), and reliable sensory information (low Σ_y^i) about the partner. Conversely, when partner predictability is low, irrespective of Σ_y^i the R-R strategy is preferred. When sensory information is less reliable, the players' behaviors become more erratic and tend to fluctuate between SS and RR. These observations are confirmed by statistical analysis. We used a two-way ANOVA to compare the four groups of simulations. We used sensory and internal noise as factors. We found a significant effect of sensory noise— $F(28, 1) = 98.88, p < 0.001$ —and an even stronger effect of internal noise— $F(28, 1) = 2428.63, p < 0.001$. The sensory and internal noise interaction is also significant— $F(28, 1) = 98.88, p < 0.001$.

The model suggests that convergence to the S-S Nash equilibrium requires two conditions: (i) reliable sensory information (low Σ_y^i) and (ii) a belief in partner predictability (low Σ_x^i). If the second condition is not satisfied, the partner model would only rely on sensory information, the actions would be purely reactive. This would prevent the build-up of trust which is necessary to establish a robust collaboration.

As a cross-check, we identified the model parameters from the simulated actions time series. To robustly recover model parameters, the action time series must exhibit some temporal variability, a situation which only occurred in the low Σ_x^i and high Σ_y^i condition; see Fig 2d. For this condition we indeed obtained a high correlation coefficient ($R^2 > 0.95$) for the estimated actions time series and small absolute errors between the simulated and estimated model parameters—details are reported in the S1 File.

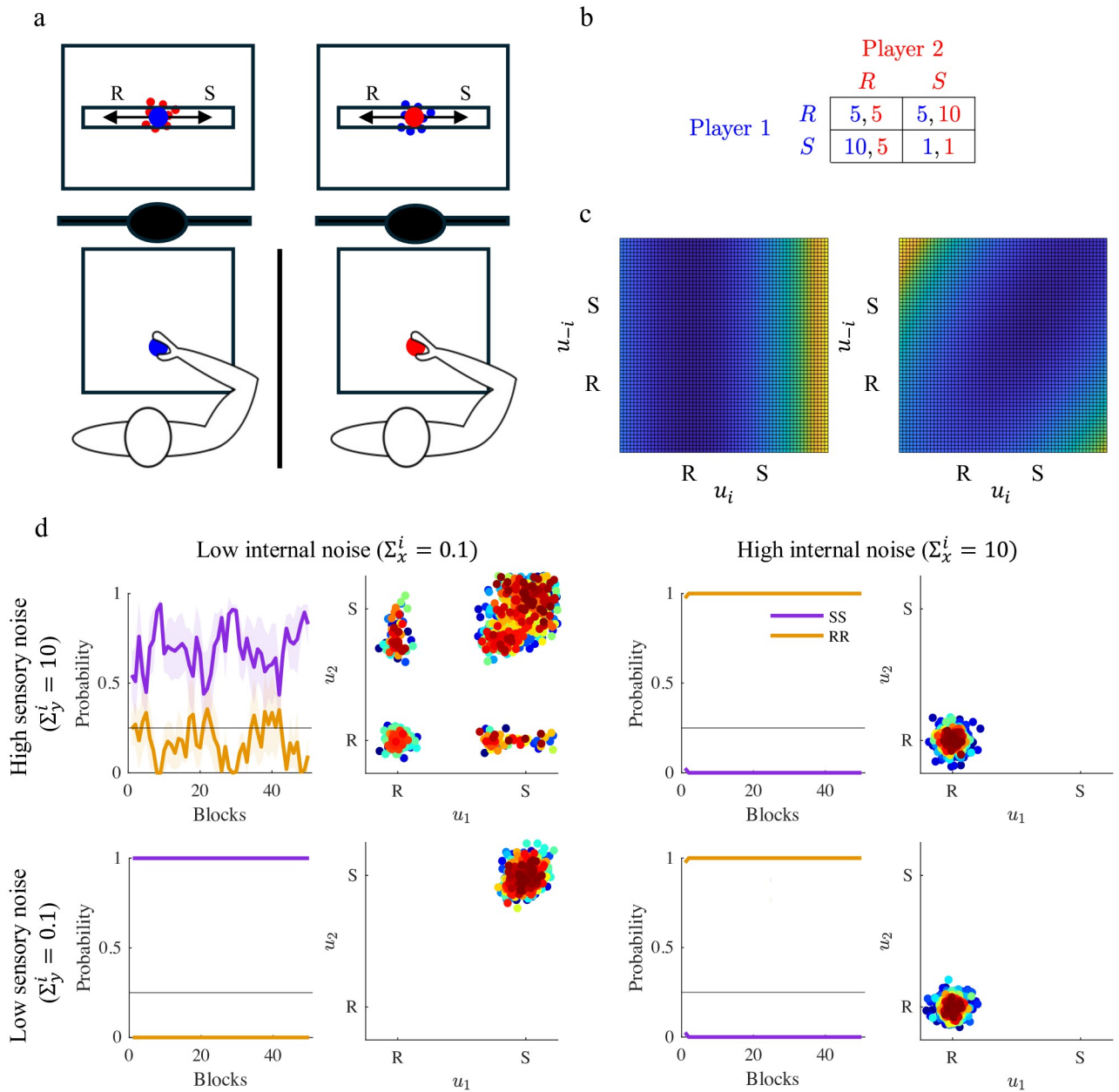


Fig 2. Stag Hunt game. a. Simulation concept. Each player knows the locations of both rabbit (R) and stag (S). They can see their actions and have imperfect information about their partner's actions (blurred circles). b. cost matrix used in the simulations, in which the cost can be interpreted as a reduced payoff. c. Costs related to the two equilibria are represented in the action space. Cold colors indicate lower costs, and hot colors indicate higher costs. On the left, whichever action the partner selects it is safe for player i to go for the Rabbit solution. On the right, the two players minimize their costs by selecting both the Stag. d. Simulated Stag Hunt game for different combinations of sensory noise (Σ_y^i) and internal noise or partner predictability (Σ_x^i). Temporal evolution of the probabilities of selecting the SS (blue), the RR (orange). The curves represent game performance at population level, where probabilities (mean \pm SE) have been computed over epochs (1 epoch = 40 trials) and averaged over multiple simulated dyads. The horizontal line indicates the chance level ($p = 0.25$). Bottom: scatter plots of the joint actions (u_1, u_2) for one representative dyad for each condition.

<https://doi.org/10.1371/journal.pcbi.1011948.g002>

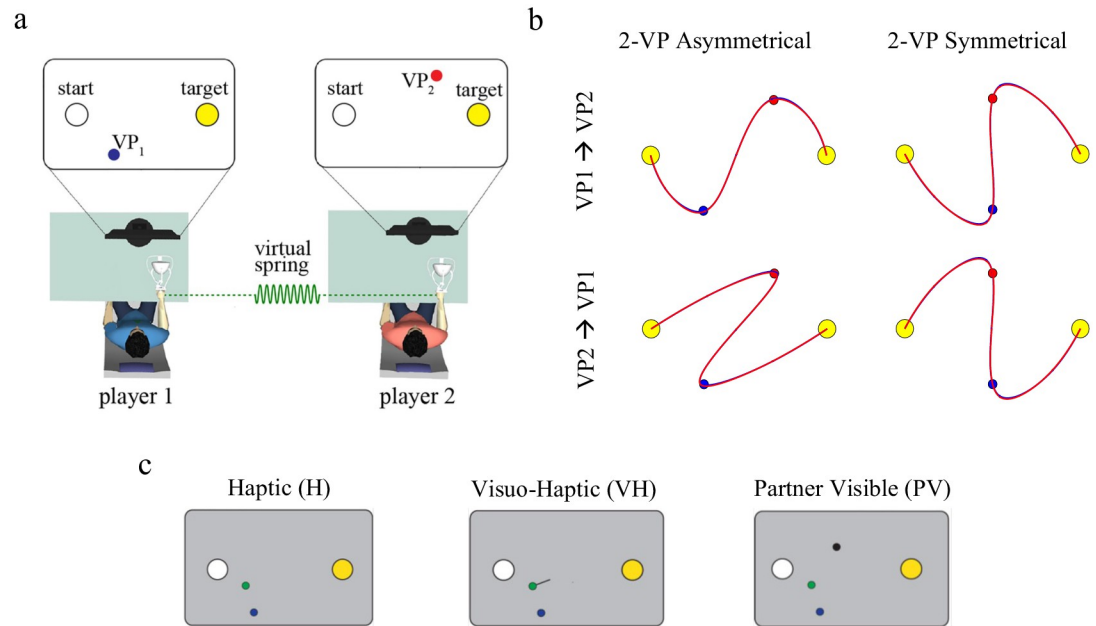


Fig 3. The 2-VP task. a. Experimental apparatus and task—modified from [11]. The partners were connected through a virtual spring. b. Nash equilibria with asymmetric and symmetric via-points locations. c. In different experimental groups, information about partner actions was provided haptically, through the interaction force alone (Haptic group, H); by additionally displaying the interaction force vector on the screen (Visuo-Haptic group, VH), or by also displaying the partner’s cursor (Partner Visible, PV).

<https://doi.org/10.1371/journal.pcbi.1011948.g003>

Two via-point game

In the two via-point (2-VP) task [11], the players are mechanically connected through a virtual spring. They perform reaching movements from the same start and end position, by crossing a via-point (VP) which is different for the two players. Each player can only see their own VP, but not their partner’s. Both are instructed to keep the spring force—and therefore the inter-player distance—as low as possible during movement; see Fig 3a.

The joint task has two Nash equilibria, in which both players follow the same path and cross both via-points in the same order, either $VP_1 \rightarrow VP_2$ or $VP_2 \rightarrow VP_1$. Depending on the VP locations, the two Nash equilibria may substantially differ in terms of path length and effort required. When the VPs are arranged along the midline of the start-end segment, the two Nash equilibria have exactly the same cost—see Fig 3b.

In terms of the model, the 2-VP task can be described by the following cost function (one per player):

$$J_i(p_i, p_{-i}, s_i) = \frac{1}{2} w_1 \|p_i(\tau_i(s_i)) - VP_i\|^2 + \frac{1}{2T} \sum_{\tau=0}^{T-1} [w_2 \|p_i(\tau) - p_{-i}(\tau)\|^2 + r \|F_i(\tau)\|^2] \quad (14)$$

Where p_i and p_{-i} refer respectively to trajectories of the two players and where $\tau = 0$ is taken as the start time and $\tau = T$ as the final time (common to both players) and we assume that $p_i(0) = p_{start}$ and $p_i(T) = p_{end}$. Where p_{start} and p_{end} corresponds to start point and target locations. Time τ_i is the time at which the i -th player crosses their own via-point VP_i . The crossing time is determined by the decision variable s_i , which specifies whether the VP is crossed early ($s_i = E$), in the middle ($s_i = M$), or late in the movement ($s_i = L$). We make the simplifying

assumption that, across trials, the crossing time $\tau_i(s_i)$ only takes three fixed values. Body dynamics of the i -th player can be described as

$$m_i \cdot \ddot{p}_i(\tau) + b_i \cdot \dot{p}_i(\tau) + k_i \cdot [p_i(\tau) - p_{-i}(\tau)] = F_i(\tau) \quad (15)$$

where $F_i(\tau)$ is the motor command.

Given the short duration of one trial, here we make the simplifying assumption that the action of one player only affects their partner's in future trials, whereas within-trial corrections are neglectable.

Under this assumption, we use a parametric representation of the players' actions within each trial. In particular, we approximate the trajectory $p_i(\cdot)$ of player i at trial t with a polynomial spline, specified by M nodes. The action u_i at time t is specified by the vector of node coordinates: $u_i = [p_i(\tau_1) \dots p_i(\tau_M)]^T$, where $\tau_1 = 0$ and $\tau_M = T$, where T is the movement duration. With spline approximation, the whole trajectory is now entirely specified by a small number of parameters (the nodes' coordinates). In particular, the whole trajectory, its subsequent derivatives (velocity, acceleration), and the body's equations of motion can be expressed as linear functions of the spline coefficients. The cost function in Eq 14 reduces to a quadratic form in u_i , and u_{-i} , so that the dynamic optimization problem—finding the optimal action $u_i(\cdot)$ —reduces to a static one—finding the optimal u_i .

In different experimental groups, in a previous study [11] we varied the information provided about the ongoing partner actions, namely haptically, through the interaction force alone (Haptic group, H); by additionally displaying the interaction force vector on the screen (Visuo-Haptic group, VH) or by also displaying the partner's cursor (Partner Visible, PV); see Fig 3c. For the first two experimental conditions (H and VH), the players see their own hand and feel the interaction force, which depends on the locations of both hands. In the PV condition, the position of their partner's hand is also displayed on the screen. Therefore, all groups have equal information on their own hand (vision and haptics), whereas the uncertainty in perceiving their partner is lower in group PV and greater in group H.

Experiment 1: Asymmetric via-points. Looking at the experimental findings, all participants developed strategies that are consistent with the low-effort Nash equilibrium (crossing VP₁ first, then VP₂), corresponding to $s_1 = E$ and $s_2 = L$. The hand trajectories turned out to be increasingly close to the Nash equilibrium when information about the partner was more reliable (PV group). In the model, we set $\tau_i(E) = 0.32 T$, $\tau_i(M) = 0.5 T$ and $\tau_i(L) = 0.68 T$ where $T = 1$ is the total movement duration. We estimated the model parameters of all individual participants from the actual experimental results—hand trajectories over trials—of the original 2-VP study [11]. These parameters describe the learning behavior of each player within a dyad.

The results of the identification procedure are summarized in Fig 4. To address how well the movement data fit the model, for each player we compared the actual and predicted actions. For both groups of players, we obtained an average $R^2 = 0.86$ —see the S1 File for details.

The predicted hand trajectories are consistent with the experimental findings—see Fig 4b—and exhibit distinct features in the different experimental conditions. Given that the players' actions are represented as vectors of spline nodes, sensory and internal noises are assumed to have spheric covariances so that both Σ_x^i and Σ_y^i reduce to a single scalar. Indeed, we assume that noises affect equally all the nodes, therefore parameters reduce to scalar values. We found significant group effects in the sensory variance parameter Σ_y^i — $F(2, 27) = 10.89$, $p < 0.001$. Post-hoc analysis revealed that Σ_y^i is significantly lower in the PV group respect to the H group ($p < 0.001$) and in VH with respect to the H group ($p = 0.004$). All the remaining parameters exhibited no significant group effects. Fig 4c summarizes these effects. The group differences

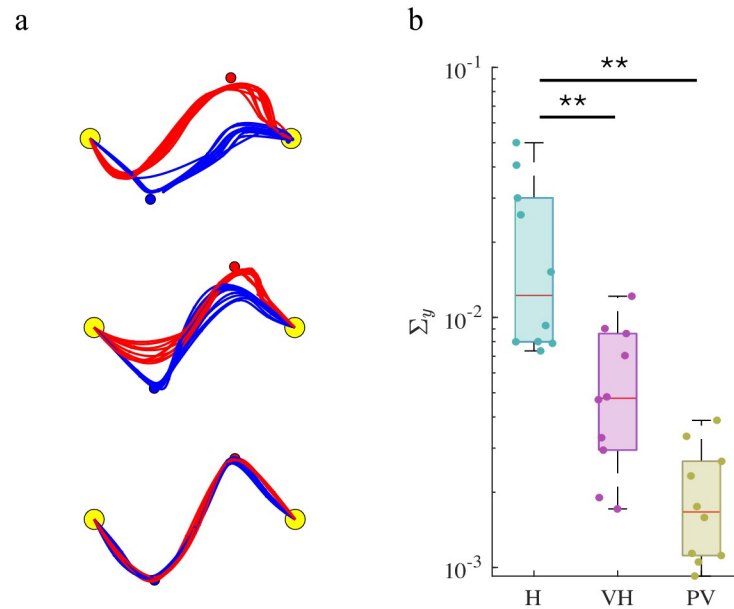


Fig 4. Fitting performances. a. Estimated trajectories in the identification procedure (last ten trials), for three representative dyads (one per group). b. Group differences in the estimated model parameters. The box plots display median, 25th and 75th percentiles of the estimated parameters, for H (light blue), VH (pink) and PV (yellow) groups. Asterisks indicate statistically significant differences (*: $p < 0.05$; **: $p < 0.01$).

<https://doi.org/10.1371/journal.pcbi.1011948.g004>

in the estimated Σ_y^i reflect the different experimental conditions, which are characterized by different amounts of sensory uncertainty—this suggests that the identification procedure has captured the latter from the movement data alone.

To further clarify the role of sensory noise in the development of joint coordination, we used the model to simulate the five dyads per each experimental group; we then compared the temporal evolution of the simulated movements in all three experimental conditions with the corresponding experimental results—see Fig 5.

The simulated trajectories qualitatively resemble the experimental results—see Fig 5a. We specifically looked at the temporal evolution of inter-player coordination in terms of the minimum distances of the i -th player's trajectory from the j -th via-point (MD_{ij} , with $i \neq j$)—see Fig 5b. Consistent with experimental findings [11], with lower sensory noise (PV group) the players converged toward a shared solution that is more similar—smaller MD_{ij} —to the theoretical Nash equilibrium—Fig 3b. The simulated MD_{ij} time series for the three groups turned out to be very similar to the experimental observations. In particular, the rate of decrease was faster when the information about the partner was more reliable.

One of the main findings of the original 2-VP study [11] was that, as the information about the partner became less reliable, the players converged to a joint strategy that was more robust to partner uncertainty, in which the players actively controlled the crossing of their own via-point and passively followed the partner toward the other via-point, thus alternating leader and follower roles within the same movement. This effect was quantified in terms of a metric—leadership index, LI_{ij} —calculated as the average power of the interaction force of the i -th player just before crossing the j -th via-point. A negative or positive LI_{ij} indicates that the i -th player behaves, respectively, as a 'leader' or as a 'follower'—see the Methods section. The experimental findings point to a dependence of the observed leadership indices on the reliability of sensory information. To understand the nature of this relation, in the experimental data

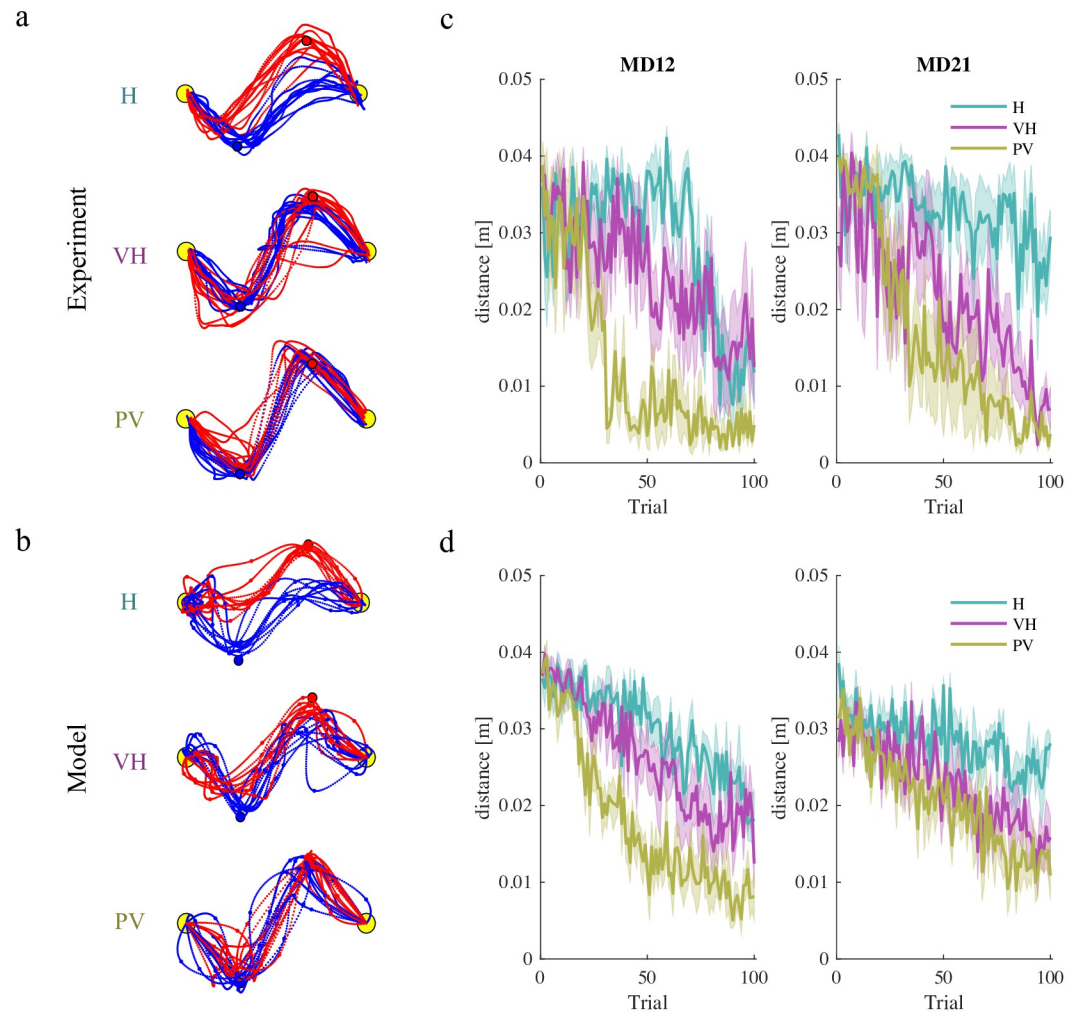


Fig 5. Experimental results (top) and simulations (bottom) of the 2VP task. a. Experimentally observed and b). simulated trajectories for selected dyads in the three experimental conditions (H, VH, PV) in the original 2-VP study [11]. c. Experimentally observed and d. simulated minimum distances of player 1 from VP₂ (MD₁₂) and player 2 from VP₁ (MD₂₁)—mean \pm SE over actual and simulated players—in all three conditions.

<https://doi.org/10.1371/journal.pcbi.1011948.g005>

we compared the leadership indices calculated for each subject at the end of the training phase and the estimated sensory noise covariance Σ_y^i for that same player. In addition, we used the model to simulate dyads playing the game in which we systematically varied Σ_y^i , while keeping the other parameters constant. The relation between Σ_y^i and the leadership indices is summarized in Fig 6a. In both experiments and simulations, when crossing their own via-pont both players behave as leaders (negative LI₁₁ and LI₂₂). Consistent with the simulations, the experimental data also suggest that in player 1 the amount of leadership (LI₁₂) is significantly correlated to noise covariance ($r = -0.632$, $p = 0.011$) whereas no such correlation is observed in player 2 (LI₂₁). When considering the crossing of the partners' via-points, both players exhibit a clear 'follower' behavior (positive LI₁₂ and LI₂₁), but only player 1 exhibits a significant correlation with the estimated noise covariance ($r = 0.5799$, $p = 0.023$). Again, the simulations—Fig 6a, red and blue lines—are consistent with this observation.

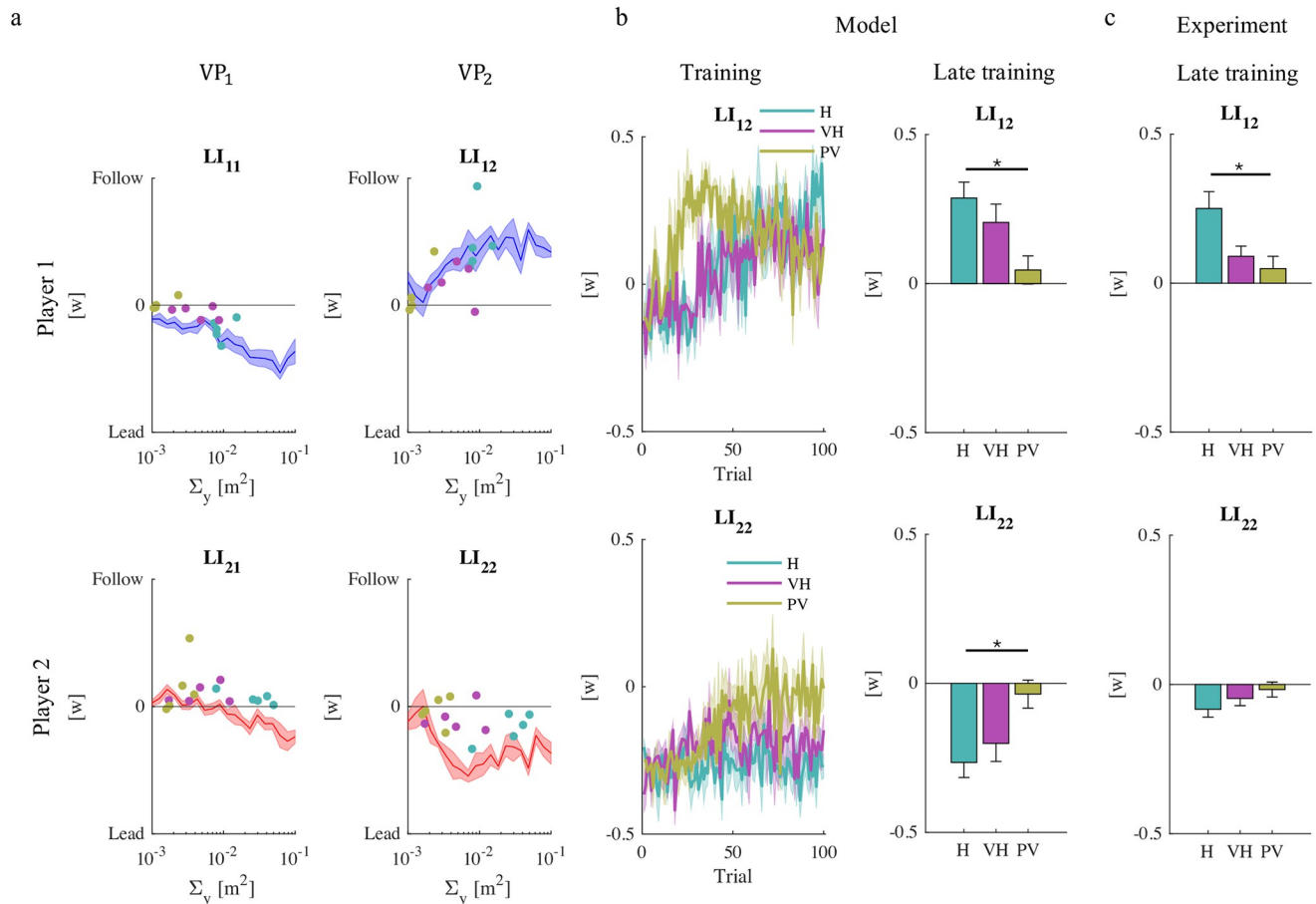


Fig 6. Sensory noise modulates the leadership index. a. Leadership indices as functions of the estimated sensory noise covariance Σ_y^i in experimental data (one data point per subject). Σ_y^i was systematically varied, and each Σ_y^i was simulated five times. The lines (mean \pm SE) denote the simulation results, for player 1 (blue) and player 2 (red). b. Simulated leadership indices over time (mean \pm SE); c. Simulated (left) and experimental (right) leadership indices for VP₂ in the three experimental groups, at the end of the training phase. Experimental data from [11].

<https://doi.org/10.1371/journal.pcbi.1011948.g006>

In summary, the experiments point at a stronger leader-follower role alternation in player 1 than in player 2. As a consequence, at VP₂ the two players exhibit a much greater role differentiation: player 2 as leader— $LI_{22} < 0$ —and player 1 as follower— $LI_{12} > 0$). The model does not fully account for the different behaviors in VP₁ and VP₂ and other factors, possibly related to within-trial interaction, may play a role. Early in the movement, player 2 may focus on achieving their subgoal, thus being less prone to follow player 1. Conversely, later in the movement, player 1—having reached their own subgoal—may be more willing to be led by player 2.

We then focused on the leadership indices for VP₂. Fig 6b depicts the simulated temporal evolution of the leadership indices for player 1 and player 2 just before the crossing of the second via-point—respectively, LI_{12} and LI_{22} . In the PV group, both indices quickly go to zero, indicating that roles tend to disappear as the interaction force disappears and the dyads get closer to the Nash equilibrium. In contrast, in the H and VH groups these indices change much more slowly and roles are preserved until the very end of the training phase. This model prediction is confirmed by statistical analysis of these quantities at the end of the training phase—significant group effect in LI_{22} ($F(2,14) = 5.03$, $p = 0.0259$) and in LI_{12} ($F(2,14) = 5.13$, $p = 0.0245$)—and is consistent with the experimental observations—see Fig 6c.

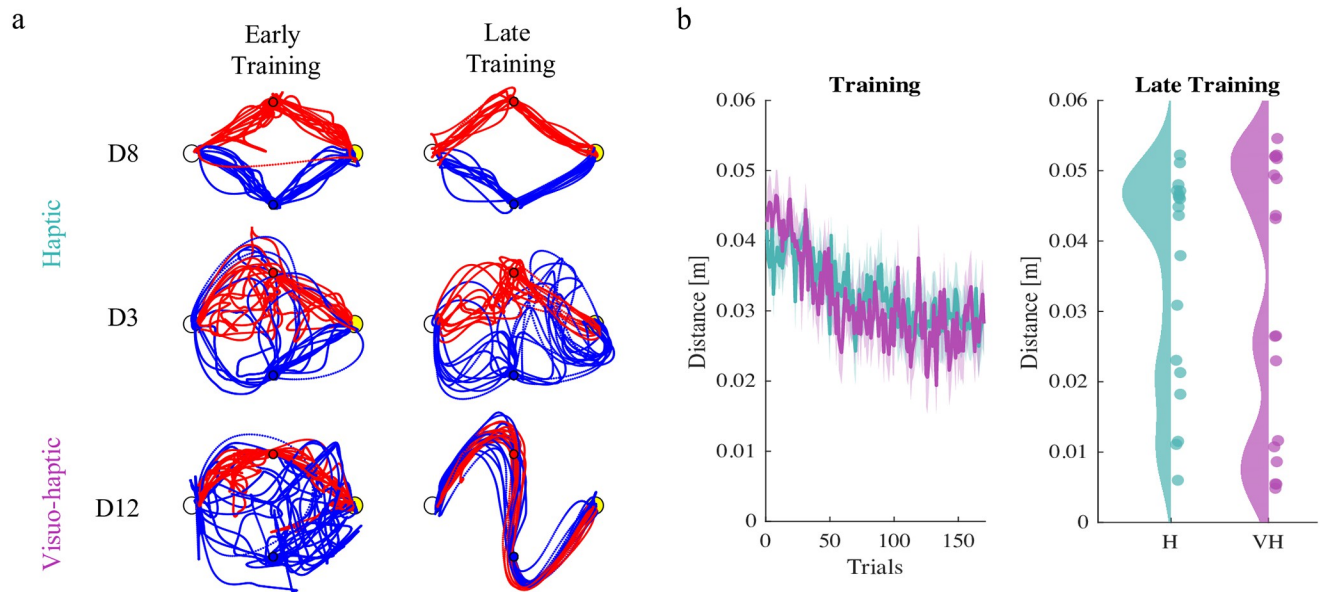


Fig 7. Coordination behavior in the symmetric 2-VP task. Dyads exhibit a variety of coordination behaviors. a. Typical dyad trajectories (from top to bottom: D8, D3, D12) at early and late training of the experiment. Player 1 and Player 2 are depicted in blue and red. These dyads exemplify typical behaviors: players ignoring each other (D8); cycling between opposite strategies (D3); and converging to a coordination (D12). b. Minimum distance of the two players from partner's via-points. We report population average and standard error over time for the two groups. On the left we report distribution of the average performances at the end on training.

<https://doi.org/10.1371/journal.pcbi.1011948.g007>

In conclusion, the model not only captures the consequences of varying the reliability of sensory information. Also, it correctly predicts the temporal evolution of the kinematics (minimum distances) and the kinetics (leadership indices) of the coordinated movements. Therefore, a simple best-response strategy seems to capture the ongoing development of joint coordination.

Experiment 2: Symmetric via-points. When the via-points are arranged symmetrically—both VPs placed on the axis of the start-target line—the two Nash equilibria are cost-equivalent—see Fig 3b. To develop stable coordination, the players face the additional challenge of converging to the same equilibrium. To investigate how this can be achieved we ran a novel study in which the experimental protocol was the same as in [11], but the via-points were arranged symmetrically. As in the previous study, in two experimental groups we manipulated the information available about the partner—either haptic alone (H) or visuo-haptic (VH).

Experiment 2 has an additional component—decision on strategy. To identify the strategy that the players followed at each trial, we used the minimum distance to the partner's via-point (MD_{i-i}) to determine whether that player got closer to their partner's VP earlier or later in the movement—thus crossing or getting closer to the via-points in the $VP_i \rightarrow VP_{-i}$ order (E strategy for short); the $VP_{-i} \rightarrow VP_i$ order (L strategy); or they only focused on their own VP (M strategy); see the Methods section for details. The overall coordination performance is summarized in Fig 7.

Joint coordination only occurs if the two players follow opposite strategies, i.e. either E-L or L-E. Fig 7a exemplifies these behavior types. In some dyads, one or both players gradually reduce their minimum distance from the opponent via-point—see Fig 7b.

In other dyads, the minimum distance does not decrease during training. Statistical analysis confirms an overall significant Time effect for both players (MD_{12} : $F(1,16) = 6.272$, $p = 0.0235$;

MD_{21} : $F(1,16) = 9.621$, $p = 0.0069$), but there is neither a significant Group effect, nor a significant Time \times Group interaction.

However, a closer look at the MD_{12} and MD_{21} at the end of the training suggest that they have a bimodal distribution, indicating that in both experimental conditions only a fraction of the players attempt to incorporate partner's action in their motor command. From this observation, we classified each player as a 'Learner' if their average MD_{ij} at the end of the training phase was below 0.02m; 'Non Learner' otherwise. Using this criterion, 7 out of 18 players were classified as Learners in the H group and 9 out of 18 in the VH group. The group difference is not statistically significant (χ^2 test).

Overall, these findings suggest that forms of coordination that require the selection of one among multiple cost-equivalent solutions are much more complex processes than the asymmetric 2-VP settings [11], and coordination performance is not determined solely by the reliability of sensory information.

Nevertheless, a fraction of the dyads succeeded in achieving stable coordination in both space and time. To characterize individual behaviors, we estimated the model parameters from the sequence of the movements of each participant over trials. In the model, we set $\tau_i(E) = 0.38 T$, $\tau_i(M) = 0.5 T$ and $\tau_i(L) = 0.62 T$. We identified the model parameters for each player in a dyad from the time course of dyad movements. The fitting results are summarized in Fig 8.

Consistent with the experimental findings, no parameter exhibited a group (H, VH) effect (one-way ANOVA). We then examined differences in the Learner vs Non-Learner groups. In this case, three model parameters exhibited a Group effect. The sensory variance (Σ_y), is smaller in the Learner group— $F(1,34) = 35.24$, $p < 0.0001$. The internal (process) noise variance (Σ_x) is greater in the Learner group— $F(1,34) = 8.43$, $p = 0.0065$. The partner model retention rate (A) is smaller in the Learner group— $F(1,34) = 8.5$, $p = 0.0062$; see Fig 8b.

As in Experiment 1, we simulated the individual dyads. The simulations closely resemble the experimental findings. The model qualitatively predicts the different behaviors of Learners

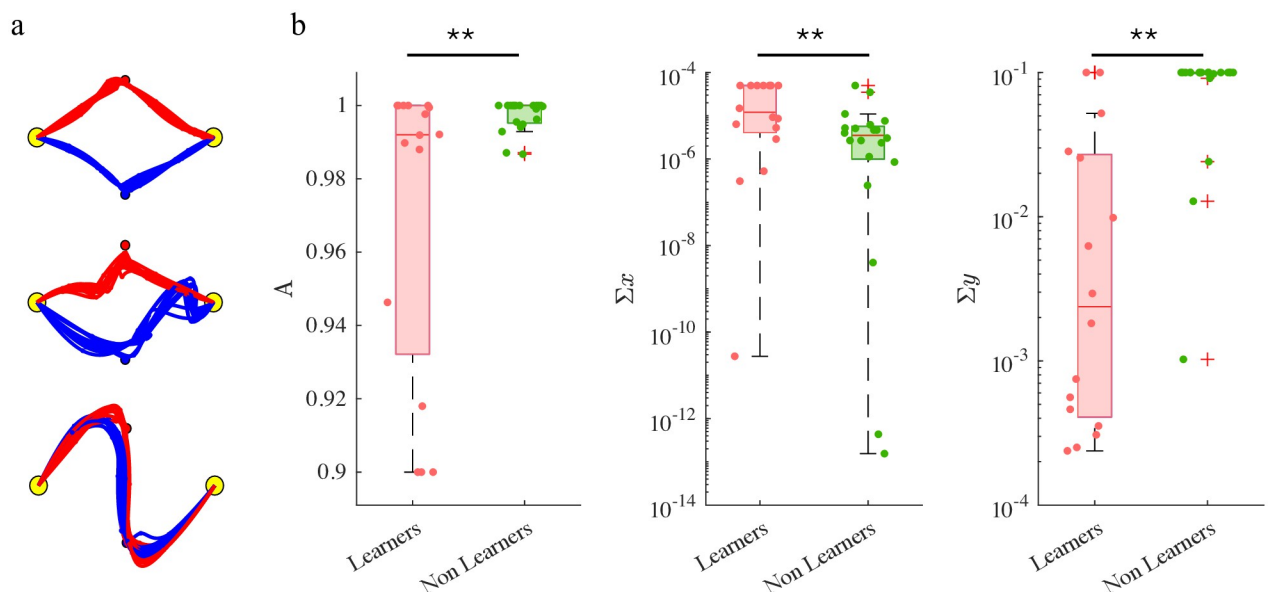


Fig 8. Determinants of learning. a. Estimated trajectories. b, c. Estimated model parameters grouped in terms of Learner and Non-Learner and in H and VH groups. Boxplots display median, 25th and 75th percentiles of the estimated parameters values, for Learners (orange) and Non-Learner (green). *: $p < 0.05$; **: $p < 0.01$.

<https://doi.org/10.1371/journal.pcbi.1011948.g008>

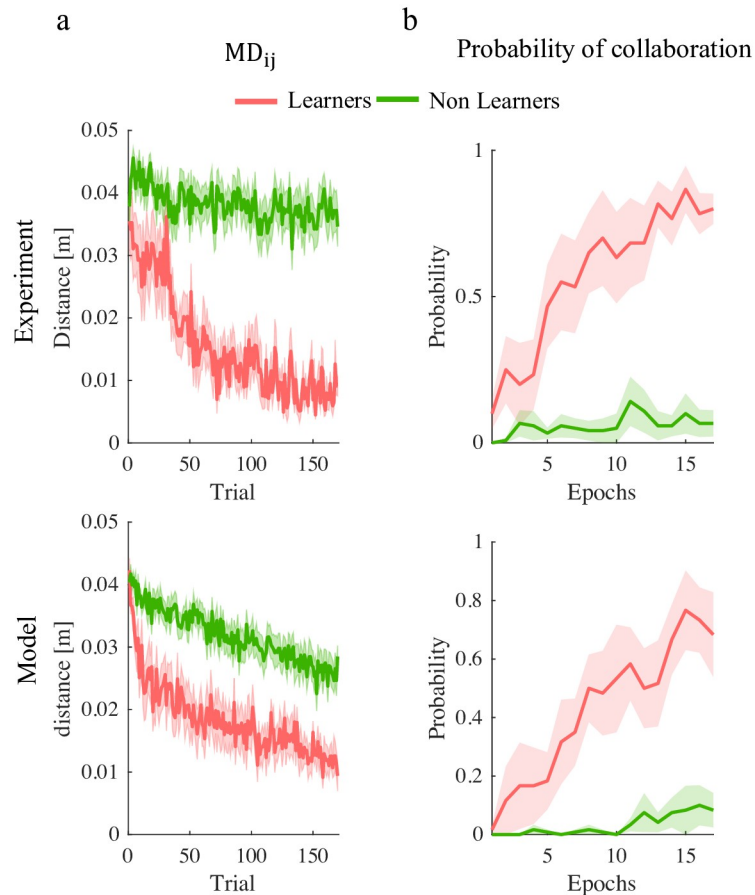


Fig 9. Simulations (bottom) closely resemble Experiments (top), in the temporal evolution of both the minimum distance to partner's via-point over trials (a) and in the probability of establishing a collaboration (b), in both Learners (pink) and Non-Learners (green).

<https://doi.org/10.1371/journal.pcbi.1011948.g009>

and Non Learners. The minimum distances from the partner via-points at the end of training did not significantly differ in the experiment and in the simulations in the Learners groups. We found a significant difference in the Non-Learners group ($p = 0.0056$)—see Fig 9a. In the experiments, Non-Learner dyads ignored their partner during the whole training phase. The model did not fully capture this behavior. This is possibly due to the fact that in these experiments the actions tend to be quite similar from trial to trial. The persistent excitation condition may not be satisfied, thus leading to inaccurate parameter estimates and ultimately to poor match between experiments and simulations.

In simulated trials, the temporal evolution of the probability to establish a collaboration—see Fig 9b—is consistent with the experimental findings. In Learners, the probability tends to increase with time, whereas in Non-Learners the probability remains low.

Discussion

We developed a general computational framework to describe how pairs of agents (dyads) gradually negotiate coordinated movements over multiple repetitions of the same interactive task. Each player is assumed to know their own goal; to have no information about their partner's, and limited information about their action. At each trial, the i -th player selects an action

(u_i) on the basis of a partner model (x_i). After both players have performed their actions, through their sensory system provides each player information (y_i) about their partner's action. This information is used to update the partner model to be used on the next trial. This modeling framework specifically focuses on how uncertainty about own task and uncertainty about the partner determine the game outcome.

The model can be used as a generative tool, to simulate joint action scenarios, or as an analytical tool to characterize individual players' behaviors and strategies in actual joint action experiments. The simulation results may provide insights on the way coordination strategies develop. We demonstrated the model in the context of two different joint action scenarios: a motor version of the Stag Hunt game [32] and a joint 2-Via-Point task [11] with asymmetric and symmetric via-points. In the first scenario, we used simulations to explore the model's parameter space in order to understand the determinants of emergent joint coordination. In the second scenario, we estimated the model parameters from experimental observations, and demonstrated that the model correctly captures the experimental manipulations and their effect on how joint coordination is learned.

Trust as the basis of cooperation

In the Stag Hunt game, the players have the option to choose a safe but less profitable action—hunting a rabbit—which requires no coordination with their partners, or a potentially more profitable action—hunting a stag—which requires that both players aim at it to succeed. As they select their actions independently, under what conditions we can expect that they end up cooperating if they play the game repeatedly? If after each round both players are provided with imperfect information about their partner's action, how is this information used to develop a propensity to behave cooperatively?

The Stag Hunt game has been used extensively as a model for the emergence of trust in social settings. Does it require an innate predisposition, or trust can be learned through repeated interaction? [33, 34]. Our model suggests an intermediate possibility. Under the assumption that own cost structure is completely known and thus there is no task learning, at each trial the prediction of partner actions optimally combines the sensory information available about the previous action, and the previous history of interaction. Hence the partner model adapts continuously over game iterations. The simulations indeed suggest that necessary conditions to 'trust' are (i) a prior belief that the partner does not behave erratically—in the model, this trait corresponds to a low Σ_x^i ; and (ii) reliable information about current partner action—low Σ_y^i ; see Fig 2. In these situations, the Bayesian partner model has more 'inertia', in the sense that it places more weight on previous history of interaction and less on the last action. In its current implementation, partner representation is minimal, in the sense that only focuses on predicting the partner's next action. More sophisticated partner representations, involving their goals and intentions, and possibly the way they represent their own (theory of mind) are likely even more effective in building up trust and thus cooperation [32].

How is joint coordination learned?

In the asymmetric 2-VP task, model simulations replicated the gradual development of coordinated trajectories that was observed empirically [11]—see Figs 5 and 6. In particular, the model correctly predicts that over trials, the hand paths undergo a gradual shaping process, in which they gradually reduce their distance from the partner's via-point. Consistently, roles gradually disappear. Uncertainty about partner action modulates rates of convergence and final performances. This is remarkable as in the model the agents have no explicit information about their partner's task requirements. The model also reproduces the three experimental

conditions tested in the original study. Consistent with experimental observations, with lower sensory noise the players converged toward a shared solution that is closer to the theoretical Nash equilibrium. These results further support the notion that lower sensory noise facilitates the development of joint coordination.

Achieving a joint coordination is not just a matter of more reliable sensory information

To further investigate how interacting agents negotiate a coordination over multiple game repetitions, we focused on a scenario in which the agents must select one among multiple strategies, and develop appropriate spatiotemporal coordination within that selected strategy. To do so, we manipulated the geometry of the 2-VP task so that the via-points were located symmetrically with respect to the horizontal axis, thus leading to two cost-equivalent Nash equilibria. This symmetric version of the 2-VP task is much more challenging than its asymmetric counterpart [11], because to achieve a collaboration, the players not only need to cross or approach their partner's via-point; they must also decide whether to do so before or after crossing their own. Hence achieving a collaboration is more difficult, and may actually never occur.

In these experiments, we observed a gradual improvement of spatiotemporal performance which indicates that, at population level, the dyads increasingly coordinate their movements over trials. However, only a fraction of the dyads actually converged to a stable coordination. Some dyads got stuck in cyclic behaviors, in which both players switch between the two coordination solutions. In other dyads, each player only focused on their own VP by ignoring the mechanical interaction with their partner, even though this implied more mechanical effort on their side. This latter finding suggests that collaboration has a cost, which may overcome the extra mechanical effort required by lack of collaboration in this task. However, we found no evidence that convergence to stable coordination is determined by reliability of information about the partner. This finding is in partial contrast with previous observations [11], that more reliable sensory information has quantitative and qualitative effects on coordination, in the sense that the latter is faster and closer to the predictions of game theory (Nash equilibrium). Nevertheless, some dyads did converge to a shared strategy. The question arises of whether some specific sensorimotor abilities and/or peculiar features of partner representation facilitate this outcome.

The estimated model parameters did not capture the experimental manipulation of the sensory information (H and VH groups); this may be due to the fact the observed macroscopic behaviors—convergence or non-convergence to a stable coordination—are little different in the two experimental groups. It should be noted, however, that the relevant model parameter—sensory covariance Σ_y^i —reflects not only the experimental manipulation but also the individual sensory capability, which may vary in different individuals. When looking at Learners and Non-learners, we found that the former are characterized by a combination of lower sensory covariance Σ_y^i , higher internal partner model noise, Σ_x^i and lower partner model retention, A^i .

In other words, Learners are more prone to integrate current sensory feedback to predict their partner's action, whereas Non Learners tend to keep their initial belief and fail to develop stable coordination. Taken together, these features indicate that Non learners are less likely to modify their partner predictions and therefore to switch strategy from one trial to another. These findings suggest that not only sensory information has quantitative and qualitative effects on coordination [11] but also some other specific sensorimotor abilities and the peculiar aspects of partner representation facilitate this outcome.

Relation to other models of joint action

Previous models of joint action were purely theoretic [32, 35] or were only capable of reproducing qualitative aspects of the observed behaviors [10, 11]. The proposed model extends approaches used to model sensorimotor adaptation and relies on optimality principles to model both the estimation of partner actions and the selection of own action.

The model can be related to the active inference framework [36]. Active inference aims at unifying perception and action selection in terms of the minimization of one single quantity—surprise, or the mismatch between observation and predictions. Previous studies have addressed specific aspect of joint action on an active inference perspective—namely, theory of mind [32], or discriminating between self and joint action [37]. Recently, active inference has been proposed as a possible general framework to study joint action [38] and to explain the collective behaviors of multiple interacting agents [39]. Very much like active inference, we model perception of the partner as Bayesian estimation—by optimally combining observations and model predictions (in our case, of partner's action) and by updating the model in terms of the prediction error. In our model, each player selects the action that minimizes their expected cost, which combines a task representation—for instance, an intended target—that is supposed to be known—and the partner model. In the active inference framework the action cost may be seen as playing the role of a prior [40] for action selection. In our model the two processes—(partner) perception and action selection are kept separate, which allows to use the tools of game theory to address the development of joint coordination.

The model addresses scenarios with shared goals

Scenarios where participants have partly distinct goals provide broader insights into the mechanisms underlying interaction, but the model also covers situations in which participants have one shared goal. The goal is completely specified by each player's cost function. If the cost only depends on the action of that player alone, no interaction is needed to achieve the common goal and there is no need to predict the partner's future action. This is the case, for instance, in which players are required to independently aim at the same spatial target. In this case, there is actually no joint action. Conversely, if the cost depends on the actions of both players, knowledge about the partner's future action becomes crucial. For instance, in a situation in which the two players are required to reach the same target by moving together. In this case, fictive play predicts that each player will imitate their partner by taking that same action.

Model limitations and possible extensions

The model can be applied to different interactive settings. We addressed a setting where the matrices defining the cost function are positive definite, then the conditional action probability is a Gaussian distribution. Otherwise—as in competitive games—the action probability may not be Gaussian although the model formulation is still valid. The only change is in the system identification procedure.

In the presented formulation we made a series of simplifying assumptions. However, the model must be considered as a basic framework, focusing on the sensorimotor level of interaction, which can be easily extended to more complex scenarios.

First, the model assumes that each players has a perfect knowledge of their own goal—the structure of the cost function—and no knowledge at all about their partner's. Decay of the action selection temperature $\lambda_i(t)$ can be interpreted as a gradual refinement of the cost function, but it is possible to add to the model a reinforcement learning mechanism in which a reward—function of the cost—provided to each player at the end of each trial—can be used to update the cost function parameters. Second, the model uses fictitious play to model the

development of coordination. Fictitious play assumes that the players select their action on the basis of information about their partner's last action or, more in general, a probabilistic model of their partner's actions, under the assumption of a stationary strategy. Our simulations suggest that this model produces plausible behaviors and—at least in two studies [9, 11]—reproduces well the reported joint action time series. However, this does not rule out the possibility that in other joint action scenarios, the players also develop a model of their partner's goals—encoded in the structure of their cost function. This information could be used, for instance, as a prior in the partner model [41]. The players may use even more sophisticated insight—e.g., the partner's mental state or Theory of Mind to make their decisions. The partner's mental state can be made part of the partner model, being inferred from observation of the action outcomes, and accounted for in the expected cost during action selection.

Conclusions

In summary, our proposed model successfully captures the experimental findings across various strategic scenarios. The model provided valuable predictive capabilities and highlighted the importance of reliable information and exploration in achieving Nash equilibria. The parameter identification process further elucidated the impact of experimental conditions on the estimated model parameters. Overall, our model serves as a useful tool for understanding and predicting strategic decision-making processes in diverse settings.

Overall, these findings confirm that model-based analysis of joint action experiments may provide valuable information on the underlying mechanisms. In particular, we addressed joint action using a pair of computational models representing the two interacting agents.

Coupled computational models that evolve over time to describe the structure and behavior of a physical system or process are referred to as digital twins. Digital twins and computational models of joint action are generally complicated frameworks which require custom implementations. The presented model relies on dynamical system theory and probabilistic graphical models. This is consistent with previous works [42] proposed a general and robust foundation for the development of digital twins. As such, the model described in this chapter is a promising tool to advance the understanding of the processes underlying joint action and a wide variety of applications can be envisioned. In healthcare, computational models or digital twins of interacting agents promise to advance medical assessment, diagnosis, and personalized treatment. In a similar way, in educational settings, they can be used to devise personalized training paths.

Methods

Ethics statement

The study conforms to the ethical standards laid down in the 1964 Declaration of Helsinki that protects research subjects and was approved by the Institutional Review Board (Comitato Etico Regione Liguria). All participants provided written informed consent to participate in the study.

Spatial Stag Hunt

The parameters of the cost function can be calculated in order to achieve the cost matrix reported in Fig 2. In particular, in Eq 13 we set $J_1(u_R, u_R, 1) = J_1(u_R, u_S, 1) = 5$, $J_1(u_S, u_R, 0) = 10$, and $J_1(u_S, u_S, 0) = 1$ —same for J_2 . These constraints uniquely determine the three

parameters: $w = 9/2$, $z_R = 5$, and $z_S = 1$. The model parameters were finally set as follows:

$$\begin{aligned}
 R_{ii}^i(s_i) &= s_i \cdot w + 2 \cdot (1 - s_i) \cdot w \\
 R_{i-i}^i(s_i) &= -(1 - s_i) \cdot w \\
 R_{-i-i}^i(s_i) &= (1 - s_i) \cdot w \\
 r_i^i(s_i) &= -s_i \cdot w \cdot u_R - (1 - s_i) \cdot w \cdot u_S \\
 r_{-i}^i(s_i) &= 0 \\
 z^i(s_i) &= s_i \cdot \left(\frac{1}{2} \cdot w \cdot u_R^2 + z_R\right) + (1 - s_i) \cdot \left(\frac{1}{2} \cdot w \cdot u_S^2 + z_S\right)
 \end{aligned} \tag{16}$$

We simulated this game scenario under two different levels of sensory noise variance (Σ_y^i) (low: 0.1; high: 10) and player's belief on partner predictability or internal noise Σ_x^i (high predictability/low internal noise: 0.1; low predictability/high internal noise: 10). For all other model parameters we set the following baseline values: $\mu^i = 0$, $P_0^i = \epsilon$, $A^i = 0.99$, $\lambda_1^i = 0.1$ and $a^i = 0.999$. For each combination of Σ_y^i and Σ_x^i we simulated eight dyads. For each player in a dyad, we applied random variations (within a $\pm 10\%$ range) to the baseline parameter values. We repeated the game a total of 2000 times (50 blocks of 40 trials each). We used a 2-way ANOVA to assess the effect of the two factors (sensory noise: low-high and internal noise: low-high).

Two via-point task

Experiments. The experimental protocol is identical to [11]. Each participant sat in front of a computer screen and grasped the handle of a haptic interface (Novint Falcon). They could not see or hear each other and were not allowed to talk. The participants were instructed to perform planar point-to-point movements in the vertical plane, between the same start and end point (displayed, respectively as a white and a yellow circle of $\odot 1$ cm), but through different via-points (a blue circle for Player 1, a red circle for Player 2, both with $\odot 0.5$ cm. In a reference frame centered on the robot workspace, with the X axis aligned with the left-right direction and the Y axis aligned with the vertical direction, for both players the start point was placed in the (-5, 0, 0) cm position and the target point was placed in the (5, 0, 0) cm position. Hence the start and the target point had a horizontal distance of 10 cm. In Experiment 1 the via-points were placed, respectively, at locations $VP_1 = (-2, -3, 0)$ cm and $VP_2 = (2, 3, 0)$ cm, asymmetric with respect to the Y axis. In Experiment 2 the via-points were placed, respectively, at locations $VP_1 = (0, -3, 0)$ cm and $VP_2 = (0, 3, 0)$ cm, symmetric with respect to the Y axis. The two participants were mechanically connected through the haptic interfaces, which generated a force proportional to the difference of the two hand positions, i.e. $F_i(\tau) = -k(p_i(\tau) - p_{-i}(\tau))$. The experimental protocol was organized into epochs of 12 movements each and consisted of three phases: (i) baseline (1 epoch), (ii) training (17 epochs), and (iii) wash-out (2 epochs) for a total of $20 \times 12 = 240$ movements. During the baseline phase, the interaction forces were turned off and each subject performed on their own ('solo' performance). During the interaction phase, the participants were told that they might experience a force while performing the task and were instructed to keep this force to a minimum. We introduced this additional requirement because of a technical limitation of our robots, which are unable to generate large stiffnesses. With a low stiffness the haptic perception of the interaction force is less reliable, and introducing an explicit requirement on low interaction forces made sure that all players

were provided with exactly the same task requirements irrespective of the information they had available about their partner.

In Experiment 1 we reanalyze the data from [11]—a total of 30 participants forming 15 dyads, matched by their body mass index (BMI). Each dyad was randomly assigned to three groups (5 dyads per group), which differed in the available sensory information. In the haptic (H) group the interaction could only be sensed haptically. In the visuo-haptic (VH) group the interaction force (magnitude, direction) was also displayed as an arrow attached to the cursor (scale factor: 10 N/cm). In the partner-visible group (PV), in addition to the interaction force the partner position was continuously displayed on the screen. In Experiment 2, we recruited 36 participants, all right-handed (Edinburgh Handedness Inventory [43]). All participants were naive to the task and had no known history of neurological or upper limb motor impairment. We formed 18 BMI-matched dyads. The two participants within the same dyad were randomly labeled as, respectively, Player 1 and Player 2. Each dyad was randomly assigned to two groups (H: 6 M + 12 F, age 25 ± 3 ; VH: 8 M + 10 F, age 24 ± 5), with the same meaning as Experiment 1.

Data analysis. During experiments, hand trajectories and the interaction forces generated by the robots were sampled at 100 Hz and stored for subsequent analysis. Movement and force trajectories were smoothed using a 4th order Savitzky-Golay filter with a 370 ms time window. The same filter was used to estimate velocity and acceleration. A sign of coordination is that each player, while passing through their own via-point, also gets very close to their partner's. This can be quantified in terms of the Minimum Via-Point Distance, MD_{ij} of player i from via-point j : $MD_{ij} = \min_{\tau} \text{dist}_{ij}(\tau)$, with $\text{dist}_{ij}(t) = \|x_i(\tau) - x_{VP_j}\|$.

A pre-requisite for coordination is that both players cross their partner's via-point, i.e. both reduce their MD_{i-i} . We classified each trial as collaborative in space if both players had $MD_{ij} < d$, with $d = 0.02$ m. The threshold was set based on the observation of dyads who clearly exhibited a coordination. We then estimated the decision variable $s_i(t)$ in which at every trial t player i selects their action $s_i(t)$ among three options: moving to VP_i and then VP_{-i} (E strategy); moving to VP_{-i} and then VP_i (L strategy); or moving to own VP while ignoring the other (M strategy). In summary:

$$s_i(t) = \begin{cases} \text{E} & \text{if } (MD_{ii} < d) \wedge (MD_{i-i} < d) \wedge (TC_{ii} < TC_{i-i}) \\ \text{M} & \text{if } (MD_{ii} < d) \wedge (MD_{i-i} > d) \\ \text{L} & \text{if } (MD_{ii} < d) \wedge (MD_{i-i} < d) \wedge (TC_{ii} > TC_{i-i}) \end{cases} \quad (17)$$

where TC_{ij} is the time instant—with respect to the dyad start time—at which player i achieves its minimum distance to via-point j : $TC_{ij} = \text{argmin}_t \text{dist}_{ij}(t)$; therefore, $MD_{ij} = \text{dist}_{ij}(TC_{ij})$. To understand the determinants of learning, we defined a player as a 'Learner' if they consistently achieved either E or L during the last epoch of the training phase; 'Non Learner' otherwise. Note that both players in a dyad being learners according to this definition does not necessarily imply that they achieve a collaboration—they may get stuck into a cyclic behavior in which they alternate E and L strategies and never converge. We then focused on dyad behaviors. We classified a dyad as collaborative if both players on that trial used either E-L or L-E:

$$\text{collaboration}(t) = [(s_1(t) = \text{E}) \wedge (s_2(t) = \text{L})] \vee [(s_1(t) = \text{L}) \wedge (s_2(t) = \text{E})] \quad (18)$$

In a similar way, we classified a trial as 'cyclic' if the two players took either E-E or L-L, and as 'non-collaborative' if at least one of the players ignores the other. Using these definitions, for each dyad we calculated the probability of collaboration, cycling and non-collaboration over each 12-trial epoch. We also calculated a leadership index (LI_{ij})—the power of the interaction

force of the i -th player, averaged over the 300 ms interval just before crossing the j -th via-point. A negative or positive LI_{ij} denotes, respectively, that the i -th player moves against the interaction ‘leader’—or just passively complies with it—a ‘follower’.

For all indicators, we ran a repeated-measures ANOVA with group (H, VH, PV or V, VH) and Time (early—training epoch 1—and late—epoch 17) as factors. We used post-hoc analysis (Tukey’s HSD) to further identify significant differences between group pairs in the case of significant group effects.

Model. We approximated the trajectory at time τ as a polyharmonic spline:

$$p_i(\tau) = \sum_{m=1}^M v_m^x |\tau - \tau_m|^3 + v_{M+1}^x \cdot \tau + v_{M+2}^x = g(\tau)^T \cdot v_i(\tau) = g(\tau)^T \cdot S \cdot u_i \tag{19}$$

where $g(\tau) = [|\tau - \tau_1|^3 \dots |\tau - \tau_M|^3 \tau 1]^T$ and S is a constant matrix. Considering that some components of u_i^x (and u_i^y) are fixed (initial and final position), we finally have

$$\begin{aligned} p_i(\tau) &= g(\tau)^T \cdot [S_1 \cdot u_i + S_0 \cdot u_0] \\ \dot{p}_i(\tau) &= \dot{g}(\tau)^T \cdot [S_1 \cdot u_i + S_0 \cdot u_0] \\ \ddot{p}_i(\tau) &= \ddot{g}(\tau)^T \cdot [S_1 \cdot u_i + S_0 \cdot u_0] \end{aligned} \tag{20}$$

where $u_i = [p_i(\tau_1) \dots p_i(\tau_{M-4})]^T$ and $u_0 = [p_i(0)p_i(\Delta t)p_i(T - \Delta t)p_i(T)]^T$, with $\Delta t = 0.01$ s. We set $p_i(\Delta t) = p_i(0)$ and $p_i(T - \Delta t) = p_i(T)$ to make sure that all trajectories started and ended with zero velocity. We defined the ‘action’ of the i -th player at a given trial as the vector u_i of nodes’ coordinates. This approximation captures the overall shape of a generic trajectory in terms of the coordinates of a small number of nodes. Subsequent derivatives of $p_i(\tau)$ (velocity, acceleration) can also be expressed as linear combinations of the u_i ’s. In all simulations we used $M = 9$ nodes. In this way, Eq 14 becomes a quadratic form in u_i and u_{-i} . Parameters w_1 , w_2 , and r were determined by the Bryson rule: $w_k = 1/d_{k_{\max}}$ where $d_{k_{\max}}$ is the maximum acceptable magnitude of the corresponding term of the cost function. In our task, we set $w_1 = 1/r_{VP}^2$, $w_2 = 1/d_{\max}^2$ and $r = 1/F_{\max}^2$, where the VP radius is $r_{VP} = 2.5$ mm, the maximum between-subject distance is $d_{\max} = 2$ mm and the maximum force is $F_{\max} = 1$ N.

For both experiments, we estimated the model parameters for all individual participants from the actual experimental data—movement sequences over the 100-trial training phase. These parameters describe the learning behavior of each player within their dyad. For each model parameter, we used a one-way ANOVA with Group (Experiment 1: H, VH, PV; Experiment 2: H, VH) as factor. In the case of significant group effects, we used post-hoc analysis (Tukey’s HSD) to further identify significant differences between group pairs.

To further clarify the role of sensory noise in the development of joint coordination, we used the model as a generative tool, by simulating the temporal evolution of the movements in all experimental conditions. To simulate individual dyads, we used the exact same parameters estimated in the fitting procedure for that.

Supporting information

S1 File. Supplementary note with model implementation details and validation results.
(PDF)

S1 Data. Zip folder containing data from all indicators used to address the interactive scenarios.
(ZIP)

Author Contributions

Conceptualization: Cecilia De Vicariis, Laura Bandini, Vittorio Sanguineti.

Data curation: Cecilia De Vicariis, Vittorio Sanguineti.

Formal analysis: Cecilia De Vicariis, Vittorio Sanguineti.

Funding acquisition: Vittorio Sanguineti.

Investigation: Cecilia De Vicariis, Vinil T. Chackochan, Eleonora Ravaschio.

Methodology: Cecilia De Vicariis, Vittorio Sanguineti.

Project administration: Cecilia De Vicariis, Vittorio Sanguineti.

Resources: Cecilia De Vicariis, Vittorio Sanguineti.

Software: Cecilia De Vicariis, Vittorio Sanguineti.

Supervision: Vittorio Sanguineti.

Validation: Cecilia De Vicariis, Vittorio Sanguineti.

Visualization: Cecilia De Vicariis, Vittorio Sanguineti.

Writing – original draft: Cecilia De Vicariis, Vittorio Sanguineti.

Writing – review & editing: Cecilia De Vicariis, Laura Bandini, Vittorio Sanguineti.

References

1. Sebanz N, Bekkering H, Knoblich G. Joint action: bodies and minds moving together. *Trends Cogn Sci*. 2006; 10(2):70–6. <https://doi.org/10.1016/j.tics.2005.12.009> PMID: 16406326
2. Vesper C, Butterfill S, Knoblich G, Sebanz N. A minimal architecture for joint action. *Neural Netw*. 2010; 23(8-9):998–1003. <https://doi.org/10.1016/j.neunet.2010.06.002> PMID: 20598504
3. Butterfill SA, Sebanz N. Editorial: Joint Action: What Is Shared? *Review of Philosophy and Psychology*. 2011; 2(2):137–146. <https://doi.org/10.1007/s13164-011-0062-3>
4. Della Gatta F, Garbarini F, Rabuffetti M, Viganò L, Butterfill SA, Sinigaglia C. Drawn together: When motor representations ground joint actions. *Cognition*. 2017; 165:53–60. <https://doi.org/10.1016/j.cognition.2017.04.008> PMID: 28501547
5. Barchiesi G, Zazio A, Marcantoni E, Bulgari M, Barattieri di San Pietro C, Sinigaglia C, et al. Sharing motor plans while acting jointly: A TMS study. *Cortex*. 2022; 151:224–239. <https://doi.org/10.1016/j.cortex.2022.03.007> PMID: 35447381
6. Bekkering H, de Bruijn ERA, Cuijpers RH, Newman-Norlund R, Van Schie HT, Meulenbroek R. Joint action: neurocognitive mechanisms supporting human interaction. *Top Cogn Sci*. 2009; 1(2):340–52. <https://doi.org/10.1111/j.1756-8765.2009.01023.x> PMID: 25164937
7. Reed K, Peshkin M, Hartmann MJ, Grabowecky M, Patton J, Vishton PM. Haptically linked dyads: are two motor-control systems better than one? *Psychol Sci*. 2006; 17(5):365–6. <https://doi.org/10.1111/j.1467-9280.2006.01712.x> PMID: 16683920
8. Ganesh G, Takagi A, Osu R, Yoshioka T, Kawato M, Burdet E. Two is better than one: physical interactions improve motor performance in humans. *Sci Rep*. 2014; 4:3824. <https://doi.org/10.1038/srep03824> PMID: 24452767
9. Braun DA, Ortega PA, Wolpert DM. Nash equilibria in multi-agent motor interactions. *PLoS Comput Biol*. 2009; 5(8):e1000468. <https://doi.org/10.1371/journal.pcbi.1000468> PMID: 19680426
10. Takagi A, Ganesh G, Yoshioka T, Kawato M, Burdet E. Physically interacting individuals estimate the partner's goal to enhance their movements. *Nature Human Behaviour*. 2017; 1(3):0054. <https://doi.org/10.1038/s41562-017-0054>
11. Chackochan VT, Sanguineti V. Incomplete information about the partner affects the development of collaborative strategies in joint action. *PLoS Comput Biol*. 2019; 15(12):e1006385. <https://doi.org/10.1371/journal.pcbi.1006385> PMID: 31830100
12. Thoroughman KA, Shadmehr R. Learning of action through adaptive combination of motor primitives. *Nature*. 2000; 407(6805):742–7. <https://doi.org/10.1038/35037588> PMID: 11048720

13. Scheidt RA, Dingwell JB, Mussa-Ivaldi FA. Learning to move amid uncertainty. *J Neurophysiol*. 2001; 86(2):971–85. <https://doi.org/10.1152/jn.2001.86.2.971> PMID: 11495965
14. Hwang EJ, Smith MA, Shadmehr R. Dissociable effects of the implicit and explicit memory systems on learning control of reaching. *Exp Brain Res*. 2006; 173(3):425–37. <https://doi.org/10.1007/s00221-006-0391-0> PMID: 16506003
15. Smith MA, Ghazizadeh A, Shadmehr R. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol*. 2006; 4(6):e179. <https://doi.org/10.1371/journal.pbio.0040179> PMID: 16700627
16. Smith MA, Shadmehr R. Intact ability to learn internal models of arm dynamics in Huntington's disease but not cerebellar degeneration. *J Neurophysiol*. 2005; 93(5):2809–21. <https://doi.org/10.1152/jn.00943.2004> PMID: 15625094
17. Casadio M, Sanguineti V, Morasso P, Solaro C. Abnormal sensorimotor control, but intact force field adaptation, in multiple sclerosis subjects with no clinical disability. *Mult Scler*. 2008; 14(3):330–42. <https://doi.org/10.1177/1352458507085068> PMID: 18208874
18. Casadio M, Sanguineti V. Learning, retention, and slacking: a model of the dynamics of recovery in robot therapy. *IEEE Trans Neural Syst Rehabil Eng*. 2012; 20(3):286–96. <https://doi.org/10.1109/TNSRE.2012.2190827> PMID: 22531822
19. De Vicariis C, Chackochan VT, Sanguineti V. Game theory and partner representation in joint action: toward a computational theory of joint agency. *Phenomenology and the Cognitive Sciences*. 2022.
20. Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science*. 1995; 269(5232):1880–2. <https://doi.org/10.1126/science.7569931> PMID: 7569931
21. Körding KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature*. 2004; 427(6971):244–7. <https://doi.org/10.1038/nature02169> PMID: 14724638
22. Braun DA, Ortega PA, Wolpert DM. Motor coordination: when two have to act as one. *Exp Brain Res*. 2011; 211(3-4):631–41. <https://doi.org/10.1007/s00221-011-2642-y> PMID: 21455618
23. Nash J. Non-Cooperative Games. *Annals of Mathematics*. 1951; 54(2):286–295. <https://doi.org/10.2307/1969529>
24. Grau-Moya J, Hez E, Pezzulo G, Braun DA. The effect of model uncertainty on cooperation in sensorimotor interactions. *J R Soc Interface*. 2013; 10(87):20130554. <https://doi.org/10.1098/rsif.2013.0554> PMID: 23945266
25. Lindig-León C, Schmid G, Braun DA. Nash equilibria in human sensorimotor interactions explained by Q-learning with intrinsic costs. *Sci Rep*. 2021; 11(1):20779. <https://doi.org/10.1038/s41598-021-99428-0> PMID: 34675336
26. Brown GW. Iterative solution of games by fictitious play. *Act Anal Prod Allocation*. 1951; 13(1):374.
27. Fudenberg D, Levine DK. *The theory of learning in games*. vol. 2. MIT press; 1998.
28. Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. *Nat Neurosci*. 2002; 5(11):1226–35. <https://doi.org/10.1038/nn963> PMID: 12404008
29. Başar T, Olsder GJ. *Dynamic noncooperative game theory*. vol. 23. 2nd ed. SIAM; 1999.
30. Todorov E. Optimality principles in sensorimotor control. *Nat Neurosci*. 2004; 7(9):907–15. <https://doi.org/10.1038/nn1309> PMID: 15332089
31. McKelvey RD, Palfrey TR. Quantal response equilibria for normal form games. *Games and economic behavior*. 1995; 10(1):6–38. <https://doi.org/10.1006/game.1995.1023>
32. Yoshida W, Dolan RJ, Friston KJ. Game theory of mind. *PLoS Comput Biol*. 2008; 4(12):e1000254. <https://doi.org/10.1371/journal.pcbi.1000254> PMID: 19112488
33. Fang C, Kimbrough SO, Pace S, Valluri A, Zheng Z. On Adaptive Emergence of Trust Behavior in the Game of Stag Hunt. *Group Decision and Negotiation*. 2002; 11(6):449–467. <https://doi.org/10.1023/A:1020639132471>
34. Zhou L, Su C, Sun X, Zhao X, Choo KKR. Stag hunt and trust emergence in social networks. *Future Generation Computer Systems*. 2018; 88:168–172. <https://doi.org/10.1016/j.future.2018.05.053>
35. Pesquita A, Whitwell RL, Enns JT. Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychon Bull Rev*. 2018; 25(5):1751–1769. <https://doi.org/10.3758/s13423-017-1393-6> PMID: 29119405
36. Friston K, Kilner J, Harrison L. A free energy principle for the brain. *J Physiol Paris*. 2006; 100(1-3):70–87. <https://doi.org/10.1016/j.jphysparis.2006.10.001> PMID: 17097864
37. Kahl S, Kopp S. A Predictive Processing Model of Perception and Action for Self-Other Distinction. *Front Psychol*. 2018; 9:2421. <https://doi.org/10.3389/fpsyg.2018.02421> PMID: 30559703

38. Lehmann K, Bolis D, Friston KJ, Schilbach L, Ramstead MJD, Kanske P. An Active-Inference Approach to Second-Person Neuroscience. *Perspect Psychol Sci.* 2023; p. 17456916231188000. <https://doi.org/10.1177/17456916231188000> PMID: 37565656
39. Heins C, Millidge B, da Costa L, Mann R, Friston K, Couzin I. Collective behavior from surprise minimization. *arXiv.* 2023;2307.14804.
40. Friston K, Samothrakis S, Montague R. Active inference and agency: optimal control without cost functions. *Biol Cybern.* 2012; 106(8-9):523–41. <https://doi.org/10.1007/s00422-012-0512-8> PMID: 22864468
41. Fudenberg D, Levine D. Learning in games. *European Economic Review.* 1998; 42(3):631–639. [https://doi.org/10.1016/S0014-2921\(98\)00011-7](https://doi.org/10.1016/S0014-2921(98)00011-7)
42. Kapteyn MG, Pretorius JV, Willcox KE. A probabilistic graphical model foundation for enabling predictive digital twins at scale. *Nature Computational Science.* 2021; 1(5):337–347. <https://doi.org/10.1038/s43588-021-00069-0> PMID: 38217207
43. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia.* 1971; 9(1):97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4) PMID: 5146491