Computations underlying people's risk-preferences in social interactions

Short Title: Computations underlying social risk preferences

Erdem Pulcu[1*], Masahiko Haruno[1]

[1] Center for Information and Neural Networks / National Institute of Information and Communications Technology; Suita, Osaka, 565-0871, JAPAN

***To whom the correspondence should be addressed: pulerd@gmail.com**

Abstract

Interacting with other individuals to negotiate how we want distribution of resources (e.g. money, time) to be made is an important part of our social life. Considering that not all of our requests from others are always granted, the outcomes of such social interactions are, by their nature, probabilistic and therefore, risky. While risk-perception has been well studied in non-social contexts, its computational mechanisms in social interactions remains unknown. Here, we investigated value computations underlying how people make unfair, fair or hyper-fair Ultimatum offers to others who accepted or rejected these offers probabilistically in relation to how they valued them. We showed that people adjust their risk-preferences dynamically in social interactions, and these can be predicted from a weighted linear combination of one's Social Value Orientation (SVO), inference about the opponent's SVO–including one's uncertainty in it; and relative prosociality (Δ SVO) interacting with one's risk attitude in the value-based domain. In tandem, our results provide the first evidence to suggest that dynamic risk taking is also a cardinal element of social interactions.

Significance

Before requesting something important from another person, we find ourselves thinking about how they would perceive our request. If we judge that our approach will not be perceived well, we consider tweaking it to make it more acceptable. This study describes the computational mechanisms underlying how people assess the risk inherent in such uncertain social situations where their requests (i.e. offers) may be accepted or rejected depending on how the others value them. Surprisingly, despite its cardinal importance, risk-perception in social interactions has not previously been studied using a computational framework. In an ecologically valid experimental design, we show that people's risk-preferences in social interactions may be adaptive to changing characteristics of their opponents.

Introduction

Using the optimal strategy in non-social and social contexts is a key challenge of our everyday life, and from a broader perspective it is closely linked to our survival/fitness. Making requests to others which might result in fair, unfair or hyper-fair resource distributions is an important aspect of human social interactions(1-4). A common psychological observation would suggest that the optimal strategy in cooperative/competitive social interactions (e.g. in the Ultimatum Game(1, 3), or in the Prisoner's Dilemma(5)) should involve considering how others will perceive our requests by simulating their valuation processes before initiating these kinds of communications. However, as our knowledge of other people's state of mind is limited, we will commonly misjudge their response and our requests will not always be granted. In other words, the outcomes of such social interactions are inherently risky. This suggests that risk perception should be a cardinal aspect of these social interactions, in the same way it is for economic decision-making under uncertainty(6). However, the role of risk perception in social decision-making has not previously been studied and, in particular, it is not known whether similar approaches to risk are used in non-social and social domains.

Previous theoretical work has laid the mathematical foundations of how risk perception may emerge naturally in social interactions(7). In their work, Press and Dyson demonstrated the mathematical possibility of sustaining mutual cooperation even when resources are distributed unfairly. Using the example of a simple social game they demonstrated that, in order to sustain mutual cooperation under unfair conditions, the player (i.e. he) who is enforcing unfair terms still needs to give enough incentive to the opponent (i.e. she). This means that he needs to have a good understanding of the opponent's underlying value function to predict at which stage she might change her strategy. In this context, human social interactions have an intrinsic element of risk; whereby one side setting the conditions of the mutual cooperation needs to consider the risk associated with the other's defection possibility (as in the Ultimatum Game rejection behaviour). Additionally, with any move that he makes towards maximising his own payoff by offering conditions that are not in harmony with the underlying value function of the opponent, he will get into a riskier domain where the opponent's defection probability would increase. Taken together, these conditions reveal a negotiation scenario in which a player who is interested in maximizing his payoff needs to make value-based decisions in a social context while incorporating the risk associated with his opponent's defection/rejection likelihood (e.g. the Balloon Analogue Task(8) in value-based; and the proposer behaviour in the Ultimatum Game(3) in the social competition domains).

In this study, we wanted to construct an experimental setting to probe these decision-making processes and to reveal the context dependent aspects of risk-preferences between non-social and social settings. We also wanted to investigate the extent to which people's risk preferences in social interactions are influenced by socially congruent variables related to one's degree of prosociality and inferences about the prosociality of their opponents. Here, by focusing on people's Ultimatum giving behaviour, we investigated the computations underlying the ways in which risk is modulated in social interactions where our participants were asked to: (i) learn the underlying value functions of two distinct computerised agents with different Social Value Orientations (SVOs[1,2]; one prosocial, the other one individualistic; categorically defined with respect to their degree of prosociality) in a social learning session by observing their Ultimatum acceptance preferences (see Fig. 1 and legends for the experimental design), and (ii) transfer these value functions to make value-based decisions while giving Ultimatum offers to those agents in social interactions, while negotiating to maximise their monetary outcomes. We tested two competing hypotheses based on the assumption that the optimal strategy in social decision-making should follow similar principles to that of value-based risk decision-making. In the first family of models, we tested the prediction that participants should integrate the inferred acceptance probability of their opponents (which is transformed by a non-linear probability weighting function) with their potential self-reward magnitudes and make their decisions according to the expected value difference between two options. In alternative models, we also considered the possibility that people's risk preferences in social interactions may be better captured by [exponential or mean-variance] utility functions which are commonly reported in the literature(9), where modulated self-reward magnitudes are integrated with opponent's inferred acceptance probability (see Materials and Methods for mathematical definitions). Finally, in order to test the prediction that risk perception may be modulated differently in non-social and social contexts, we used a probabilistic value-based risk decision-making paradigm as a control condition (Fig. 1).

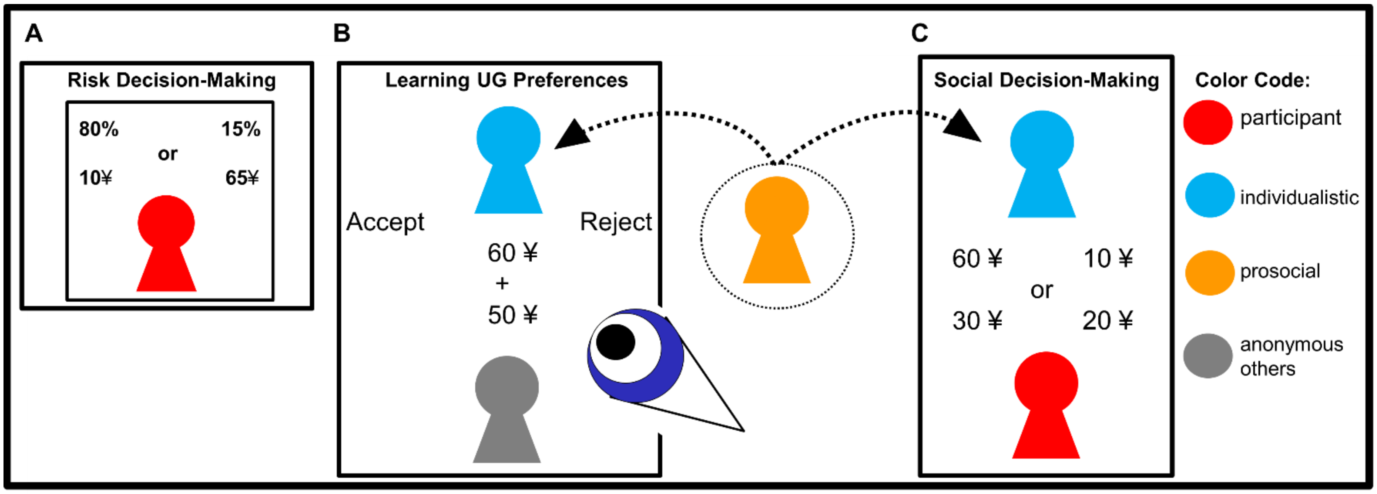

Fig. 1. Details of experimental paradigms. In total participants completed 700 trials: 100 trials in task A, 360 trials in task B and 240 trials in task C. (A) Outline of the control task, where participants' risk perception in a non-social context was evaluated. (B) Participants completed two observational social-learning sessions (represented by the schematic blue eye observing the Ultimatum Game interaction), where they were asked to predict the Ultimatum acceptance preferences of two social agents with different Social Value Orientations (SVO) who were responding to offers coming from different anonymous individuals, blue: individualistic; orange: prosocial agent (colour coding is consistent in all of the subsequent figures). (C) Following the observational social-learning sessions, participants completed a social decision-making experiment in which they were asked to give Ultimatum offers to those social agents from a binary selection. All tasks were self-paced.

Our *a priori* prediction was that at the population level, people's risk attitudes would not be significantly different in each of the studied domains, but we anticipated that there may be within-subjects differences in the social decision-making domain, by which people display a risk attitude that is adaptive to changing characteristics of their opponents.

Results

Social Learning Session. Participants were able to predict the Ultimatum acceptance preferences of social agents with different SVOs ~ 70% correctly. Prediction accuracy was significantly higher than random guessing for both of the social agents (Fig. 2A; t-tests from 0.5; all $t>22$; $p<0.001$, Bonferroni corrected). Participants were able to predict the decisions of the prosocial agent significantly more frequently than the individualistic agent ($t=-9.94$; $p<0.001$), which may be due to the fact that we had higher percentage of prosocial individuals in this cohort (n=29 vs n=5 individualistic participants according to the categorical classification of the SVO Triple Dominance Measure(10) which penalises inconsistent responses). Participants' prediction accuracy closely followed the subjective valuation of the social agents: predictive accuracy was low around the indifference point of agents' subjective valuations ($\tilde{v} = 0$) and it increased whenever the subjective valuation of agents was either very negative or positive (Fig. 2B).

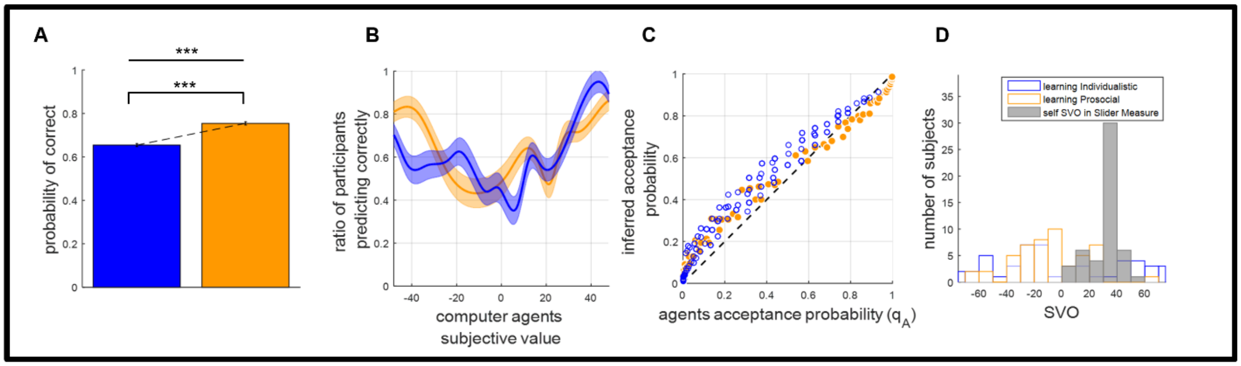

Fig. 2. Summary of results from the observational social-learning session. (A) Participants' prediction accuracy was significantly higher than random guessing and higher for the prosocial agent relative to the individualistic agent (* P<0.001, error bars show SEM). (B) Prediction accuracy followed the subjective valuation of the social agents, irrespective of their SVOs. (C) Encoded value functions were highly significantly correlated with social agents' actual value functions (r>.97, P<0.001), and these had non-linear properties. (D) Participants calculated SVOs based on their decisions in the learning sessions were distributed significantly differently than their actual SVOs in real-life (*D*>0.60, P<0.001), suggesting they actively made decision which are not in accordance with their SVOs to learn about these social agents.**

The decisions of the computerised social agents were generated by a model which was derived from the SVO framework(1) (also see Materials and Methods). We fitted the same model to participants' predictions in the learning session to estimate their inference of their opponents' value function ($\tilde{q}_A$). The encoded value functions estimated from the learning session ($\tilde{q}_A$) correlated highly significantly with agents' actual value functions (q_A; all r> .97, all p<0.001, Bonferroni corrected, see Fig. 2C); and the learning model had highly significantly better fitting relative to a model which makes random predictions in terms of –log likelihood values (t-test from 0.69; all t<-81, all p<0.001, Bonferroni corrected) for both of the social agents, and had pseudo-R^2 values ($\bar{R}^2$; adjusted for the sample size and the number of free parameters(11, 12)) 0.233 and 0.378 for the individualistic and the prosocial agents, respectively. McFadden (1974) suggests that $\bar{R}^2$ values between 0.20 and 0.40 indicate highly desirable model fitting. Plots of the encoded value functions against the actual value functions revealed that these had non-linear properties (Fig. 2C). This model also had better fitting relative to two alternative reinforcement learning models which were fitted to the partcipants' behaviour in the learning sessions (see Materials and Methods, all $F_{2,147}$>168, all p<0.001, Bonferroni corrected)

The comparison between participants' SVO in degrees as measured by the SVO Slider Measure(2) and based on their decisions in the learning sessions suggested that participants should actively be making decisions to predict the choices of social agents which were not in accordance with their own SVO (i.e. not choosing for oneself; 2-sample Kolmogorov-Smirnov tests, all D>0.62; all p<0.001, Bonferroni corrected; see Fig. 2D). Similarly, the distribution of participants' calculated SVO based on their predictions in the learning sessions were also significantly different between the individualistic and prosocial agents (D= 0.28; p=0.03), suggesting that participants relied on different selections to learn about their opponents' underlying value functions.

We were also interested in understanding participants' affective reaction to these social agents. In order to address this issue, after each learning session, we asked our participants to rate the imagined personalities of these social agents on a number of different domains. These questions were related to social constructs such as the SVOs of these social agents and how much our participants would like them in real life. These items showed that the prosocial agent was rated consistently higher relative to the individualistic agent, which conforms to the general intuition that prosocial individuals would be regarded more positively in real life (2x4 Multivariate ANOVA showing main effect of agent $F_{3,294}$=3.01, p=0.03 and main effect of the interaction term $F_{3,294}$=10.163, p<0.001; see Fig. S1A). Particularly the participants' responses to Q3 in which we asked how many people they know in real-life who behave similarly to the computerised agents whose decisions they observed, suggests that our experimental manipulation successfully mimicked interactions with real human opponents (1-sample t-test relative to 0 (i.e. computerised agents' decisions do not resemble any people that I know); all t>13; all p<.001, Bonferroni corrected).

Social Decision-Making. Upon completion of the learning sessions, our participants progressed with the social decision making experiment(s), where they interacted with these social agents by giving them Ultimatum offers for 120 trials each (see Fig. 1C for task screen).

After each social decision-making block, participants were asked to rate how much weight they put on inferred acceptance probabilities and/or their self-reward magnitudes on a scale from 0 to 10. Here, a rating of 0 would refer to making decisions only based on inferred acceptance probabilities; a rating of 10 would refer to relying solely on self-reward magnitudes and 5 would mean their equally weighted integration. In accordance with our predicted value-based social decision-making model (i.e. making offers based on their expected value difference), our participants reported that they considered both the other's inferred acceptance probability ($\tilde{q}_A$) and how much they would win if their offer is accepted (i.e. self-reported integration weights were significantly different than relying on either self-reward magnitudes or encoded acceptance probabilities; all p<0.001, Bonferroni corrected, see Fig. 3A). However, their self-reported integration of these decision variables while making their offers were also significantly suboptimal (i.e. participants reported putting more weight on inferred acceptance probability ($\tilde{q}_A$)) relative to an integration with equal weighting (all t>10.46; all p<0.001, Bonferroni corrected), whereas there were no significant differences between the integration weights reported for the prosocial and the individualistic agents suggesting that our proposed value-based social decision-making model should apply irrespective of opponents' SVO (also in Fig. 3A).

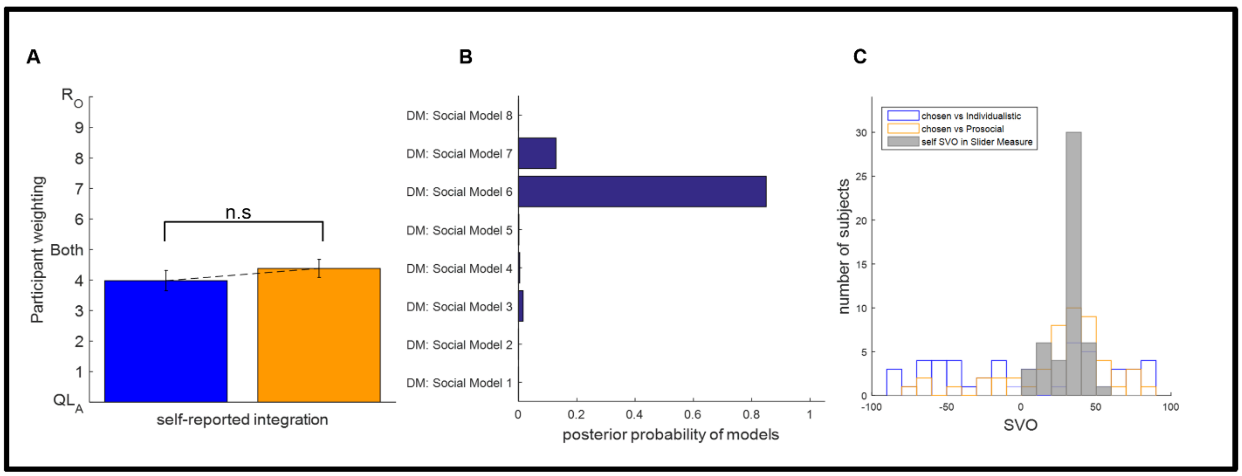

Fig. 3. Participants' self reported value integration and model selection in the social decision-making experiment. (A) In line with our proposed model, participants reported that they considered both their self-reward magnitudes and others' inferred acceptance probabilities. Mean of bars are highly significantly different than both 0 (only consider inferred acceptance probability, i.e. QL_A) and 10 (only consider self-reward magnitude, i.e. R_O; all P<0.001). Self-reported integration weights were comparable against the prosocial and the individualistic agents (n.s: not significant, error bars show SEM). (B) Model selection based on Bayesian posterior probability weights recommends Model 6 (longer bars indicate better fitting), which is based on a value-based decision-making model making use of the exponential utility function. (C) Participants' SVOs calculated on the basis of their decisions in the Ultimatum giving experiments were distributed significantly differently than the distribution of their actual SVOs in real-life (all *D*>0.26; all P≤0.056). These vectors of SVOs calculated based on chosen options were not correlated with participants' actual SVO as assessed by the Slider Measure (all P>0.32).

We selected between competing social decision-making models based on their Bayesian posterior probability weights given the data(13), by the following formula:

$$weight_i = \frac{\exp(-(BIC_i - BIC_{\min})/2)}{\sum_k \exp(-(BIC_k - BIC_{\min})/2)}$$

while considering each model's Bayesian Information Criterion (BIC) value, which penalises for additional free parameters. Model fits revealed that the social decision-making model which is based on integration of inferred acceptance probability ($\tilde{q}_A$) with self-reward magnitudes modulated by an exponential utility parameter capturing participants' risk preferences was the best fitting model (Fig. 3B, a selection based on group-wise sum of BIC values also favours this model, see Fig. S2). In Social Decision Making Models 4 and 5, we used participants' self-reported weights and added a free parameter to allow unequally weighted integration of nonlinearly perceived probabilities and magnitudes respectively, but these models did not improve the fitting of the probability weighting function family of models any further. Bayesian posterior probability weights suggested that models which mainly relied on either reward magnitudes or inferred acceptance probabilities could not account for our participants' choice behaviour.

One might argue that our approach to modelling the social decision-making task neglects any learning which might be happening in parallel, during this stage. If learning continues to take place during the social decision-making period, the predictive accuracy of our model which considers the inferred acceptance probability ($\tilde{q}_A$) solely based on the learning session should gradually decay down the trials. In order to investigate this possibility, we segmented the trials in the social decision-making task into three temporal sections: early (1-40); middle (41-80); late (81-120) trials and compared the predictive accuracy of our model across these sections. This control analysis showed that model predictions were highly stable (all $F_{2,147}$ <0.27; all p > 0.77), and suggested that any additional learning which might take place during the decision-making period would not have a profound effect on inferred acceptance probabilities participants use to compute the expected value of offers.

Further analysis done by comparing participants' SVOs calculated based on their offers in the Ultimatum giving experiments and their SVOs measured by the Slider Measure suggested that these came from significantly different distributions (Kolmogorov-Smirnov tests; all *D*>0.26, p≤ 0.056, uncorrected; see Fig. 3C) and these vectors of SVOs calculated from the decision-making experiments were not correlated with participants own SVO as obtained from the Slider Measure (-.141< r <-.035, 0.32< p <0.81). In tandem, these results limit the possibility that in the social decision-making experiments participants performed using other cognitive models linked to the SVO framework that we did not consider here (including those which were based on reinforcement learning); and lends further support for our hypothesis that participants would use a cognitive model from the risk/value-based decision-making framework.

Risk modulation in non-social and social settings. In order to evaluate whether risk perception is modulated differently between non-social and social contexts, we conducted a simple value-based risk decision-making experiment as a control condition before the learning sessions and established our participants' risk taking preference at the baseline (see Fig. 1A for the control task screen).

First, we conducted a model-free analysis on participant's choices by analysing the proportion of risky decisions made in each domain, focusing on the trials where participants were asked to choose between a low probability-high magnitude and a low magnitude-high probability option. This analysis revealed that the frequency of risky choices, after controlling for the number of trials meeting the criteria described above, were not significantly different across each domain (Fig. 4A; $F_{2,147}$=3.263, p=0.073). The best fitting model to participants' choice behaviour in the value-based risk decision-making experiment, which also utilised the same exponential utility function as in the social decision-making experiment, had –log likelihood value of 0.242/trial and group-wise sum of BIC score 2699.

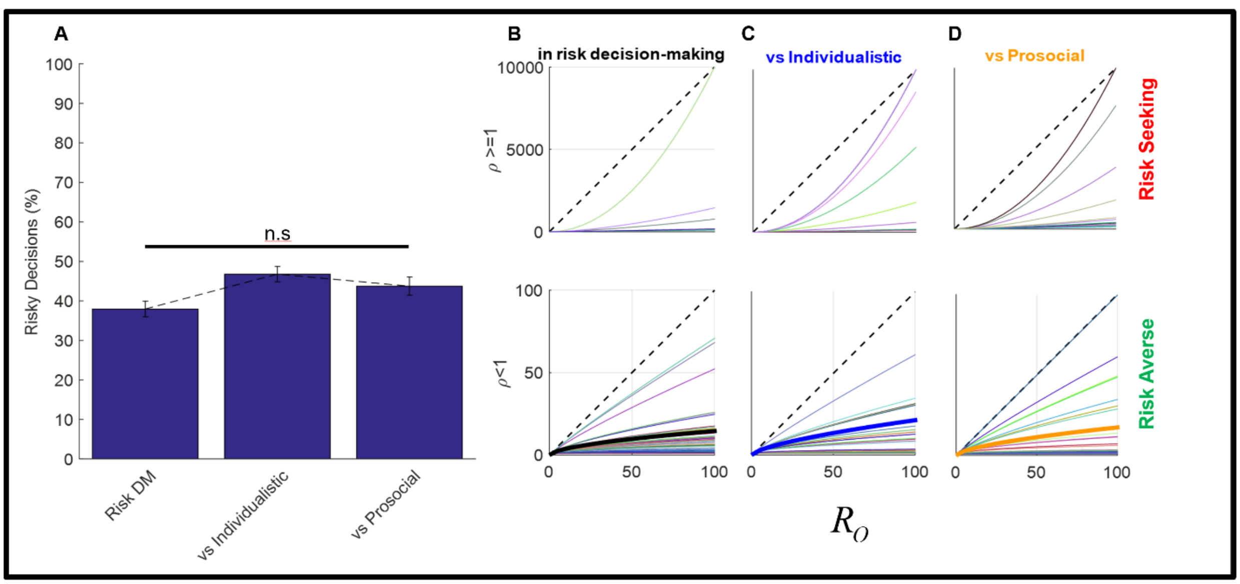

Fig. 4. Risk decision-making across non-social and social domains. (A) Proportion of risky decisions made in each domain was not significantly different (P=0.07). Parameter estimates modulating the participants gains (R_O) in each domain: (B) value-based risk decision-making; (C) Ultimatum giving against the Individualistic opponent; (D) Ultimatum giving against the Prosocial opponent. In B-D, upper panels show the exponential utility curves for risk-seeking preferences, whereas the lower panels show the risk-averse preferences. In B-D, thin lines with the same [R,G,B] colour coding specify each subject's exponential utility curve and thick lines in black, blue and orange colours show the population means for each domain, where we observed a pronounced risk aversion across all domains (P<.002, Bonferroni corrected).

By estimating the risk parameters separately for the value-based risk decision-making experiment (i.e. the non-social context) and the two Ultimatum giving experiments against different social agents (i.e. social context), we were able to show that on average our participants displayed a pronounced risk aversion in all domains (all p<0.002, Bonferroni corrected relative to risk neutrality $\rho = 1$, see Fig. 4B-D for utility curves). These risk parameters estimated separately in each domain correlated highly significantly with the proportion of risky decisions the participants made in each domain, where parameter values higher than 1 indicate a risk-seeking preference (all r > .71, all p<.001, Bonferroni corrected for pairwise comparisons).

The parameter estimates from the value-based experiment were correlated with the estimates from the social decision-making experiments (all r>.40, all p<.004, Bonferroni corrected). However, risk parameters were not correlated within the social decision-making domain (r=.22, p=.122), indicating that people's risk-preferences in social interactions may be adaptive to the changing characteristics of their opponents. In the social decision-making domain, although our participants displayed a slightly more risk seeking tendency against the individualistic agent relative to the prosocial agent, this was not significantly higher (p=0.74, see Fig. S3), supporting our *a priori* prediction at the population level that people should not be acting in a consistently risk-seeking manner against one type of social agent.

Next, we wanted to explain the "social" risk parameters estimated from our Ultimatum giving experiments by a number of predictive parameters to evaluate the extent to which people's risk-preferences are influenced by social variables describing degrees of prosociality. Our hypothesis was that people's risk-preferences in social interactions should depend on their SVO; their inference about the SVO of their opponents ($S\tilde{V}O$; also including one's uncertainty about this estimate); and the interaction between their baseline risk perception in non-social settings and how prosocial they think their opponent is relative to themselves. Here, the interaction term allows us to model the extent of risk-seeking tendencies in social competitions; particularly in situations where people judge their opponent is more prosocial than themselves, also considering their baseline attitude to uncertainty. To investigate the accuracy of our hypothesis we constructed two multiple linear regression models with these predictive variables:

$$\rho^S = \beta_0 + \beta_1 * SVO + \beta_2 * S\tilde{V}O_\mu^S + \beta_3 * S\tilde{V}O_\sigma^S + \beta_4 * (\rho_R (S\tilde{V}O_\mu^S - SVO)) + \varepsilon$$

where ρ^S is the risk parameter estimated separately from the Ultimatum giving experiments and $S \in \{i, p\}$, defines whether the opponent is individualistic or prosocial. In this linear regression model ρ_R is the risk parameter estimated from the value-based risk decision-making experiment and it is orthogonalised with respect to each social risk parameter. Here, it is important to point out that participants' inferences about their opponent's SVO ($S\tilde{V}O^S$) depend on their encoded value function ($\tilde{q}_A$) from the learning sessions, which has a stochastic nature (i.e. a sigmoid function). In order to get the best estimate of the opponent's inferred SVO in degrees, we ran the learning model with each participant's estimated parameters through the learning stimuli 1000 times per participant per social agent and calculated the resulting SVOs in degrees (see Fig. S4A). We used both the mean (μ) and the standard deviation (σ) of the distribution of calculated SVOs from these 1000 simulations as the best estimate of opponent's inferred SVO and the participant's uncertainty about this estimate, and included these scalar values for each participant in our regression model (see Fig. S4B). For robustness, we performed a leave-one-out cross-validation procedure and repeated the multiple linear regression analysis. Further analysis which was done on the regression coefficients by performing 1-sample t-tests from 0 suggested that the regression coefficients for all of these predictive variables were highly significant ($33<|t|< 301$ for the individualistic, $4 <|t|< 302$ for the prosocial, all $p<0.001$, all Bonferroni corrected, set level for the p-value is 0.006 for 8 comparisons, see Fig.5). Subsequent analysis conducted on the actual versus predicted risk parameters suggested that these were highly significantly correlated for both of the agents (all $r>.76$, all $p<.001$, Bonferroni corrected, see Fig. S5).

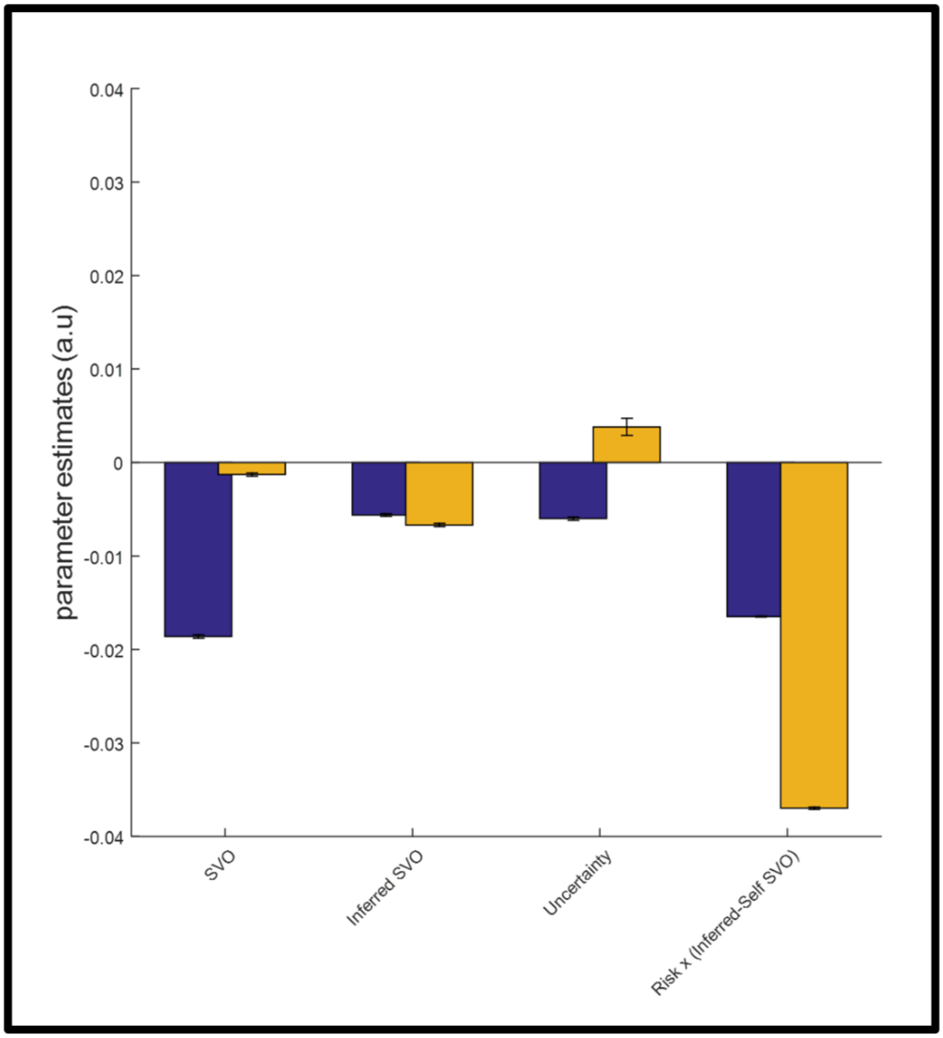

Fig. 5. Parameter estimates $\{\beta_1 ... \beta_4\}$ from two leave-one-out cross-validated multiple linear regression models accounting for the social risk parameters (y-axis). All regression coefficients were highly significantly different than 0 (all P<0.001, Bonferroni corrected; error bars show SEM). Variables in the x-axis refer to those described in the main text, SVO: participants' Social Value Orientation; Inferred SVO: $S\tilde{V}O_\mu$; Uncertainty: $S\tilde{V}O_\sigma$; Risk x (Inferred-Self SVO):

$\rho_R * (S\tilde{V}O_\mu - SVO)$.

Response Times. Response times from the risk decision-making experiment were comparable to those from the social learning and decision-making experiments (all p>0.15, see Fig. 6). Additionally, there were no statistically significant differences between the prosocial and the individualistic agents in terms of participants' response times in either the social learning or the social decision-making experiments (all p>0.39), suggesting that choice difficulty was well controlled throughout these sessions. However, response times in the social decision-making experiments were significantly longer than those in the social-learning experiments, supporting our hypothesis regarding the suitability of the value-based risk model applied to the social decision-making context, where we predicted that our participants should simulate the underlying value functions of their opponents and integrate it with their self-reward magnitudes while they were making Ultimatum offers (all $t>2.28$, all $p<0.025$; also in Fig. 6).

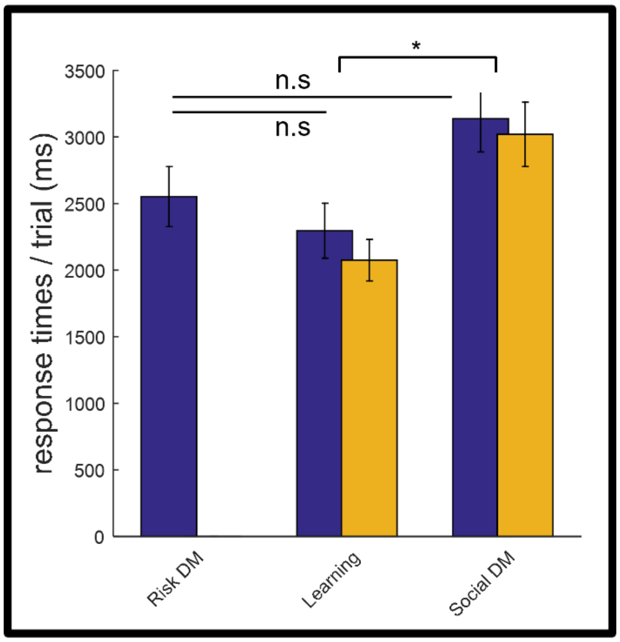

Fig. 6. Response times across all experimental sessions. Response times from the value-based risk decision-making experiment were comparable to those from the social-learning and social decision-making experiments (n.s: not significant). Response times were longer in the social decision-making experiments relative to the social-learning experiments (* P<0.05). Across social-learning and social decision-making experiments, there were no significant differences between prosocial/individualistic conditions.

Discussion:

In line with our prediction, our results suggest that in social interactions which may result in resource distributions between two individuals, people employ a cognitive model which shares properties with well-established computational models concerning value-based decision-making under uncertainty(14-16). However, unlike risk decision-making paradigms in which probability and reward information is usually given explicitly(17-20), in social interactions people need to infer others' valuation processes (Fig. 2C) in order to make value-based decisions(21). Here, we showed evidence to suggest that people's risk-preferences in social interactions may be adaptive to the changing characteristics of their opponents, considering that risk parameters estimated in the same parameter space ($0 < \rho \leq 2$) but separately from each experiment were not correlated. Particularly, it is important to point out that with careful sequencing using an approach similar to a gradient decent, we were able to minimise the correlation between decision variables(22) over 2x10^6< iterations, and we were able to make the face values of the reward distributions in the Ultimatum giving experiments the same for both of the opponents (i.e. prosocial/individualistic social agents, where the order of presentation was randomised for each subject and for each type of opponent), such that our participants needed to utilise their opponent's encoded value function to compute the expected value difference between two options only cued by the colour of the icon representing their opponents (Fig. 1C). This approach minimises the possibility that within-subject differences in risk parameters (Fig. 4) could have been observed due differences in the numerical components of the stimuli.

In mainstream economics and finance, people's risk preference is often regarded as a hard-wired trait. However, a number of recent studies have suggested that value-based decision parameters may be subjected to influence after observing others' decisions which were performed in the same context(9, 23, 24), challenging this point of view. Here, by focusing on social interactions, we provide evidence to suggest that even in the absence of observations of comparable decisions as used by these previous studies(9, 23, 24), people's risk-preferences may be subject to change in a social context, depending on the nature of the social interaction they are engaged in. While the impact of social framing on risk-preferences has previously been investigated in the context of the trust game where outcome probabilities were explicitly stated(25), to the best of our knowledge the present study is the first to describe the computations underlying risk-preferences in human social interactions where outcome probabilities were directly related to participants' inferences about the value functions of their opponent.

It is important emphasise that our experimental design did not have any element of observing others' risk-preferences. Instead, as we have shown, we anticipated that social context specific risk-preferences should emerge naturally due to the fact that the outcomes of these social interactions were probabilistic. We provide further evidence for "social" risk perception by showing that in social interactions the change in risk parameters may be adaptive, considering that in our multiple linear regression analyses (Fig. 5) we showed an overarching and consistent pattern of significant contributions from key social variables which predicts people's risk-preferences in the social context, irrespective of the SVO of their opponents. Although designing two different computerised social agents increased our experimental difficulty in terms of number of trials our participants needed to complete (t=700), it also allowed us to reveal these overarching contributions which we think are cardinal for cognitive/computational models of human social interactions.

Here, it is worthwhile to comment on why we decided against the inclusion of a “competitive” agent (based on the definition of Murphy et. al) in our experimental design. Previous studies with relatively large sample sizes investigating Social Value Orientation in the population showed that the population density of competitive individuals are only around 9% (26). Furthermore, people with competitive SVO are driven by achieving superiority over the others, which limits them to only accept the offers that satisfy this superiority criterion, making their underlying value function not very suitable for probing risk perception in social interactions. Additionally, the inclusion of a “competitive” agent would require our participants to complete at least an additional 300 trials [across social learning and decision-making sessions], making it rather difficult to achieve. Considering that in our cohort participants’ SVO and their inferences about the SVO of their opponents showed a healthy degree of variability, and focussing on each pairwise combination (n=100, see Fig. S6), we think our proposed model suitably meets the generalisability criteria to account for risk perception in many different social interactions which naturally occur in real-life.

One interpretation of the multiple linear regression analyses is that, keeping the contributions of all other social variables fixed, people’s positive inferences about the SVOs of their opponents will decrease the value of their risk parameters; such that they would be likely to display risk-aversion. Additionally, our multiple linear regression model also predicts that, keeping the contribution of all other social variables fixed, people with higher SVO (i.e. prosocial individuals) will have a significantly lower risk parameter against an individualistic opponent relative to the value of their risk parameter against a fellow prosocial opponent. Under these conditions, if the value of the social risk parameter converge to 1 (i.e. risk neutrality), this would mean that a prosocial individual would make more optimal value integration against an individualistic opponent, while still maintaining a risk-seeking tendency against a prosocial opponent. On the other hand, highly significantly negative regression coefficients for the interaction term indicates that people would display risk-aversion in situations where they judge their opponent as more prosocial relative to themselves. While keeping the contributions of all other social variables fixed, when the value of the interaction term is increasing, significantly lower regression coefficients against a prosocial opponent will lead to smaller risk parameter values, indicating a risk-aversion, whereas the model predicts that people would show a tendency for risk-seeking attitude towards others whom they judge to be less prosocial than themselves. Finally, our analysis suggests that one's uncertainty about the opponent's SVO will influence risk perception in social interactions differently depending on the actual degree of prosociality of the opponent. We show that a unit of uncertainty will increase the value of the risk parameter against a prosocial opponent, whereas decreasing it against an individualistic opponent. Thus, our results reveal that people's risk-preferences in social interactions are under the influence of a number predictive variable defining different aspect of the nature of a social encounter in terms of degrees of prosociality.

Our study also has implications for understanding the interactions in the Ultimatum Game, where the wide majority of the previous literature focused on responders' behaviour(3, 27-32). In our study, the participants were explicitly instructed to treat the binary options they were presented with like thoughts in their mind, such that they knew their opponents can only see the chosen offer, and can never know whether the unchosen option was better or worse. The structure of our experiment which allowed our participants to make offers from a binary selection complements previous Ultimatum Game studies where responders were commonly asked to make decisions about a single offer per trial. Under these conditions our proposed social decision-making model including an exponential utility term, was the best fitting model to subjects' choice behaviour (Fig. 3B), and our participants' self-reports also suggested that both the self-reward magnitudes and the others' inferred acceptance probabilities needed to be considered while making offers (Fig. 3A). To the best of our knowledge, the current study is also the first to address value computations underlying Ultimatum giving, and it provides evidence to suggest that proposers' do not solely rely on responders' acceptance probabilities, but make offers based on their expected value. Based on these findings, we recommend that future cluster(33) or hyper-scanning(34) studies of Ultimatum bargaining in neuroeconomics, should consider computational models which explicitly parametrize participants' risk-preferences.

Finally, although the social-learning session was not our main focus in this work, we showed behavioural computational evidence to suggest that our participants could suitably transfer the encoded value functions of others from one [learning] environment (i.e. observing social agents' responses to singular Ultimatum offers) in order to utilise them for solving decision problems in another [decision-making] environment (i.e. making Ultimatum offers from binary options). Thus, our behavioural study uncovers an important aspect of human social interactions and highlights the need for further research in three main streams involving functional magnetic resonance imaging (fMRI): (i) what are the regions involved with neural computations underlying how people transfer the encoded value functions of their opponents in social interactions; (ii) what are neural mechanisms responsible for tracking the value functions of opponents with different SVOs; and (iii) which brain regions encode the estimated trial-by-trial variability in the social risk parameters.

Materials and Methods

Participants. In total 50 healthy individuals (54% males) who reported no history or current neurological or psychiatric disorders, or use of any psychotropic medication were recruited from the general population. The average age of this cohort was 31.5 (range: 20-56 years; STD=±9.49). On average the participants had 16 years of education (STD= 2.1 years) and reported annual income of 1.91 million Japanese Yen (STD=1.71 million Yen).

Procedure. The study took place in Center for Information and Neural Networks (CiNet) and it was approved by the CiNet Research Ethics Committee. Participants who met the inclusion criteria were given an appointment for the behavioural experiments. The testing session began by an explanation of the research procedures, followed by obtaining an informed consent. First, the participants completed a battery of questionnaires related to their demographic information and Social Value Orientation (SVO) in pen and paper format. After that they completed the value-based risk decision-making task which lasted for 100 trials (Fig. 1A). The amount of money participants won in the risk decision-making task was added to their performance related reimbursement. Then, the participants completed the learning sessions for both the individualistic and prosocial agents in which they were asked to predict whether the social agents will accept or reject the offers coming from other anonymous individuals. The participants were told that the social agents whose Ultimatum responses they needed to predict were two participants from a previous study conducted by our research group. In fact, they were computerised agents making decisions following an underlying value function (see below). Similar methodology is frequently used in behavioural studies conducted by other research groups (9, 35). The learning sessions contained 180 trials each and the order of these learning sessions was counterbalanced across participants. The participants won ¥25 for every correct prediction which was added to their performance based reimbursement amount. After completing the learning sessions, participants were asked to respond to various descriptive questions while considering the imagined personalities of these social agents (see Fig. S1). Finally, the participants completed the social decision-making experiments, for 120 trials against each social agent, in the same order they completed the learning sessions. In the Ultimatum giving experiment, participants obtained the monetary amount (R_O) in all of their accepted offers, which was added to their performance based reimbursement. All of the behavioural experiments were self-paced. All experiments took place in a comfortable room designated for testing purposes and all tasks were presented by PsychToolbox 3.0 running on MATLAB (MathWorks, Inc.).

Description of the computerised social agents. Two distinct computerised social agents, whose behaviours were guided by the way they computed the subjective value of the conditions they faced (which are in fact unfair, fair, and hyper-fair Ultimatum offers) were defined by a model from the social value orientation (SVO) framework, based on a model derived from a previous publication of our research group (1). Here the subjective value ($\tilde{v}$) of a condition is calculated as:

$$\tilde{v} = \alpha R_S - \delta R_O + \rho \left| R_S - R_O \right| \tag{1}$$

where R_S (always from the perspective of the social agents) and R_O depicts self and other's reward magnitude, respectively. The agents make decisions following a stochastic choice model where q_A is the probability of accepting a condition (36):

$$q_A = 1 / (1 + \exp^{(-\beta \tilde{v})}) \tag{2}$$

Here, β is the inverse temperature term, which gives the shape of the sigmoid function, based on our *a priori* assumption about the shape of the sigmoid function in humans, which should particularly be the case if the number of trials in the learning session approach to infinity; that is if the participants had the opportunity of observing the behaviour of the computerised social agents for a very long time.

The hyper-parameters defining the valuation of the agents, $\alpha, \delta, \rho, \beta$ were set to [1.096, 0.382,-2.512, 0.037] for the individualistic; and [1.368,-0.644,-3.798, 0.045] for the prosocial agent. The key difference between these two agents were that the prosocial agent valued conditions cooperatively, whereas the individualistic agent valued them competitively; and that the prosocial agent was more sensitive to the absolute value difference between the self and other's reward magnitude. We generated a vector of responses to 180 trials in the learning sessions by the defined model where the SVO of the social agents were calculated by the following formula derived from Murphy et al. 2011:

$$SVO^\circ = \arctan\left(\frac{(\sum R_O)/n_A - 50}{(\sum R_S)/n_A - 50}\right) \tag{3}$$

where n_A is the total number of accepted conditions. After extensive simulations to evaluate how agents would behave, a selection was made such that the SVO of the prosocial agent was 31.47° and the SVO of the individualistic agent was 12.36°; clearly falling into the categorical boundaries described by Murphy et al., 2011. Therefore, these strategies were labelled as "prosocial" and "individualistic" throughout this manuscript.

Social-learning. There are a few models in the literature to account for how people learn during social interactions(21, 37, 38). Due to the widely-known complexity of this process, we did not focus too much on specific models of social-learning by performing detailed trial-by-trial analyses here, but we assumed that learning occurs through successful simulation of other's valuation model(39) (achieved by model-free, reinforcement or Bayesian learning; or their weighted combination) such that as the number of trials (t) approach to infinity, $\tilde{q}_A \to q_A \mid t \to \infty$, that is subjects inferred acceptance probability ($\tilde{q}_A$) will fully converge to the social agent's true acceptance probability (q_A). We confirmed that social learning occurs suitably well by performing model-free analysis of the data from the learning sessions and also by fitting the proposed valuation model of the social agents to the participants' choice data to generate participants' inferred choice probability for each social agent ($\tilde{q}_A$; see Fig. 2).

Although the learning session was not our primary focus, we still wanted to compare the performance of the valuation model with two alternative Rescorla-Wagner models(40) which were fitted to participants' choice behaviour in the learning sessions. In alternative Model 1, the participant updates his/her estimate of the social agent's overall acceptance probability on trial *i* in proportion to the prediction errors(ε) on trial *i-1* on trial-by-trial basis:

$$\tilde{q}_{A(i+1)} = \tilde{q}_{A(i)} + \eta_{\tilde{q}_A} \varepsilon_{(i)} \tag{3}$$

where $\eta_{\tilde{q}_A}$ is the learning rate. In the alternative Model 2, each of social agent's parameters in the described SVO model (Eq.1 and 2) is updated on trial-by-trial basis, for example:

$$\alpha_{(i+1)} = \alpha_{(i)} + \eta_{\alpha} \varepsilon_{(i)} \tag{4}$$

where η_α is the learning rate updating the value of the free-parameter α from trial *i* to *i*+1 . This second model had free parameters to account for the learning rates, updating each of the parameters of the SVO-based valuation model (Eq.1 and 2).

Model comparison for the learning session favoured the SVO-based valuation model, which had significantly lower –log likelihood values relative to the reinforcement learning models (all $F_{2,147}$>168, all p<0.001, Bonferroni corrected). Furthermore, close relationship observed between q_A and $\tilde{q}_A$ from the

SVO model, as well as model fitting metrics falling into the highly desirable range (11) suggest that the present approach is suitable to obtain participants' inferred choice probabilities.

Description of the control experiment and the risk model. Considering that our social agents were designed to make choices following specified value functions, interacting with them in the main experiment becomes a risky business. Consequently, we decided to select a value-based risk decision-making task (Fig. 1A) as our control experiment. This selection enabled us to understand whether people process risk in non-social and social contexts differently. We fitted computational models to participants' data, as described below and in the following sections.

In line with the previous literature(41), we assumed that people process uncertainty in probabilistic gambles nonlinearly. The probability weighting function, which was based on previous research (39, 42), is defined by the following formula:

$$\hat{p} = 2^{(-(-\log_2(p)^{\gamma_R})} \tag{5}$$

where γ_R>0 is the risk parameter. The $\log_2$ function always crosses the p/p diagonal at 0.5 and in our point of view accurately captures the intuition that people should have a somewhat accurate perception of the 50/50 odds. Participants then compute the expected value of a gamble they face accordingly:

$$\pi = m * \hat{p} \tag{6}$$

where m is the reward magnitude, and make their choices in relation to the subjective value difference between each gamble (i.e. here, the difference between left and right options):

$$\Delta\tilde{\pi} = \pi_L - \pi_R \tag{7}$$

and trial-wise stochastic choice probabilities for each gamble are generated by a sigmoid function:

$$q_L = 1/(1+\exp^{(-\beta(\Delta\tilde{\pi}))}) \tag{8}$$

where $\beta > 0$, is the inverse temperature term adopted from thermodynamics and it determines the degree of stochasticity in participants' choices.

Description of the social decision-making models. In the main experiment, the participants were asked to make Ultimatum offers to these social agents. If the offer was accepted, the participant would receive the amount R_0, whereas if the offer is rejected both sides got nothing for that trial (as in a typical Ultimatum Game experiment).

Participants could complete the social decision-making tasks using a number of different strategies and here we considered 6 cognitive models with variable complexity, where our preferred model is based on simulating the social agents' choice probability for both the chosen and the unchosen options modulated by a social risk parameter and computing the expected value difference by integrating self-reward magnitude information.

According to Model 1, the participant's decision value ($\Delta\tilde{v}$) is depends on the difference between social agent's inferred choice probability ($\tilde{q}_A$) associated with the offers on each side $\{L, R\}$, whereby participant makes decisions only considering other's acceptance probability (i.e. choosing the offer which they think is more likely to get accepted):

$$\Delta\tilde{v} = \tilde{q}^S_{A,L} - \tilde{q}^S_{A,R} \tag{9}$$

where $S \in \{i, p\}$, defining whether the opponent is individualistic or prosocial.

Due to the probabilistic nature of this decision-making, in Model 2, we hypothesised that the inferred acceptance probabilities could also undergo a similar probability weighting transformation as in the value-based risk decision-making task (Eq. 4). Thus, in Model 2 participants' decision value ($\Delta\tilde{v}$) is computed by the following two equations:

$$Q^S_{A,L} = 2^{(-(-\log_2(\tilde{q}^S_{A,L})^{\gamma^S})} \tag{10}$$

$$\Delta\tilde{v} = Q^S_{A,L} - Q^S_{A,R} \tag{11}$$

where $\gamma^S > 0$ is the social-risk parameter estimated separately for each social decision-making experiments (i.e. against different social agents). Here, the introduction of the risk parameter is critical and it allows us to evaluate differences between the modulation of risk perception in social and non-social contexts.

In Model 3, we considered a more sophisticated value computation which also integrates participants' own payoff (R_0). This computation assumes that participants employ a cognitive model with a large degree of overlap with the previously defined risk decision-making model, whereby the participant can choose to make an offer with lower $Q^{S}_{A,R\sim L}$ if the overall expected value is higher.

$$\Delta\tilde{v} = (Q^{S}_{A,L} \cdot R_{O,L}) - (Q^{S}_{A,R} \cdot R_{O,R}) \tag{12}$$

After each social decision-making experiment, we asked participants to rate how much they considered other's acceptance probability, their own payoff, or both on a 0-to-10 scale (e.g. a rating of 5 meaning integration with equal weighting). In Model 4, we used this rating as a linear weighting information, where the weight parameter, w, takes a value in the normalised space (i.e. between 0 to 2) with 0.2 increments as it was directly derived from participants' own report on a 0-to-10 scale (e.g. w=1 meaning integration with equal weighting, w=0.4 meaning more weight is given to the self-reward magnitude, etc.). Here, the decision value is computed as follows:

$$\Delta\tilde{v} = (wQ^{S}_{A,L} \cdot (2-w)R_{O,L}) - (wQ^{S}_{A,R} \cdot (2-w)R_{O,R}) \tag{13}$$

In Model 5, we considered a different version of the Model 4 in which the weight parameter, w_f, is estimated freely in the normalised space (i.e. between 0 and 2) and the decision value is computed accordingly:

$$\Delta\tilde{v} = (w_f Q^{S}_{A,L} \cdot (2-w_f)R_{O,L}) - (w_f Q^{S}_{A,R} \cdot (2-w_f)R_{O,R})$$

(14)

We also fitted two additional models to participants' choice behaviour, which would account for their risk-preferences by revealing the curvature of the utility functions(9). In Model 6, the decision value is computed by the following formula for the exponential utility function:

$$\Delta\tilde{v} = \tilde{q}^{S}_{A,L} R^{\rho}_{O,L} - \tilde{q}^{S}_{A,R} R^{\rho}_{O,R} \tag{15}$$

where ρ is the utility parameter, $\rho > 1$ indicating a risk-seeking and $\rho < 1$ a risk-averse preference.

In Model 7, the decision value is based on the mean-variance utility function, which assumes that participants would make decisions in relation to the utility associated with each option:

$$\Delta\tilde{v} = (\tilde{q}^S_{A,L} R_{O,L} + \lambda R^2_{O,L} \tilde{q}^S_{A,L}(1-\tilde{q}^S_{A,L})) - (\tilde{q}^S_{A,R} R_{O,R} + \lambda R^2_{O,R} \tilde{q}^S_{A,R}(1-\tilde{q}^S_{A,R})) \quad (16)$$

where negative values of the λ account for risk-averse and positive values of the λ account for risk-seeking tendencies. Finally, as a baseline control condition we also investigated the fitness of a model which makes decisions based on self-value difference alone (i.e. Model 8; $\Delta\tilde{v} = R_{O,L} - R_{O,R}$).

At the very last step, choice probabilities under each model $M \in \{1,2,3,4,5,6,7,8\}$ were generated by a sigmoid function:

$$Q_L^M = 1/(1+\exp^{(-\beta_S(\Delta\tilde{v}))}) \quad (17)$$

where $\beta_S > 0$, is the inverse temperature term estimated separately for the social decision-making experiments.

We used a Maximum Likelihood Estimation procedure to evaluate how well the proposed cognitive models explained our participants' choice behaviour. The free parameters were estimated using a non-linear optimization method over a numerical grid which covered whole parameter space, using MATLAB's (MathWorks, Inc.) *fmincon* function with random starts. We chose between competing models (43) in the social decision-making experiments by their Bayesian Information Criterion (BIC) score which penalises more complex models with additional free parameters and Bayesian posterior probability weights as previously described in the main text.

References

1. Haruno M, Frith CD. Activity in the amygdala elicited by unfair divisions predicts social value orientation. Nature neuroscience. 2010;13(2):160-1.
2. Murphy RO, Ackermann KA, Handgraaf M. Measuring social value orientation. Judgment and decision making. 2011;6(8):771-81.
3. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. The neural basis of economic decision-making in the ultimatum game. Science. 2003;300(5626):1755-8.
4. Fehr E, Schmidt KM. A theory of fairness, competition, and cooperation. Quarterly journal of Economics. 1999:817-68.
5. Axelrod RM. The complexity of cooperation: Agent-based models of competition and collaboration: Princeton University Press; 1997.
6. Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. Econometrica: Journal of the Econometric Society. 1979:263-91.
7. Press WH, Dyson FJ. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. Proceedings of the National Academy of Sciences. 2012;109(26):10409-13.
8. Lejuez CW, Read JP, Kahler CW, Richards JB, Ramsey SE, Stuart GL, et al. Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). Journal of Experimental Psychology: Applied. 2002;8(2):75.
9. Suzuki S, Jensen EL, Bossaerts P, O'Doherty JP. Behavioral contagion during learning about another agent's risk-preferences acts on the neural representation of decision-risk. Proceedings of the National Academy of Sciences. 2016:201600092.
10. Van Lange PA, De Bruin E, Otten W, Joireman JA. Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. Journal of personality and social psychology. 1997;73(4):733.
11. McFadden D. Conditional logit analysis of qualitative choice behavior. 1973.
12. McFadden D. Quantitative methods for analyzing travel behavior of individuals: some recent developments: Institute of Transportation Studies, University of California; 1977.
13. Burnham KP, Anderson DR. Multimodel inference understanding AIC and BIC in model selection. Sociological methods & research. 2004;33(2):261-304.
14. Tversky A, Fox CR. Weighing risk and uncertainty. Psychological review. 1995;102(2):269.
15. Hsu M, Krajbich I, Zhao C, Camerer CF. Neural response to reward anticipation under risk is nonlinear in probabilities. The Journal of Neuroscience. 2009;29(7):2231-7.
16. Wu G, Gonzalez R. Common consequence conditions in decision making under risk. Journal of Risk and uncertainty. 1998;16(1):115-39.
17. Schonberg T, Fox CR, Poldrack RA. Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. Trends in Cognitive Sciences. 2011;15(1):11-9.
18. Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF. Neural systems responding to degrees of uncertainty in human decision-making. Science. 2005;310(5754):1680-3.
19. Samanez-Larkin GR, Wagner AD, Knutson B. Expected value information improves financial risk taking across the adult life span. Social Cognitive and Affective Neuroscience. 2010:nsq043.
20. Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, Behrens TE. Mechanisms underlying cortical activity during value-guided choice. Nature neuroscience. 2012;15(3):470-6.
21. Yoshida W, Dolan RJ, Friston KJ. Game theory of mind. PLoS computational biology. 2008;4(12):e1000254.
22. Hayden BY, Heilbronner SR. All that glitters is not reward signal. Nature neuroscience. 2014;17(9):1142-4.

23. Nicolle A, Klein-Flügge MC, Hunt LT, Vlaev I, Dolan RJ, Behrens TE. An agent independent axis for executed and modeled choice in medial prefrontal cortex. Neuron. 2012;75(6):1114-21.
24. Chung D, Christopoulos GI, King-Casas B, Ball SB, Chiu PH. Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. Nature neuroscience. 2015;18(6):912-6.
25. Lauharatanahirun N, Christopoulos GI, King-Casas B. Neural computations underlying social risk sensitivity. Frontiers in human neuroscience. 2012;6:213.
26. Van Lange PA. The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. Journal of personality and social psychology. 1999;77(2):337.
27. de Quervain DJ-F, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, Buck A, et al. The Neural Basis of Altruistic Punishment. Science. 2004;305(5688):1254-8.
28. Pulcu E, Thomas E, Trotter P, McFarquhar M, Juhasz G, Sahakian B, et al. Social-economical decision making in current and remitted major depression. Psychological Medicine. 2014:1-13.
29. Xiang T, Lohrenz T, Montague PR. Computational substrates of norms and their violations during social exchange. The Journal of Neuroscience. 2013;33(3):1099-108.
30. Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E. Diminishing Reciprocal Fairness by Disrupting the Right Prefrontal Cortex. Science. 2006;314(5800):829-32.
31. Koenigs M, Tranel D. Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. Journal of Neuroscience. 2007;27(4):951-6.
32. Haruno M, Kimura M, Frith CD. Activity in the nucleus accumbens and amygdala underlies individual differences in prosocial and individualistic economic choices. Journal of Cognitive Neuroscience. 2014;26(8):1861-70.
33. Fischbacher U. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental economics. 2007;10(2):171-8.
34. Babiloni F, Astolfi L. Social neuroscience and hyperscanning techniques: past, present and future. Neuroscience & Biobehavioral Reviews. 2014;44:76-93.
35. Campbell-Meiklejohn DK, Bach DR, Roepstorff A, Dolan RJ, Frith CD. How the opinion of others affects our valuation of objects. Current Biology. 2010;20(13):1165-70.
36. Daw ND. Trial-by-trial data analysis using computational models. Decision making, affect, and learning: Attention and performance XXIII. 2011;23:1.
37. Hampton AN, Bossaerts P, O'Doherty JP. Neural correlates of mentalizing-related computations during strategic interactions in humans. Proceedings of the National Academy of Sciences. 2008;105(18):6741-6.
38. Behrens TE, Hunt LT, Woolrich MW, Rushworth MF. Associative learning of social value. Nature. 2008;456(7219):245-9.
39. Suzuki S, Harasawa N, Ueno K, Gardner JL, Ichinohe N, Haruno M, et al. Learning to simulate others' decisions. Neuron. 2012;74(6):1125-37.
40. Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical conditioning II: Current research and theory. 1972;2:64-99.
41. Prelec D. The probability weighting function. Econometrica. 1998:497-527.
42. Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. Nature neuroscience. 2007;10(9):1214-21.
43. Burnham KP, Anderson DR. Model selection and multimodel inference: a practical information-theoretic approach: Springer Science & Business Media; 2003.

Supplementary Figures and Legends:

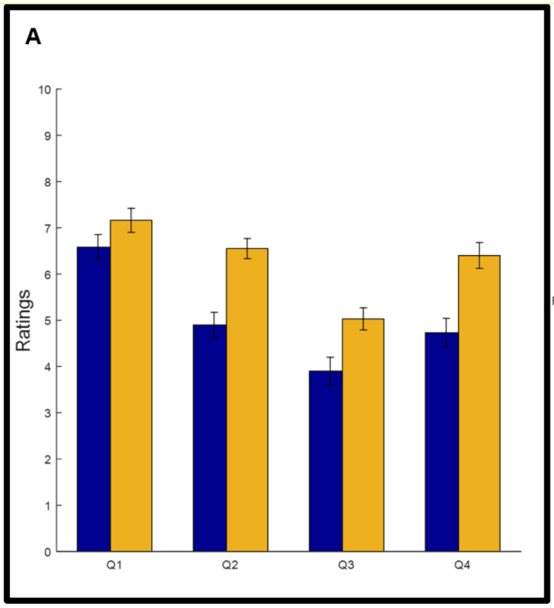

Fig. S1. Descriptive ratings participants made about social agents following the observational social-learning sessions. (A) Participants responded to a number of questions while thinking about the personality of these social agents (all rated on a Likert scale from 0 to 10).

Q1: How much do you think this person cares about the rewards to others?

Q2: How much would you like this person if you spent 1 hour with him/her in real-life?

Q3: How many people do you know in real-life who resembles this person?

Q4: How socially close do you feel towards those people that you know?

A fitted 4x2 MANOVA suggests there is a significant main effect of social agent (F=3.01, P=0.03) and a significant effect of the interaction term (F=10.16, P<0.001).

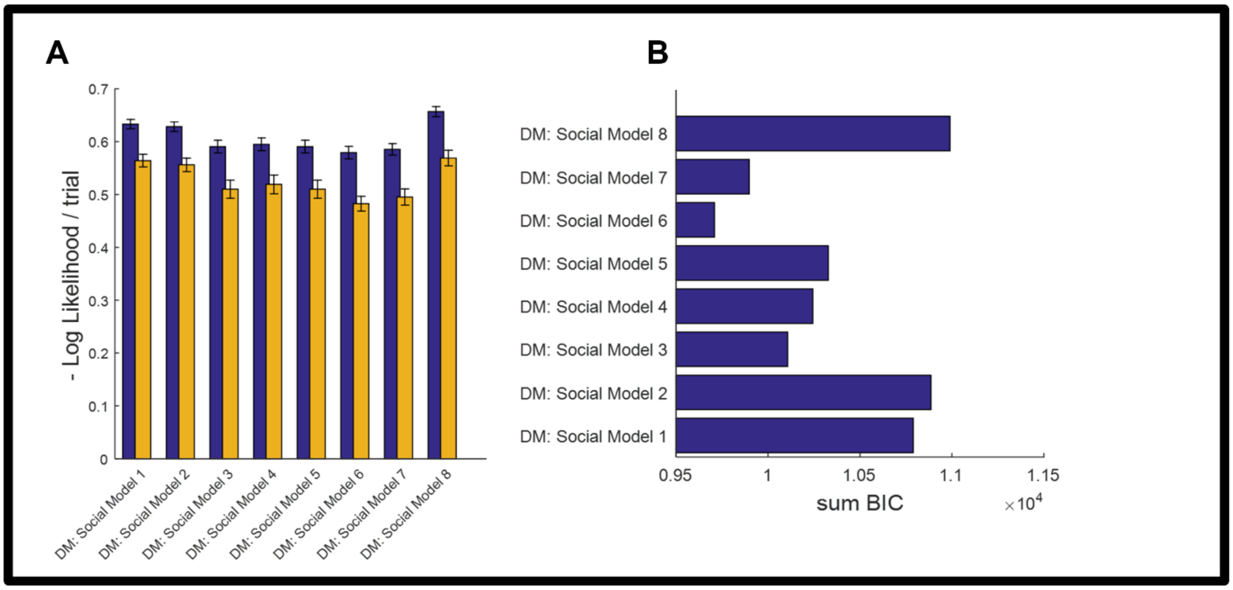

Fig. S2. (A) -Log likelihood values and (B) sum of BIC values for social decision-making models (shorter bars indicate better fitting).

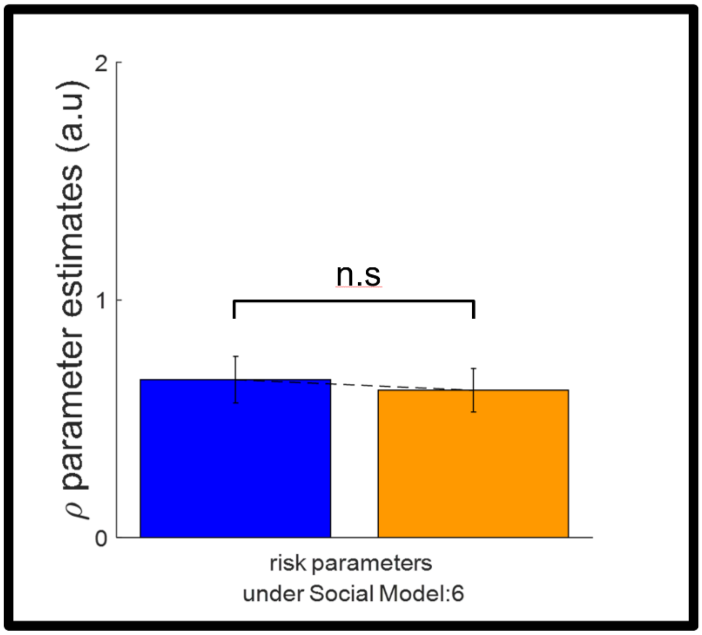

Fig. S3. Risk parameters (ρ_i^p) estimated from the Ultimatum giving experiments against two different social agents with different SVOs were not significantly different (P=0.74, n.s: not significant).

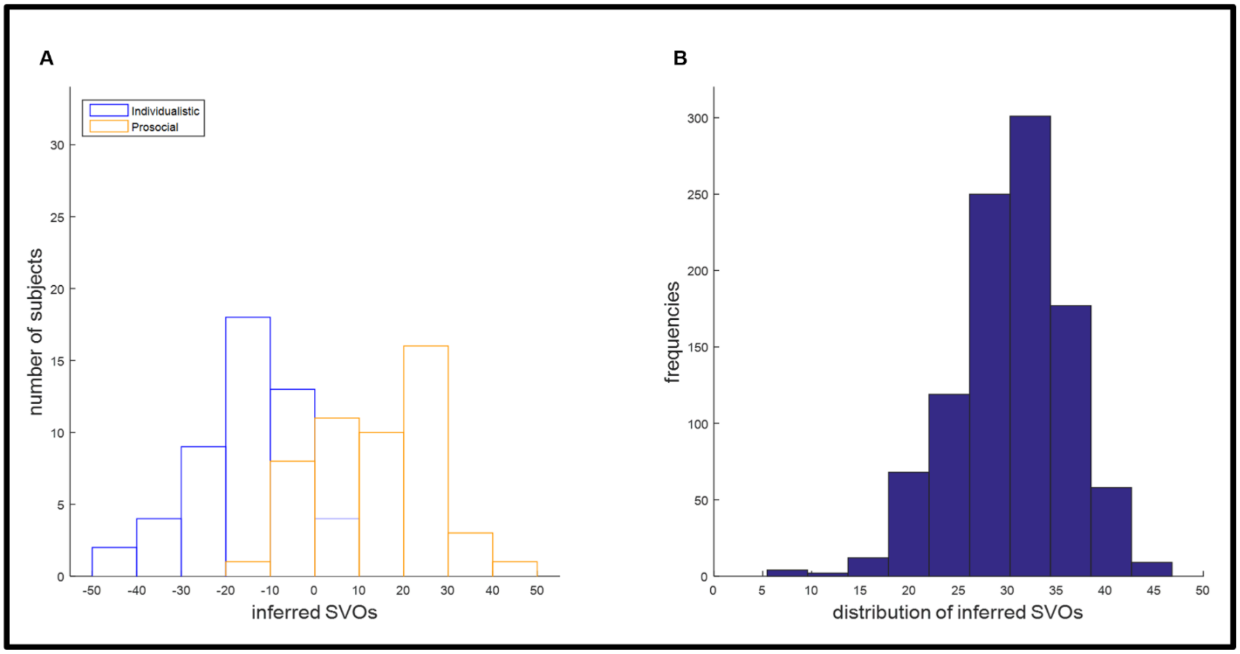

Fig. S4. Distributions of opponents' inferred SVOs ($S\tilde{V}O$) based on simulations of the encoded value functions ($\tilde{q}_A$) from the social-learning sessions. (A) Distribution of resulting inferred SVOs for the whole cohort (i.e. distribution of means obtained from each subject's simulation) based on 1000 simulations of the encoded value functions per participant per social agent (i.e. prosocial and individualistic). These distributions were highly significantly different from each other (*D*=0.78, P<0.001), suggesting that participants were able to make distinct inferences about the SVO of their opponents. (B) Distribution of inferred SVOs in 1000 simulations from a single subject gives a normal distribution of inferred SVOs, of which the mean and the standard deviation were used as input variables for the multiple linear regression analysis. Here the true SVO of the prosocial agent was 31.5° falling close to the mean of this distribution (μ =30.77° ± 5.4).

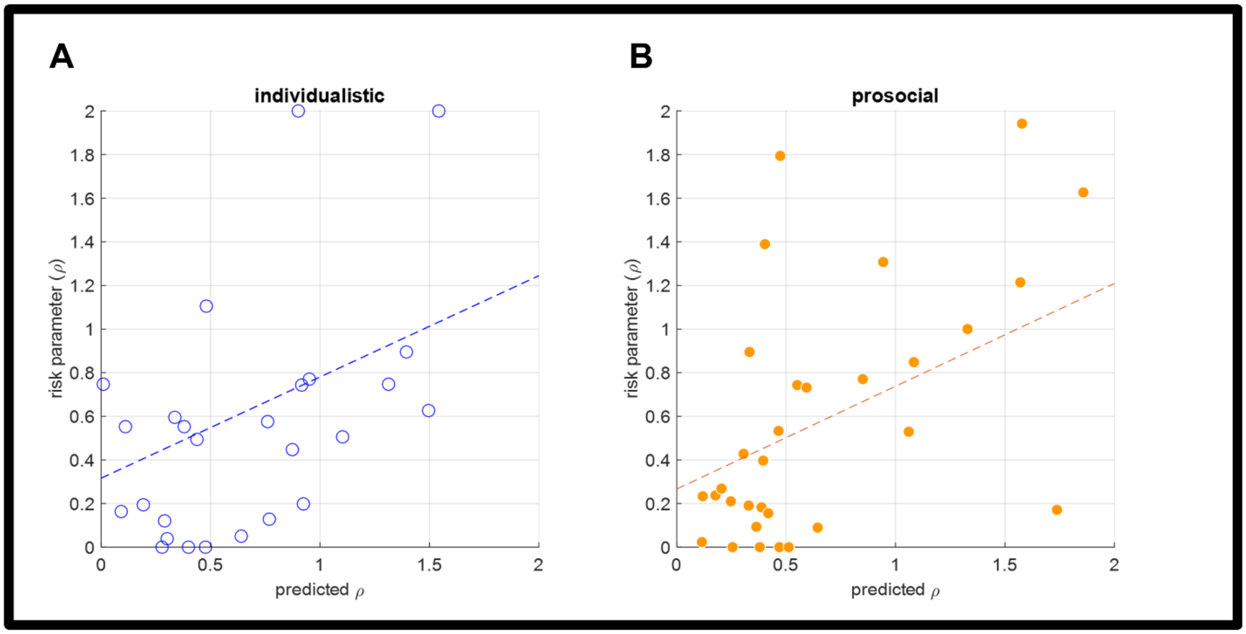

Fig. S5. Actual versus predicted values of the risk parameter (ρ) in the Ultimatum giving experiments. For both agents the predicted values correlated highly significantly with the actual parameter estimates (r>.79, p<.001, Bonferroni corrected).

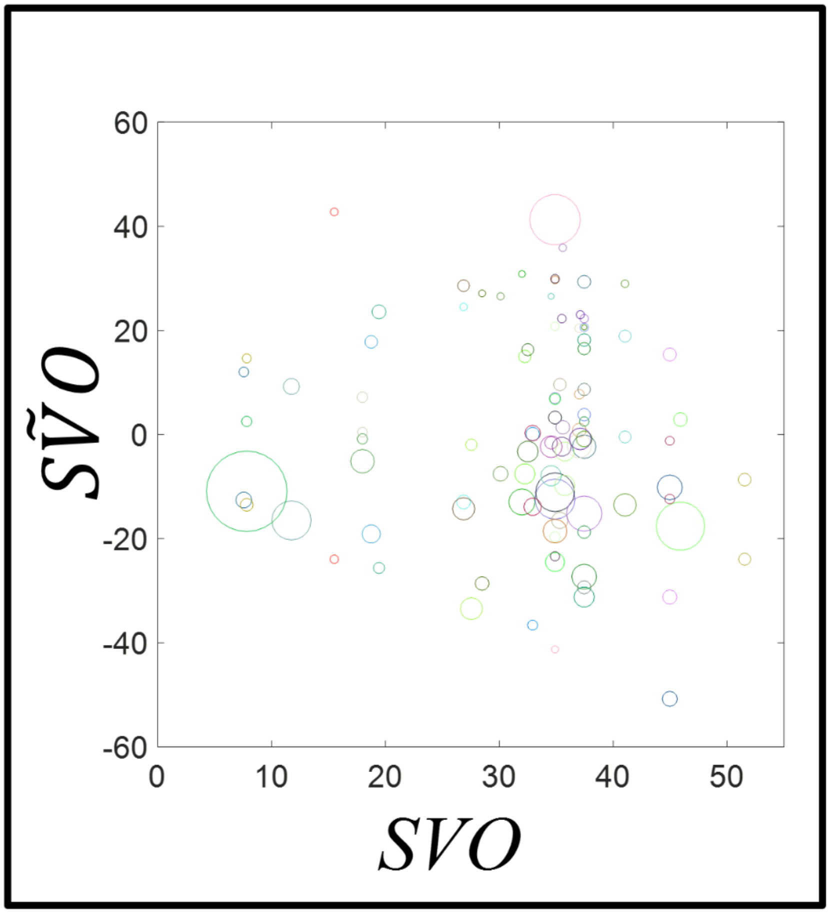

Fig. S6. Combinations of social interactions covered by the current study (n=100) with respect to participants' own (x-axis) and their inference of their opponents' SVO (y-axis). Markers with the same [R, G, B] colour coding refer to a single subject's data point and marker sizes are proportional to the uncertainty estimates ($\tilde{SVO}_\sigma$) which was included as an input variable in the multiple linear regression model described in the main text.