



CrossMark  
click for updates

## Research

**Cite this article:** Grau-Moya J, Hez E, Pezzulo G, Braun DA. 2013 The effect of model uncertainty on cooperation in sensorimotor interactions. *J R Soc Interface* 10: 20130554.  
<http://dx.doi.org/10.1098/rsif.2013.0554>

Received: 21 June 2013

Accepted: 25 July 2013

### Subject Areas:

computational biology

### Keywords:

model uncertainty, risk sensitivity, cooperation, stag hunt

### Author for correspondence:

J. Grau-Moya

e-mail: [jordi.grau@tuebingen.mpg.de](mailto:jordi.grau@tuebingen.mpg.de)

# The effect of model uncertainty on cooperation in sensorimotor interactions

J. Grau-Moya<sup>1,2</sup>, E. Hez<sup>1,2</sup>, G. Pezzulo<sup>3,4</sup> and D. A. Braun<sup>1,2</sup>

<sup>1</sup>Max Planck Institute for Biological Cybernetics, and <sup>2</sup>Max Planck Institute for Intelligent Systems, Spemannstrasse 38, Tübingen 72076, Germany

<sup>3</sup>Institute of Computational Linguistics, "Antonio Zampolli", CNR, Via Giuseppe Moruzzi 1, Pisa 56124, Italy

<sup>4</sup>Institute of Cognitive Sciences and Technologies, CNR, Via S. Martino della battaglia 44, Roma 00185, Italy

Decision-makers have been shown to rely on probabilistic models for perception and action. However, these models can be incorrect or partially wrong in which case the decision-maker has to cope with model uncertainty. Model uncertainty has recently also been shown to be an important determinant of sensorimotor behaviour in humans that can lead to risk-sensitive deviations from Bayes optimal behaviour towards worst-case or best-case outcomes. Here, we investigate the effect of model uncertainty on cooperation in sensorimotor interactions similar to the stag-hunt game, where players develop models about the other player and decide between a pay-off-dominant cooperative solution and a risk-dominant, non-cooperative solution. In simulations, we show that players who allow for optimistic deviations from their opponent model are much more likely to converge to cooperative outcomes. We also implemented this agent model in a virtual reality environment, and let human subjects play against a virtual player. In this game, subjects' pay-offs were experienced as forces opposing their movements. During the experiment, we manipulated the risk sensitivity of the computer player and observed human responses. We found not only that humans adaptively changed their level of cooperation depending on the risk sensitivity of the computer player but also that their initial play exhibited characteristic risk-sensitive biases. Our results suggest that model uncertainty is an important determinant of cooperation in two-player sensorimotor interactions.

## 1. Introduction

When interacting with its environment, the human sensorimotor system has been shown to use predictive models for control and estimation [1–6]. These models are thought to be probabilistic in nature, and considerable evidence suggests that learning of such models is consistent with the process of Bayesian inference [7–10]. Such probabilistic models are not only important for perception but they can also be used for decision-making and motor control [11–16]. Importantly, decision-makers who maximize expected gain (or minimize expected costs) require probabilistic models of their environment so that they can determine an expectation value. However, such optimal decision-makers have no performance guarantees if their model happens to be partially incorrect or plain wrong [17]. This raises the issue of decision-making strategies that do not rely on accurate probabilities. An extreme example strategy that completely dispenses with probabilities altogether are maximin strategies where the decision-maker picks an action that is optimal under the assumption of a worst-case scenario (or minimax strategies in the case of costs). Such a decision-maker, for example, would take out insurance not for the calamity with highest expected costs, but for the most disastrous (possibly low-probability) calamity, because of not knowing the probability. Similarly, an extremely optimistic decision-maker would assume a best-case scenario following a maxmax strategy (or a minmin strategy in the case of costs), for example by buying lottery tickets with the highest prize, independent of the presumed winning probabilities. Risk-sensitive decision-makers strike a compromise between the two extremes: they have a probabilistic model that they distrust

to some extent, but they do not completely dismiss it—though the extreme cases of robust or optimistic and expected gain decision-making can also be considered as risk-sensitive limit cases [18].

More formally, we can think of a decision-maker who considers model uncertainty in the following way [17,19]. Initially, the decision-maker has a probabilistic model  $p_0$ , but knowing that this model may not be entirely accurate, the decision-maker allows deviations from it, which leads to a new effective probabilistic model  $p$ . The transformation between  $p_0$  and  $p$  has to be constrained when the decision-maker is very confident about the model. Conversely, when the decision-maker is very insecure about the correctness of the model, there should be leeway for larger deviations. The effective value of a choice set with outcomes  $x$  under the effective probability  $p$  can then be stated as

$$V = \text{ext}_{p(x)} \left[ \int dx p(x) U(x) - \frac{1}{\beta} \int dx p(x) \log \frac{p(x)}{p_0(x)} \right], \quad (1.1)$$

where the utility  $U(x)$  quantifies the desirability of  $x$ . The first term is the expected utility under  $p$ , and the second term—formed by the cost factor  $1/\beta$  times the Kullback–Leibler (KL) divergence—captures the cost of the transformation from  $p_0$  to  $p$ . When  $1/\beta > 0$ , we have to replace the extremum operator with a max operator (concave maximization), when  $1/\beta < 0$ , we have to replace it with a min operator (convex minimization). Sensitivity to model uncertainty is modulated by  $\beta$ . When  $\beta \rightarrow 0$ , we recover a decision-maker without model uncertainty. For  $\beta \rightarrow -\infty$ , we get a maximin decision-maker who picks the choice set with maximum  $V$ , where each  $V$  considers the worst-case scenario of the choice set. In fact, the quantity  $V$  is a free energy difference, and equation (1.1) can be motivated by statistical physics (see §2.1).

Recently, it was found that model uncertainty also affects decision-making in sensorimotor integration tasks where subjects have to form beliefs about latent variables, for example the position of a hidden target [20]. However, latent variables do not only play an important role in single-player environments but also in multi-player sensorimotor interactions [21,22], where the policy of the other player can be considered as a particular latent variable. Sensorimotor interactions in humans range from hand shaking and avoiding collisions with another passer-by to tandem riding, tango dancing and arm wrestling. As in the case of single-player environments, the presence of a latent variable suggests the formation of a belief model that can be exploited for prediction and planning [7–10]. And as in the case of the single-player environment, decision-makers might exhibit model uncertainty [20]. Especially when meeting a player for the first time, only a little information about this player's strategies is available. The initial trust or distrust with respect to this player can be thought of as an optimistic or pessimistic bias. However, as more information about the unknown player becomes available, such deviations should vanish and be replaced by accurate statistical estimates.

Sensorimotor interactions can be of cooperative nature, as in the case of dancing, or of competitive nature as in the case of arm wrestling. To investigate the effect of model uncertainty on cooperation, we study sensorimotor interactions similar to the stag-hunt game. In the stag-hunt game, each player decides whether to hunt a higher-valued stag or a lower-valued hare. However, the stag is caught only if both

players have decided to hunt stag. By contrast, a hare can be caught by each player independently. The stag-hunt game is a coordination game with two pure Nash equilibria, given by the pay-off-dominant stag solution, where both players hunt stag and achieve the highest possible pay-off, and the risk-dominant hare solution, where both players hunt hare and obtain a lower pay-off. The latter solution is called risk-dominant, because a player hunting hare knows exactly the pay-off he will receive, which is higher than he would get if he hunted stag by himself. The stag-hunt game is therefore often used to study the emergence of cooperation.

In our study, we investigate a decision-making model that forms Bayesian beliefs about the other player's strategy based on empirically observed choice frequencies. In simulations, we study how model uncertainty with respect to these beliefs affects cooperation in a stag-hunt-like setting. To test human behaviour in stag-hunt-like sensorimotor interactions, we use a previously developed paradigm that allows translating  $2 \times 2$  matrix games into sensorimotor interactions [21,22]. In the experiment, one of the players is simulated by a virtual computer player who is based on our risk-sensitive decision-making model. This way, we can directly manipulate the risk sensitivity of the artificial player and observe the response of the human player.

## 2. A risk-sensitive model of interaction

Classic models in game theory are usually equilibrium models that predict the occurrence of Nash equilibria, that is joint settings of strategies where no individual player has any incentive to deviate unilaterally from their strategy [23]. In evolutionary game theory, this problem is addressed by developing dynamic learning models that converge to the equilibria [24]. One of the simplest classes of such learning models is *fictitious play* [25–27]. In fictitious play, it is assumed that the other player plays with a stationary strategy, which is estimated by the hitherto observed empirical choice frequencies. In our model, we also adopt the assumption of modelling the other player with a stationary strategy, but form a Bayesian belief about this strategy. In the case of the stag-hunt game, this strategy is a distribution over a binary random variable that indicates the two possible actions, namely whether to hunt stag or hare. This distribution can be expressed as a beta distribution. After observing  $s$  choices of stag and  $h$  choices of hare from the opponent, the decision-maker's belief about the strategy  $x$  of the other player is then given by

$$P(x|s, h) = \frac{x^s (1-x)^h}{\int_0^1 dx x^s (1-x)^h}, \quad (2.1)$$

where the opponent's stationary strategy is represented by the probability  $x$  of choosing stag. For a known strategy  $x^*$  of the opponent where  $p(x) = \delta(x - x^*)$ , the decision-maker faces the following expected pay-off

$$EU_1(a_1|x^*) = x^* U(a_1, a_2 = S) + (1 - x^*) U(a_1, a_2 = H), \quad (2.2)$$

with  $U(a_1, a_2)$  denoting the player's pay-off if he chooses action  $a_1$ , and the opponent chooses  $a_2$ . Under strategy  $x^*$ , the opponent chooses  $a_2 = S$  with probability  $x^*$  and  $a_2 = H$  with probability  $1 - x^*$ . In fictitious play, the decision-maker simply gives a best response to this expected pay-off, where

$x^*$  is given by the empirical frequencies and corresponds to the mean of the beta distribution. By contrast, we construct a decision-maker that takes the uncertainty over the  $x$ -estimate into account and exhibits risk sensitivity with respect to this belief over  $x$ . This can be achieved by inserting (2.1) as  $p_0$  and (2.2) as  $U(x)$  into equation (1.1), which results in

$$V(a_1) = \frac{1}{\beta} \log \int_0^1 dx P(x|s, h) e^{\beta(xU(a_1, a_2=S) + (1-x)U(a_1, a_2=H))}.$$

The value  $V(a_1)$  assigned to each action depends on the parameter  $\beta$ , that in our case represents the risk sensitivity. For action selection, we assume a soft-max decision rule

$$P(a_1) = \frac{e^{\alpha V(a_1)}}{\sum_{a'_1} e^{\alpha V(a'_1)}}, \quad (2.3)$$

where  $\alpha$  is a rationality parameter that regulates how deterministic the response is. Soft-max decision rules are prevalent in quantal response equilibrium models to formalize the bounded rationality of decision-makers in games [28]. This includes the theoretically best response in the limit  $\alpha \rightarrow \infty$  that corresponds to a perfect rational agent that is able to distinguish between tiny differences in the values  $V$ . At the other end of the spectrum is a decision-maker with  $\alpha \rightarrow 0$ , which leads to  $P(a_1) \rightarrow 0.5$  corresponding to an irrational agent that only produces random actions. In the remainder of the paper, we will refer to  $P(a_1)$  also as  $\lambda_1$  if chosen by player 1 and  $\lambda_2$  if chosen by player 2.

The expression for the value  $V$  also models the learning process of the parameter  $x$ . In the limit, when  $x$  is determined completely, then the distribution  $p(x|H, S)$  is going to approach a delta function in  $x$ . In that case, the integral collapses, and the free energy becomes equal to the expected pay-off. Fictitious play is therefore obtained in the limit of  $p(x|H, S) \rightarrow \delta(x - x^*)$  and  $\alpha \rightarrow \infty$ . Before this limit is reached, the distribution  $p(x|H, S)$  captures the uncertainty over the opponent, and the temperature parameter beta determines the risk sensitivity with respect to this distribution. In the infinitely risk-seeking limit  $\beta \rightarrow \infty$ , the decision-maker is so optimistic about the stag outcome that he will ignore any information to the contrary, and such a player will always cooperate independent of the history of the game. This is because

$$\begin{aligned} \lim_{\beta \rightarrow \infty} V(a_1) &= \max_{a_1} (xU(a_1, a_2 = S) + (1-x)U(a_1, a_2 = H)) \\ &= U(a_1, a_2 = S). \end{aligned}$$

Similarly, an infinitely risk-averse decision-maker ( $\beta \rightarrow -\infty$ ) is so pessimistic that he will only expect the worst-case scenario. This decision-maker will never cooperate independent of any experienced play. For any finite settings of  $\alpha$  and  $\beta$ , both cooperative and non-cooperative solutions can occur.

## 2.1. Model uncertainty and statistical physics

The central idea of having model uncertainty is that we do not fully trust our probabilistic model  $p_0(x)$  of a latent variable  $x$ . We therefore bias our estimates of  $x$  taking into account our utility function  $U(x)$ . If we are extremely pessimistic and cautious, for example, we will completely dismiss our probability model and simply assume a worst-case scenario. We then pick the action with the best worst-case scenario. If we fully trust our probability model, then we will pick the action with the highest *expected* utility. But if we are a risk-averse decision-maker with a finite amount of model

uncertainty, we compromise between the two extremes and bias our probability model towards the worst-case to some extent.

This decision-making scenario can be translated in terms of state changes in physical systems, where we start with a probability distribution  $p_0(x)$  and end up with a new distribution  $p(x)$ , because we have added an energy potential  $\Delta\phi(x)$  to the system. In this analogy, energy plays the role of a negative utility. In physics, a statistical system in equilibrium can be described by a Boltzmann distribution  $p_0(x) = 1/Z_0 e^{-\beta\phi_0(x)}$  with inverse temperature  $\beta = 1/kT$ , energy potential  $\phi_0(x)$  and partition sum  $Z_0$ . The distribution  $p_0$  is called an equilibrium distribution, because it minimizes the free energy

$$F_0[q] = \sum_x q(x)\phi_0(x) + \frac{1}{\beta} \sum_x q(x) \log q(x),$$

such that  $p_0 = \arg \min_q F[q]$  with  $F[p_0] = -1/\beta \log Z_0$ . If an energy potential  $\Delta\phi(x)$  is now added to the system, then the new equilibrium distribution that will arise is given by  $p(x) = 1/Z_1 e^{-\beta(\phi_0 + \Delta\phi(x))} = 1/Z_1 q(x) e^{-\beta\Delta\phi(x)}$ . This equilibrium distribution minimizes a free energy  $F_1[q]$

$$F_1[q] = \sum_x q(x)(\phi_0(x) + \Delta\phi(x)) + \frac{1}{\beta} \sum_x q(x) \log q(x).$$

The distribution  $p = \arg \min_q F_1[q]$  can be interpreted as the biased model. If the inverse temperature  $\beta$  is low, then  $p$  is going to be very similar to  $p_0$ , if the inverse temperature  $\beta$  is high, then  $p$  is going to be biased towards low-energy outcomes of the added potential  $\Delta\phi$ . In the KL-control setting [29–31],  $p_0$  is the equilibrium distribution resulting from the uncontrolled dynamics, whereas  $p$  corresponds to the controlled dynamics.

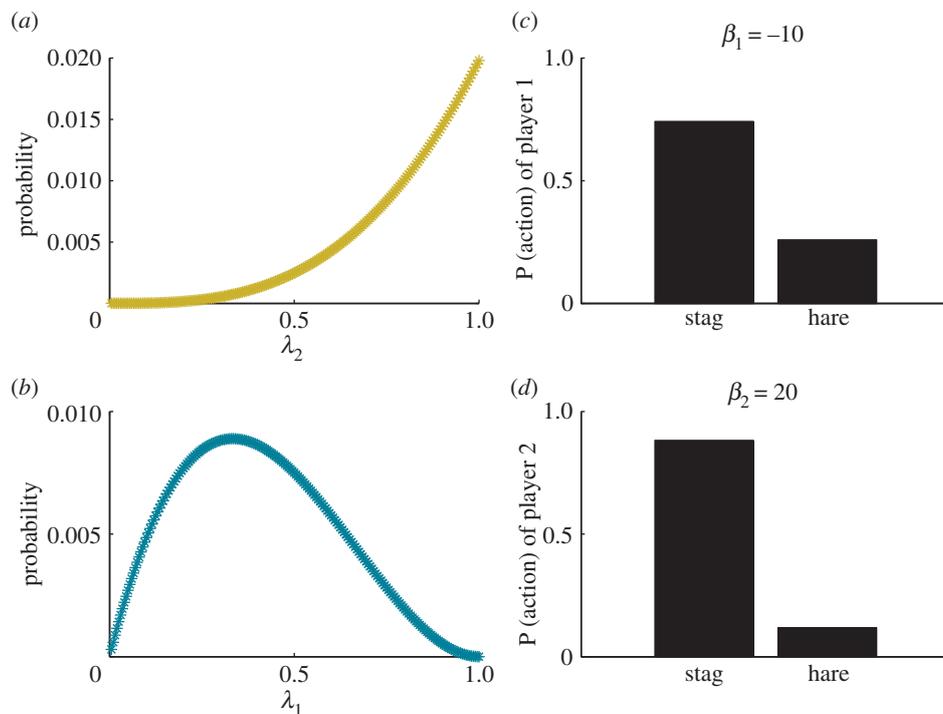
Both free energies can be combined into a free energy difference as a single variational principle such that

$$\begin{aligned} \Delta F[q] &= F_1[q] - F_0[p_0] \\ &= \sum_x q(x)\Delta\phi(x) + \frac{1}{\beta} \sum_x q(x) \log \frac{q(x)}{p_0(x)} \end{aligned}$$

and  $p = \arg \min_q \Delta F[q]$  such that  $\Delta F[p] = -1/\beta \log \sum_x q(x) e^{-\beta\Delta\phi(x)}$ . When replacing  $\Delta\phi(x) = -U(x)$ , we recognize in  $-\Delta F[q]$  the same variational principle as suggested in equation (1.1) to describe model uncertainty. This variational principle has recently been suggested as a principle for decision-making with information-processing costs [32–34]. Moreover, in non-equilibrium thermodynamics, the same expression for the free energy difference  $\Delta F[p]$  can be obtained from the Jarzynski equation for infinitely fast switching between the two states. Crucially, the Jarzynski equation holds for any switching process between the two states, and generalizes classical results for infinitely slow and fast switching [35]. When the utilities are negative log-likelihoods of outcomes under a generative model, this becomes the free energy principle that has recently been proposed to model action and perception in living systems trying to minimize surprise [36].

## 3. Simulation results

To illustrate the behaviour that arises when two decision-makers interact following equation (2.3), we simulated two model players with rationality parameter  $\alpha_1 = \alpha_2 = 10$  and risk-sensitive parameters  $\beta_1 = -10$  and  $\beta_2 = 20$  for player



**Figure 1.** Belief and action probabilities. Belief probability of player 1 (a) and player 2 (b) after observing three actions of the other player. Player 1 observed three cooperative actions and player 2 observed one cooperative and two non-cooperative actions. Accordingly, player 1 has more probability mass on the right half, whereas player 2 has more probability mass on the left half. Action probability of player 1 (c) and player 2 (d) resulting from the beliefs and the player's risk sensitivity. Player 1 has a higher probability to cooperate even though he is risk-averse, due to the strong evidence of cooperation. Player 2 also places high probability on cooperation, because he is strongly risk-seeking, even though the evidence points more towards a non-cooperative opponent. (Online version in colour.)

1 and 2, respectively. In figure 1, we depict beliefs and action probabilities of the two players after the pessimistic player 1 played stag once and hare twice, and the very optimistic player 2 played stag three times in a row. Accordingly, player 1's belief about player 2 is biased towards cooperative strategies (figure 1a), whereas player 2's belief about player 1 is biased towards non-cooperative strategies (figure 1b). Despite being risk-averse, player 1 has a higher probability for cooperation, given the strong evidence of cooperative behaviour of player 2. By contrast, player 2 has evidence of non-cooperativeness of player 1, but because he is optimistic, he most probably chooses to cooperate anyway. In figure 2, it can be seen how both players converge to a cooperative equilibrium after 25 interactions. In the left panel, the mean and standard deviations of the beta distribution beliefs of the two players are shown over the course of the 25 trials. It can be seen that both beliefs converge towards cooperative strategies, implying that both players believe in the cooperativeness of the other player. In the right panel, the action probabilities of choosing stag for both players are shown. Both action probabilities converge to cooperative strategies.

In the bottom row of figure 3, we show the probability of a cooperative equilibrium after 25 interactions depending on all possible combinations of risk sensitivities of the two players ranging from risk-averse ( $\beta = -20$ ) to risk-seeking ( $\beta = +20$ ). In this simulation, the rationality of player 1 was always set to  $\alpha_1 = 10$ , whereas the rationality of player 2 was set to  $\alpha_2 = 2$  (right panels) or  $\alpha_2 = 10$  (left panels). The prior probability of cooperation before any interaction has taken place is shown in the upper panels. For uninformative priors, the probability of cooperation in the first trial is greater than one half

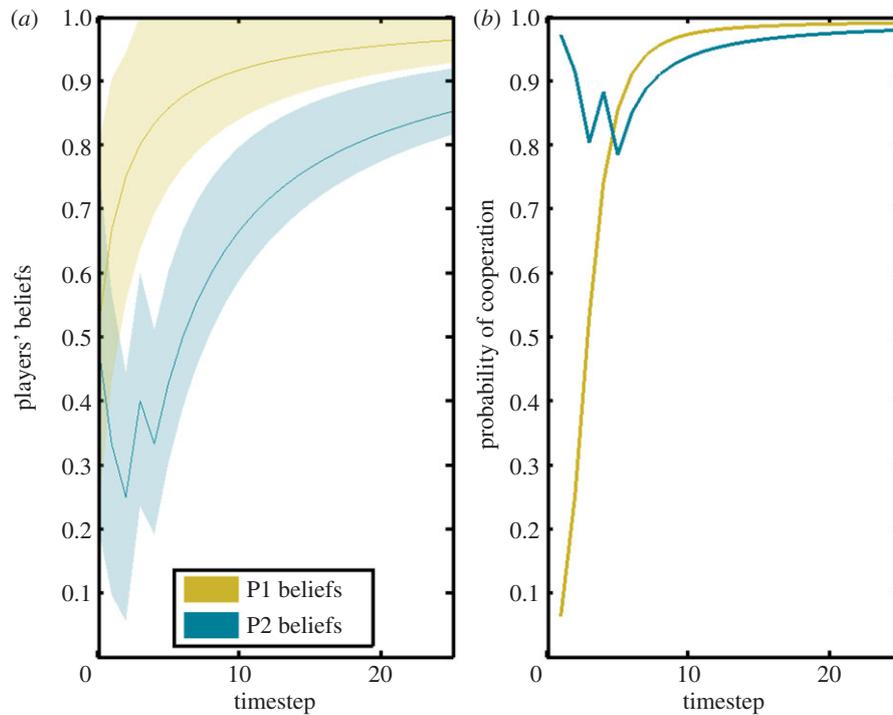
for all risk-seeking decision-makers and lower than one half for all risk-averse decision-makers independent of the opponent's risk sensitivity. Naturally, in later interactions, the opponent's risk sensitivity comes to bear. If both players have positive risk sensitivities, then there is a higher probability they will end up cooperating, and similarly if both players have negative risk sensitivities there is an increased probability they will end up with a non-cooperative equilibrium. If one of the players is risk-seeking and the other one risk-averse, then the player whose risk sensitivity has higher absolute value will more probably drive the behaviour of the interaction towards cooperation if risk-seeking or non-cooperation if risk-averse. If player 2 has a low rationality  $\alpha_2 = 2$ , the overall pattern is similar, but more noisy.

## 4. Experimental methods

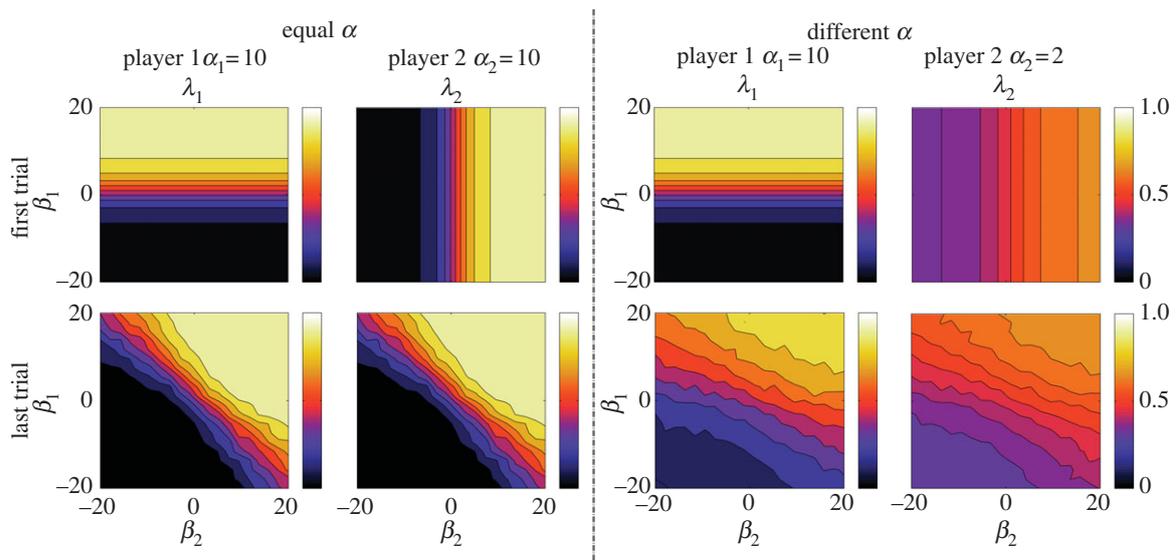
To investigate the effect of risk sensitivity in sensorimotor interactions in human subjects, we used a previously developed virtual reality paradigm to translate  $2 \times 2$  matrix games into sensorimotor games [21,22]. One of the players was always simulated by a virtual agent modelled by equation (2.3). This way, we could directly manipulate the risk sensitivity of the virtual player and record subjects' responses to these changes.

### 4.1. Experimental design

As illustrated in figure 4b, participants held the handle of a robotic interface with which they could control the position of a cursor on a display. On each trial, participants had to move the cursor from a start bar to a target bar and back.



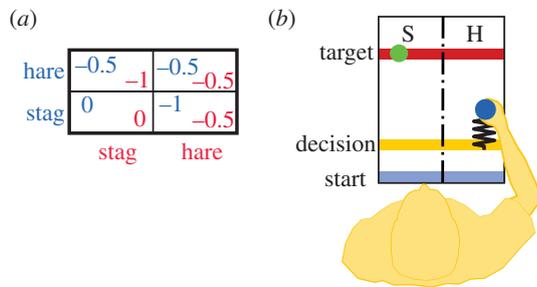
**Figure 2.** Evolution of (a) belief and (b) action probabilities over 25 trials. (a) Mean and standard deviation of the beta distribution reflecting each player's beliefs. (b) Action probabilities of the players according to equation (2.3). The third trial corresponds to the beliefs and actions displayed in figure 1. (Online version in colour.)



**Figure 3.** Probability of choosing the cooperative action for both players with different risk-sensitivities and rationality parameters. (top) Prior probability of cooperating. In the first trial, the probability of cooperation only depends on the risk sensitivity of the player and does not depend on the risk sensitivity of the opponent. (bottom) Probability of cooperating after 25 trials. In later trials, the probability of cooperation depends on the risk sensitivity of both the player and the opponent. (first two columns) Probability of cooperation when both players have equal rationality  $\alpha$ . (last two columns) Probability of cooperation when players have different rationality  $\alpha$ . The probability of cooperating was computed according to equation (2.3). (Online version in colour.)

Importantly, they could do so choosing any lateral position within the width of the target bar. Therefore, participants could achieve the task with their final hand position anywhere between the left and right target bounds. During the forth-and-back movement to the target, subjects had to cross a yellow decision line at 3 cm into the movement. Once the line was crossed, both the subject's and the virtual player's decisions were made. The left half of the subject's lateral workspace represented the cooperative stag solution, whereas the right half represented the non-cooperative hare solution.

An implicit pay-off was placed on the movements beyond the decision line by using the robot to generate a resistive force opposing the forward motion of the handle. The forces were generated by simulating springs that acted between the handle and the yellow decision bar. The stiffness of the spring during the movement depended on the lateral position of the handle at the time of crossing the decision line and the computer player's choice. The spring constant was determined by the pay-off indicated in figure 4a and multiplied by a constant factor of  $1.9 \text{ N cm}^{-1}$ . For successful trial completion, the target bar had to be reached within



**Figure 4.** The sensorimotor stag-hunt game. (a) Pay-off matrix of the game. (b) Experimental methods. Subjects had to move a cursor from the start bar to the target bar. The left half of the workspace corresponded to selecting 'stag', and the right half of the workspace corresponded to selecting 'hare'. Once they crossed, the decision line a circle on the target line indicated the choice of the virtual player, and subjects experienced a force opposing their forward movement that depended on their selection and the virtual opponent's action selection which followed equation (2.3) and subjects experienced a force opposing their forward movement that depended on their action selection and the virtual opponent's action selection as indicated by the pay-off matrix. (Online version in colour.)

1200 ms. The distance of the target bar from the start bar was sampled randomly each trial from a uniform distribution between 15 and 25 cm. Subjects performed two sessions where they faced virtual players with two different rationality parameters. In the first session, the rationality of the virtual player was  $\alpha_2 = 10$  and subjects performed 40 sets of 25 trials, where the virtual player could assume one of five different  $\beta_2$ -values from the set  $[\pm 20, \pm 10, 0]$ . At the beginning of each set, the  $\beta_2$ -parameter of the virtual player was determined and remained constant throughout the set. Each  $\beta_2$ -parameter was chosen eight times, but in randomized order. In the second session, the rationality of the virtual player was set to  $\alpha_2 = 2$  and subjects performed again 40 sets of 25 trials each with different  $\beta$ -parameters. At the start of every session, they had between 100 and 125 training trials where they could see the degree of risk sensitivity of the virtual player displayed on a bar.

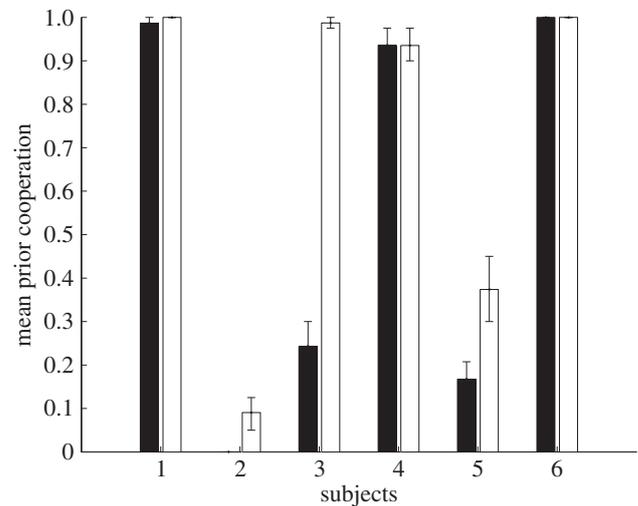
## 4.2. Experimental apparatus

The experiments were conducted using a planar robotic manipulandum (vBOT) [37]. Participants held a handle of the vBOT, which constrained hand movements to the horizontal plane. A virtual reality system was used to overlay visual feedback onto the plane of movement and players were prevented from seeing their own hand. The vBOT allowed us to record the position of the handle and to generate forces on the hand with a 1 kHz update rate.

## 4.3. Participants

Six naive participants from the student pool of the Eberhard-Karls-Universität Tübingen participated in the study. All experimental procedures were approved by the ethics committee of the medical faculty at the University of Tübingen.

The precise instructions given to subjects are described below. Subjects were told that they were playing a game against a virtual player and that they could choose between two actions in every trial: either to cooperate or not to cooperate. They were instructed to make their choice by moving the handle across the decision line either in the



**Figure 5.** Prior cooperation probabilities in human subjects playing virtual opponents with high ( $\alpha_2 = 10$ , white colour) and low ( $\alpha_2 = 2$ , black colour) rationality. In the first trial, when facing a new opponent, subjects knew the rationality of the opponent, but not their risk sensitivity.

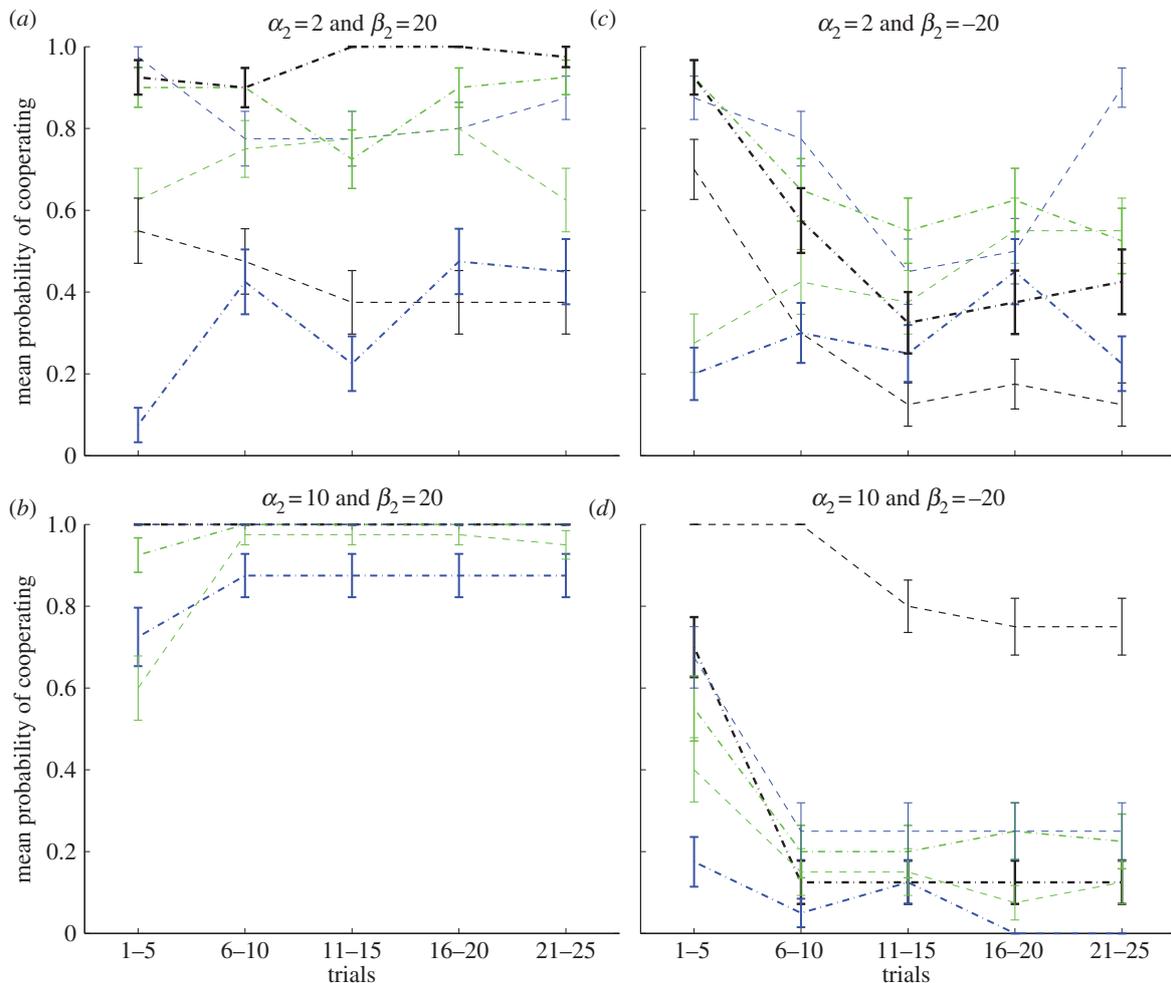
right or left half of the workspace and that the left half corresponded to cooperation, whereas the right half corresponded to non-cooperation. They were also informed that there would be a force opposing their movement between the decision line and the target line. They were told that in the case of non-cooperation they would always experience the same medium force, but that in the case of cooperation the force would depend on the choice of the virtual player, who could choose to cooperate or not to cooperate. In case both players cooperate, there would be no force, but if the virtual player chooses not to cooperate there would be a very high force. Subjects were also told that the virtual player can learn and adapt to the subject's play.

At the beginning of each block of training trials, subjects could see a bar displaying the degree of the virtual player's risk sensitivity and they were told that the bar indicates the virtual player's attitude towards cooperation. They were also told that there was a different player with a different attitude every 25 trials. After the training trials, they were told that the bar would be no longer displayed and that they can learn the player's attitude towards cooperation only from actual play. Between blocks of 25 trials, there was a short break to mark the transition between different virtual players clearly.

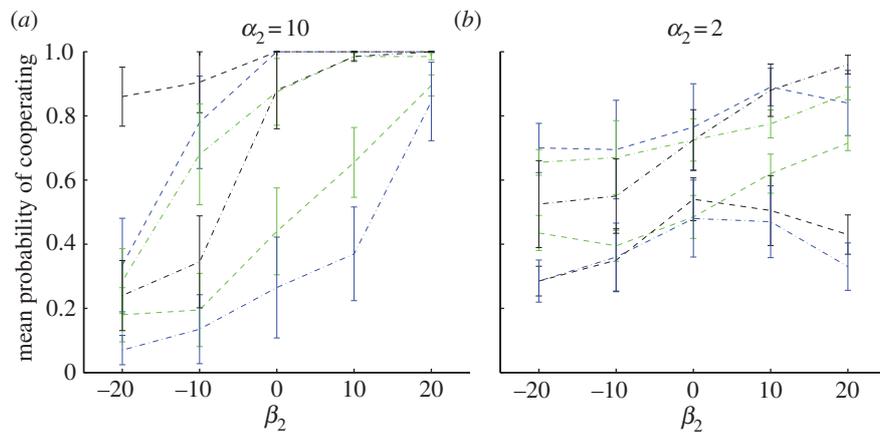
## 5. Results

In figure 5, we show subjects' prior cooperation probabilities in the first trial of every set of 25, when they face a novel virtual player. This is shown in white for virtual players with rationality  $\alpha_2 = 10$  and, in black for virtual players with rationality  $\alpha_2 = 2$ . In the  $\alpha_2 = 10$  condition, we found that four of six subjects chose to cooperate most of the time in the first trial. In the  $\alpha_2 = 2$  condition, only three out of six subjects chose to cooperate. This implies that about half of our subjects were risk-seeking and optimistic about cooperation, whereas the others were risk-averse and pessimistic.

After the first trial, subjects received feedback about the choice of the virtual player and could make a first inference



**Figure 6.** Evolution of cooperation over the course of 25 trials of human subjects facing virtual opponents with low ( $\alpha_2 = 2$ , (a,c)) or high ( $\alpha_2 = 10$ , (b,d)) rationality and positive ( $\beta_2 = 20$ , (a,b)) and negative ( $\beta_2 = -20$ , (c,d)) risk sensitivity. Different lines indicate different subjects. (Online version in colour.)

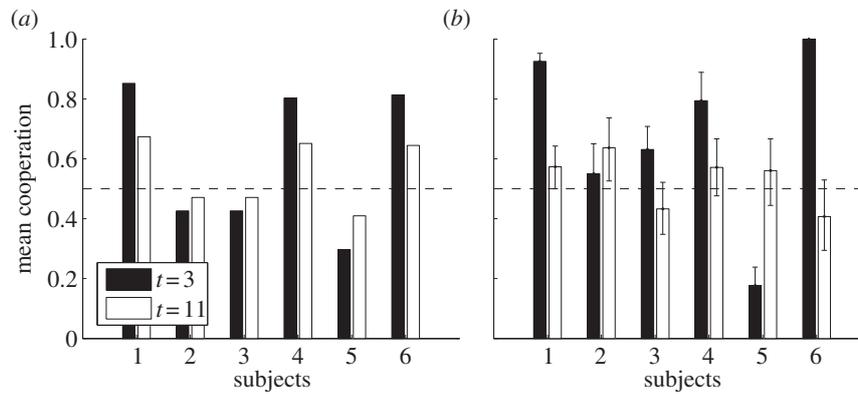


**Figure 7.** (top row) Average probability of cooperation depending on risk sensitivity of the opponent with either high ( $\alpha_2 = 10$ , (a)) or low ( $\alpha_2 = 2$ , (b)) rationality for the six different subjects. Different lines indicate different subjects. (Online version in colour.)

about the virtual player's willingness to cooperate. Accordingly, subjects' probability of cooperation in subsequent trials in a set of 25 needs to be investigated separately for the different risk sensitivities of the virtual players. For the extreme risk sensitivities of  $\beta_2 = 20$  and  $\beta_2 = -20$ , this is depicted in figure 6. When playing a risk-averse opponent ( $\beta_2 = -20$ ), subjects mostly converged to non-cooperative behaviour (figure 6c,d), whereas when playing a risk-seeking opponent ( $\beta_2 = 20$ ), subjects mostly converged to cooperative behaviour (figure 6a,b). This pattern is clearly demonstrated when facing virtual players with high rationality  $\alpha_2 = 10$  (figure 6b,d), but

much more diffuse in the case of virtual players with low rationality  $\alpha_2 = 2$  (figure 6a,c).

To directly assess the effect of risk sensitivity on the cooperative behaviour of human subjects over all trials, we computed the mean probability of cooperation averaged over all trials where the opponent had the same risk sensitivity  $\beta_2$  and rationality  $\alpha_2$ . In figure 7, this is shown for all six subjects playing an opponent with rationality  $\alpha_2 = 10$  (figure 7a) and rationality  $\alpha_2 = 2$  (figure 7b), respectively. For both rationalities, the risk sensitivity  $\beta_2$  of the opponent has a significant effect on the probability of cooperation



**Figure 8.** Comparison of risk-sensitive predictions to fictitious play and human subjects' behaviour. (a) Predictions of probability of cooperation when observing a sequence with 50% cooperation after two trials (filled bar) or 10 trials (open bar). The dashed line is the prediction of fictitious play. (b) Subjects' cooperation probabilities when observing a sequence with roughly 50% cooperation after two trials (black) or 10 trials (white).

(non-parametric Jonckheere–Terpstra trend test  $p < 0.05$  for  $\alpha_2 = 2$  and  $p < 0.001$  for  $\alpha_2 = 10$ ). However, in the case of high rationality  $\alpha_2$  of the virtual player, this effect is stronger and clearer than in the case of inconsistent play resulting from an opponent with low  $\alpha_2$ . The general trend is that subjects' tendency to cooperate increases for higher  $\beta_2$  and decreases for lower  $\beta_2$ . Importantly, most subjects deviated on average from a 50:50 cooperation probability when playing a risk-neutral opponent of high rationality ( $\alpha_2 = 10$ ), which is another signature of subjects' risk sensitivity.

To compare the predictive power of our model with the traditional fictitious play model, we investigated the ratio of cooperation after subjects had experienced an (approximately) 50:50 sequence of actions of the virtual player, i.e. the opponent had (roughly) cooperated half the time and refused cooperation the other half of the time. Importantly, we did this at two different stages of the game such that the 50:50 ratio was the result of either a small number of trials (after two trials) or a large number of trials (after 10 trials). In the two-trial case, only trials with one stag- and one hare-choice were included, however, for the 10-trial case, there were not enough instances with an exact 50:50 ratio. Therefore, we also included trials between 40% and 60% of cooperation, but still this analysis was only possible in the case of a virtual player with low rationality ( $\alpha_2 = 2$ ). The crucial observation is that in the case of two trials the estimate of the other player's cooperation is highly uncertain, whereas in the case of 10 trials, this estimate is much more consolidated. In both cases, fictitious play makes the same prediction, which is the best response to the ratio—compare dashed line in figure 8. By contrast, a risk-sensitive model predicts that the best response should depend on the uncertainty of the estimate of the ratio. For our model predictions, we fitted to each subject an  $\alpha_1$ - and a  $\beta_1$ -parameter by maximizing the log-likelihood of subjects' choices given the predicted choice probabilities of equation (2.3). In particular, this predicts that a risk-seeking player will deviate towards cooperation in early trials, whereas a risk-averse player will deviate towards non-cooperation in early trials—compare figure 8a. In late trials, when a large part of the uncertainty has been removed, both players converge to fictitious play.

In figure 8b, it can be seen that most subjects' behaviour was inconsistent with fictitious play. Subjects 1, 4 and 6 were risk-seeking and deviated significantly towards cooperation in the third trial (one-sided  $t$ -test  $p < 0.01$ ). Subject 5 was risk-averse and refused cooperation in early trials ( $p < 0.01$ ).

Subjects 2 and 3 were risk-neutral and consistent with fictitious play and therefore the deviation from 0.5 choice probability was not significant ( $p > 0.1$ ). Importantly, after 10 trials, all subjects were consistent with fictitious play and were best-responding to the observed sequence of the opponent's play, hence the deviation from 0.5 choice probability was not significant for all of them ( $p > 0.1$ ).

## 6. Discussion

Most current theoretical frameworks of motor control rely on probabilistic models that are used for prediction, estimation and control. However, when such models are partially incorrect or wrong, there are usually no performance guarantees [17]. Model uncertainty is therefore an important factor in real-world control problems, because, in practice, one can never be absolutely sure about one's model. In this paper, we investigated risk-sensitive deviations arising from having model uncertainty in sensorimotor interactions. We found that human subjects adapted their cooperation depending on the risk sensitivity of a virtual computer player. Furthermore, we found that subjects did not only best-respond to the frequency of observed play, but that they were sensitive to the certainty of this estimate. In particular, they allowed for risk-sensitive deviations in initial interaction trials when uncertainty was high. This behaviour is consistent with a risk-sensitive decision-maker with model uncertainty.

Recently, it was found that risk sensitivity is an important determinant in human sensorimotor behaviour [38]. Risk-sensitive decision-makers do not base their choices exclusively on the expectation value of a particular cost function, but they also consider higher-order moments of this cost function. This can be seen when approximating the risk-sensitive cost function with a Taylor series

$$J = \frac{1}{\beta} \log \int dx p(x) e^{\beta U(x)} \\ \approx \mathbb{E}[U] + \frac{\beta}{2} \text{VAR}[U],$$

assuming that  $\beta \text{VAR}[U]$  is small [39]. Sensitivity to the second-order moment of the cost function was found, for example, in motor tasks with speed–accuracy trade-off [40]. Such risk-sensitive decision-makers can be thought of as trading off the mean cost versus the variability of the cost. A mean–variance trade-off in effort was found, for example, in a motor task where

subjects had to decide between hitting differently sized targets that were associated with different levels of effort [41]. Sensitivity to the variance of the control cost was also found in continuous motor tasks, where subjects had to control a cursor undergoing a random walk [42]. The sensitivity to the variance can also be exploited by assistive technologies that consider the human as a (useful) noise source [43,44].

When  $x$  is a latent variable that needs to be inferred, risk sensitivity also allows decision-makers to take model uncertainty into account. This can be seen when rewriting the risk-sensitive cost function as in equation (1.1) yielding

$$J = \frac{1}{\beta} \log \int dx p(x) e^{\beta U(x)}$$

$$= \text{ext}_{p(x)} \left[ \int dx p(x) U(x) - \frac{1}{\beta} \int dx p(x) \log \frac{p(x)}{p_0(x)} \right],$$

where  $J$  can be re-expressed as a variational principle that trades off the maximization of a utility term and the deviation from  $p_0$  to  $p$  [17]. Such model uncertainty was recently found to play a role in a sensorimotor integration task, where subjects had to infer the position of a hidden target (the latent variable) [20]. When given feedback information about the target position with varying degree of reliability, subjects' estimates of the target position was consistent with a Bayesian estimator that optimally combines prior knowledge of the distribution of target positions with the actual feedback information. Subjects' behaviour was therefore also consistent with previous reports on information integration in sensorimotor tasks [9]. However, when subjects' beliefs were associated with control costs, study [20] found that subjects exhibited characteristic deviations from the Bayes optimal response that could be described by a risk-sensitive decision-making model that depended on the level of model uncertainty, the reliability of the feedback and the control cost. These risk-sensitive deviations were particularly prominent in trials with high uncertainty and vanished in the absence of uncertainty as more and more information about the latent variable becomes available.

In the context of model uncertainty, risk sensitivity can be distinguished from risk attitudes modelled by the curvature of the utility function, both theoretically and experimentally [45,46]. Utility functions generally express the subjective desirability of an outcome and not necessarily its nominal value. For example, the subjective value of money does typically not increase linearly with the nominal amount. Accordingly, receiving a monetary increase of \$1000 has more utility for a beggar than for a millionaire. The utility function is said to be marginally decreasing. Intriguingly, this property can also be used to model risk attitudes. For example, people with a marginally decreasing utility function of money will prefer \$50 for sure over a gamble between a 50:50 lottery, where one outcome is \$0 and the other is \$100, because  $U(\$50) > 1/2U(\$100)$ , assuming that  $U(\$0) = 0$ . Importantly, these risk attitudes are independent of the level of information about the probabilities. In fact, the probabilities are assumed to be perfectly known. Thus, risk attitudes are conceptually very different from model uncertainty that vanishes in the limit of perfect information about the probabilities. Model uncertainty captures the lack of information about a lottery.

The effect of risk attitudes on cooperation in the stag-hunt game is investigated in behavioural economics tasks [47–49] in which the risk attitude of subjects is determined by subjects' choice behaviour when deciding between risky and safe lotteries. In these studies, it was found that subjects'

risk attitude does not predict their cooperation in the stag-hunt game, although players consider information about the other player's risk attitude. In particular, subjects are less likely to cooperate if they know that their opponent is risk-averse. However, the fact that subjects' risk attitude is a poor predictor of their cooperation in the game suggests that not risk attitude, but model uncertainty might be a stronger factor affecting cooperation in the game.

In the traditional stag-hunt game, pay-offs are usually framed as gains, whereas, in our experiment, the pay-offs are framed as losses in the shape of forces subjects have to exert. In the economics literature, it is well known that the framing of losses versus gains can have a strong influence on human choice behaviour [50]. It is therefore not surprising that different pay-off levels have also been found to influence choice behaviour in the stag-hunt game [51]; in particular, it was found that having losses increases players' probability of choosing the more risky stag. Crucially, our results showing sensitivity to model uncertainty do not depend on the exact shape of the utility function. Expected utility players that have experienced 50:50 play of their opponent after  $N$  amount of trials will choose between  $a_1 = S$  and  $a_1 = H$  according to equation (2.2), where  $x^* = 0.5$ . The decision-maker's preference depends of course on the utilities  $U(a_1, a_2)$ , but crucially these utilities and the resulting expected utility do not change with varying the amount of trials  $N$  as long as the empirical frequency is 50:50. The fact that we have used a loss scenario does therefore not invalidate our results on model uncertainty, although the exact choice probabilities might look different in a gain scenario.

Fictitious play is one of the earliest models that were developed to explain learning in games [25,52]. Crucially, it assumes stationary strategies for both players. It can be shown to converge for a wide class of problems, including all two-player interactions [53]. However, it can also be shown that fictitious play can lead to non-converging limit-cycles for very simple games [54]. In our study, we found that subjects were not simply best-responding to the observed frequency of the opponent's play, as presumed by fictitious play. Rather, subjects were sensitive to the amount of information they had gathered about the other player when deciding whether to cooperate or not—compare figure 8. Our risk-sensitive model of cooperation can account for this dependency. However, it still makes the simplistic assumption of stationary strategy beliefs. This limitation may be overcome in the future by considering more complex belief models.

An important objection to risk-sensitive models is often that they could be replaced by a standard risk-neutral Bayesian model under a different (post hoc) prior [17]. This is also true in our case: subjects could develop biased prior beliefs about the population of virtual players. Importantly, the population of virtual players was statistically balanced and there is therefore no statistical reason why subjects should develop biased priors. However, if the prior is thought to reflect not only the (prior) statistics of the environment but also traits of the decision-maker, then a risk-neutral Bayesian model with a biased prior could, in principle, also explain our data. This is sometimes also discussed in the context of so-called complete class theorems, in which the existence of priors is investigated when modelling Bayesian decision-makers with different loss functions [55,56].

The results of our study also speak to cognitive theories of (dyadic) social interactions and joint actions. Several recent

studies have investigated how humans mutually adjust and synchronize their behaviour during online joint actions, revealing the role of several mechanisms that range from automatic entrainment to action prediction [22,57–59]. An open research question is if and how sensorimotor interactions are influenced by the co-actors' goals and attitudes. Given that socially and culturally relevant information (e.g. facial expression, racial or social group membership) is automatically processed in the brain [60] and can automatically modulate imitation [61] and empathy [62], most studies have focused on the impact of socially relevant variables in joint actions, with the hypothesis that it could favour pro-social or anti-social behaviour. It has been shown that interpersonal perception and (positive and negative) attitude towards the co-actor modulate cooperation and joint actions [63,64]. In turn, sensorimotor interactions can modulate a co-actors' attitude; for example, it has been reported that dyads engaged in synchronous interactions improve their altruistic behaviour [65].

The aforementioned studies focus on social attitudes and leave unanswered the issue of how personal traits and non-

social attitudes influence sensorimotor interactions. Here, we studied the influence of model uncertainty on the evolution of sensorimotor interactions. We designed a sensorimotor task that is equivalent to the stag-hunt game. Our results show that model uncertainty modulates sensorimotor interactions and their success. In particular, optimistic (risk-seeking) adaptive agents are much more likely to converge to cooperative outcomes. Furthermore, humans adaptively change their level of cooperation depending on the risk sensitivity of their co-actor (in our study, a computer player). Effects of model uncertainty are particularly strong in early interactions with a novel player. In summary, our results indicate that interacting agents can build sophisticated models of their co-actors [66] and use them to modulate their level of cooperation taking model uncertainty into account.

**Funding statement.** This study was supported by the DFG, Emmy Noether grant no. BR4164/1-1 and EU's FP7 under grant agreement no. FP7-ICT-270108 (goal-leaders).

## References

- Shadmehr R, Mussa-Ivaldi FA. 1994 Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.* **14**, 3208–3224.
- Wolpert DM, Ghahramani Z, Jordan MI. 1995 An internal model for sensorimotor integration. *Science* **269**, 1880–1882. (doi:10.1126/science.7569931)
- Flanagan JR, Wing AM. 1997 The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. *J. Neurosci.* **17**, 1519–1528.
- Blakemore SJ, Goodbody SJ, Wolpert DM. 1998 Predicting the consequences of our own actions: the role of sensorimotor context estimation. *J. Neurosci.* **18**, 7511–7518.
- Kawato M. 1999 Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* **9**, 718–727. (doi:10.1016/S0959-4388(99)00028-8)
- Mehta B, Schaal S. 2002 Forward models in visuomotor control. *J. Neurophysiol.* **88**, 942–953.
- Ernst MO, Banks MS. 2002 Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433. (doi:10.1038/415429a)
- Knill DC, Pouget A. 2004 The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719. (doi:10.1016/j.tins.2004.10.007)
- Körding KP, Wolpert DM. 2004 Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247. (doi:10.1038/nature02169)
- Doya K (ed.) 2007 *Bayesian brain: probabilistic approaches to neural coding*. Cambridge, MA: MIT Press.
- Todorov E, Jordan MI. 2002 Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* **5**, 1226–1235. (doi:10.1038/nn963)
- Trommershäuser J, Maloney LT, Landy MS. 2003 Statistical decision theory and trade-offs in the control of motor response. *Spat. Vis.* **16** 255–275. (doi:10.1163/156856803322467527)
- Trommershäuser J, Maloney LT, Landy MS. 2003 Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. A* **20**, 1419–1433. (doi:10.1364/JOSA.A.20.001419)
- Todorov E. 2004 Optimality principles in sensorimotor control. *Nat. Neurosci.* **7**, 907–915. (doi:10.1038/nn1309)
- Trommershäuser J, Maloney LT, Landy MS. 2008 Decision making, movement planning and statistical decision theory. *Trends Cogn. Sci.* **12**, 291–297. (doi:10.1016/j.tics.2008.04.010)
- Wu S-W, Delgado MR, Maloney LT. 2009 Economic decision-making compared with an equivalent motor task. *Proc. Natl Acad. Sci. USA* **106**, 6088–6093. (doi:10.1073/pnas.0900102106)
- Hansen LP, Sargent TJ. 2008 *Robustness*. Princeton, NJ: Princeton University Press.
- Peter W. 1990 *Risk-sensitive optimal control*. London, UK: John Wiley & Sons.
- Maccheroni F, Marinacci M, Rustichini A. 2006 Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* **74**, 1447–1498. (doi:10.1111/j.1468-0262.2006.00716.x)
- Grau-Moya J, Ortega PA, Braun D. 2012 Risk-sensitivity in Bayesian sensorimotor integration. *PLoS Comput. Biol.* **8**, e1002698. (doi:10.1371/journal.pcbi.1002698)
- Braun DA, Ortega PA, Wolpert DM. 2009 Nash equilibria in multi-agent motor interactions. *PLoS Comput. Biol.* **5**, e1000468. (doi:10.1371/journal.pcbi.1000468)
- Braun DA, Ortega PA, Wolpert DM. 2011 Motor coordination: when two have to act as one. *Exp. Brain Res.* **211**, 631–641. (doi:10.1007/s00221-011-2642-y)
- Osborne MJ, Rubinstein A. 1999 *A course in game theory*. Cambridge, MA: MIT Press.
- Weibull JW. 1995 *Evolutionary game theory*. Cambridge, MA: MIT Press.
- Brown GW. 1951 Iterative solutions of games by fictitious play. In *Activity analysis of production and allocation*. London, UK: Wiley.
- Krishna V, Sjöström T. 1998 On the convergence of fictitious play. *Math. Oper. Res.* **23**, 479–511. (doi:10.1287/moor.23.2.479)
- Berger U. 2007 Brown's original fictitious play. *J. Econ. Theory* **135**, 572–578. (doi:10.1016/j.jet.2005.12.010)
- McKelvey RD, Palfrey TR. 1995 Quantal response equilibria for normal form games. *Games Econ. Behav.* **10**, 6–38. (doi:10.1006/game.1995.1023)
- Kappen HJ. 2005 A linear theory for control of non-linear stochastic systems. *Phys. Rev. Lett.* **95**, 200201. (doi:10.1103/PhysRevLett.95.200201)
- Todorov E. 2009 Efficient computation of optimal actions. *Proc. Natl Acad. Sci. USA* **106**, 11 478–11 483. (doi:10.1073/pnas.0710743106)
- Kappen HJ, Gmez V, Opper M. 2012 Optimal control as a graphical model inference problem. *Mach. Learn.* **87**, 159–182. (doi:10.1007/s10994-012-5278-7)
- Braun DA, Ortega PA, Theodorou E, Schaal S. 2011 Path integral control and bounded rationality. In *IEEE Symp. Adaptive Dynamic Programming and Reinforcement Learning, Paris, France, 11–15 April 2011*, pp. 202–209. New York, NY: IEEE.
- Ortega PA, Braun DA. 2011 Information, utility and bounded rationality. In *Lecture notes on artificial intelligence*, vol. 6830, pp. 269–274. Berlin, Germany: Springer.
- Ortega PA, Braun DA. 2013 Thermodynamics as a theory of decision-making with information-processing costs. *Proc. R. Soc. A* **469**, 20120683. (doi:10.1098/rspa.2012.0683)
- Jarzynski C. 1997 Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2690–2693. (doi:10.1103/PhysRevLett.78.2690)

36. Friston K. 2010 The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138. (doi:10.1038/nrn2787)
37. Howard IS, Ingram JN, Wolpert DM. 2009 A modular planar robotic manipulandum with end-point torque control. *J. Neurosci. Methods* **181**, 199–211. (doi:10.1016/j.jneumeth.2009.05.005)
38. Braun DA, Nagengast AJ, Wolpert D. 2011 Risk-sensitivity in sensorimotor control. *Front. Hum. Neurosci.* **5**, 1. (doi:10.3389/fnhum.2011.00001)
39. Whittle P. 1981 Risk-sensitive linear/quadratic/Gaussian control. *Adv. Appl. Probab.* **13**, 764–777. (doi:10.2307/1426972)
40. Nagengast AJ, Braun DA, Wolpert DM. 2011 Risk sensitivity in a motor task with speed-accuracy trade-off. *J. Neurophysiol.* **105**, 2668–2674. (doi:10.1152/jn.00804.2010)
41. Nagengast AJ, Braun DA, Wolpert DM. 2011 Risk-sensitivity and the mean-variance trade-off: decision making in sensorimotor control. *Proc. Biol. Sci.* **278**, 2325–2332. (doi:10.1098/rspb.2010.2518)
42. Nagengast AJ, Braun DA, Wolpert DM. 2010 Risk-sensitive optimal feedback control accounts for sensorimotor behavior under uncertainty. *PLoS Comput. Biol.* **6**, e1000857. (doi:10.1371/journal.pcbi.1000857)
43. Medina JR, Lee D, Hirche S. 2012 Risk-sensitive optimal feedback control for haptic assistance. In *2012 IEEE Int. Conf. on Robotics and Automation (ICRA), St. Paul, MN, 14–18 May 2012*, pp. 1025–1031. New York, NY: IEEE.
44. Saida M, Medina JR, Hirche S. 2012 Adaptive attitude design with risk-sensitive optimal feedback control in physical human-robot interaction. In *RO-MAN, 2012 IEEE, Paris, France, 9–13 September*, pp. 955–961. New York, NY: IEEE.
45. Gilboa I, Marinacci M. 2011 Ambiguity and the Bayesian paradigm. Technical report. Universita Bocconi. See [http://www.tau.ac.il/~igilboa/pdf/Gilboa\\_Marinacci\\_Ambiguity\\_and\\_Bayesian\\_Paradigm.pdf](http://www.tau.ac.il/~igilboa/pdf/Gilboa_Marinacci_Ambiguity_and_Bayesian_Paradigm.pdf).
46. Chakravarty S, Roy J. 2009 Recursive expected utility and the separation of attitudes towards risk and ambiguity: an experimental study. *Theory Decis.* **66**, 199–228. (doi:10.1007/s11238-008-9112-4)
47. Neumann T, Vogt B. 2009 Do players' beliefs or risk attitudes determine the equilibrium selections in  $2 \times 2$  coordination games. FEMM Working Paper No. 24, August 2009.
48. Al-Ubaydli O, Jones G, Weel J. 2013 Patience, cognitive skill and coordination in the repeated stag-hunt. *J. Neurosci. Psychol. Econom.* **6**, 71–96. (doi:10.1037/npe0000005).
49. Büyükboyacı M. In preparation. Risk attitudes and the stag hunt game. See <http://hss.caltech.edu/~muruvvet/staghunt.pdf>.
50. Kahneman D, Tversky A. 1979 Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–291. (doi:10.2307/1914185)
51. Feltovich N, Iwasaki A, Oda SH. 2012 Payoff levels, loss avoidance, and equilibrium selection in games with multiple equilibria: an experimental study. *Econ. Inq.* **50**, 932–952. (doi:10.1111/j.1465-7295.2011.00406.x)
52. Fudenberg D, Levine DK. 1998 *The theory of learning in games*. Cambridge, MA: MIT Press.
53. Berger U. 2005 Fictitious play in  $2 \times n$  games. *J. Econ. Theory* **120**, 139–154. (doi:10.1016/j.jet.2004.02.003)
54. Shapley L. 1964 Some topics in two-person games. In *Advances in game theory*. Princeton, NJ: Princeton University Press.
55. Brown LD. 1981 A complete class theorem for statistical problems with finite sample spaces. *Ann. Stat.* **9**, 1289–1300. (doi:10.1214/aos/1176345645)
56. Friston K, Samothrakis S, Montague R. 2012 Active inference and agency: optimal control without cost functions. *Biol. Cybernet.* **106**, 523–541. (doi:10.1007/s00422-012-0512-8)
57. Pezzulo G, Dindo H. 2011 What should I do next? using shared representations to solve interaction problems. *Exp. Brain Res.* **211**, 613–630. (doi:10.1007/s00221-011-2712-1)
58. Sebanz N, Bekkering H, Knoblich G. 2006 Joint action: bodies and minds moving together. *Trends Cogn. Sci.* **10**, 70–76. (doi:10.1016/j.tics.2005.12.009)
59. Vesper C, van der Wel RP, Knoblich G, Sebanz N. 2013 Are you ready to jump? predictive mechanisms in interpersonal coordination. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 48–61. (doi:10.1037/a0028066)
60. Cosmides L, Tooby J, Kurzban R. 2003 Perceptions of race. *Trends Cogn. Sci.* **7**, 173–179. (doi:10.1016/S1364-6613(03)00057-3)
61. Losin EAR, Iacoboni M, Martin A, Cross KA, Dapretto M. 2012 Race modulates neural activity during imitation. *Neuroimage* **59**, 3594–3603. (doi:10.1016/j.neuroimage.2011.10.074)
62. Avenanti A, Sirigu A, Aglioti SM. 2010 Racial bias reduces empathic sensorimotor resonance with other-race pain. *Curr. Biol.* **20**, 1018–1022. (doi:10.1016/j.cub.2010.03.071)
63. Iani C, Anelli F, Nicoletti R, Arcuri L, Rubichi S. 2011 The role of group membership on the modulation of joint action. *Exp. Brain Res.* **211**, 439–445. (doi:10.1007/s00221-011-2651-x)
64. Sacheli LM, Candidi M, Pavone EF, Tidoni E, Aglioti SM. 2012 And yet they act together: interpersonal perception modulates visuo-motor interference and mutual adjustments during a joint-grasping task. *PLoS ONE* **7**, e50223. (doi:10.1371/journal.pone.0050223)
65. Valdesolo P, Ouyang J, DeSteno D. 2010 The rhythm of joint-action: Synchrony promotes cooperative ability. *J. Exp. Soc. Psychol.* **46**, 693–695. (doi:10.1016/j.jesp.2010.03.004)
66. Yoshida W, Dolan RJ, Friston KJ. 2008 Game theory of mind. *PLoS Comput. Biol.* **4**, e1000254. (doi:10.1371/journal.pcbi.1000254)