TITLE

Neurocomputational mechanisms of confidence in self and others

AUTHORS

Dan Bang[1,2,*], Rani Moran[1,3], Nathaniel D. Daw[4] & Stephen M. Fleming[1,3,5,*]

AFFILIATIONS

[1]Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3BG, UK

[2]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

[3]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London WC1B 5EH, UK

[4]Princeton Neuroscience Institute and Department of Psychology, Princeton University, Princeton, New Jersey 08544, USA

[5]Department of Experimental Psychology, University College London, London WC1H 0AP, UK

*Correspondence: danbang.db@gmail.com and stephen.fleming@ucl.ac.uk

ABSTRACT

Computing confidence in one's own and others' decisions is crucial for success in many social situations. There has been substantial progress in our understanding of confidence in oneself, but little is known about how we form confidence in others. Here, we address this question by asking subjects undergoing fMRI to place bets on perceptual decisions made by themselves or one of three other players of varying ability. We show that subjects' compute confidence in another player's decisions by combining distinct estimates of player ability and decision difficulty – allowing them to predict that a good player may get a difficult decision wrong and that a bad player may get an easy decision right. We find that this computation is supported by an interaction between brain systems traditionally implicated in decision-making (LIP) and theory of mind (TPJ, dmPFC). These results reveal a neurocomputational interplay between self- and other-related processes during social cognition.

INTRODUCTION

A key feature of adaptive behaviour is an ability to estimate confidence in our decisions in the absence of immediate feedback. For example, we may recognize that a choice was based on weak evidence and change our mind as conflicting but stronger evidence comes to light[1,2]. Recent years have seen substantial progress in our understanding of decision confidence. At a computational level, an emerging consensus is that our confidence in a choice reflects the probability that the choice is correct as computed using Bayesian inference[3–7]. At a neural level, converging evidence points to a central role for prefrontal cortex in the computation of decision confidence [8–10], the mapping of this internal estimate onto explicit reports[11,12], and the integration of single-trial confidence estimates into a long-run estimate of task performance[13].

Humans are, however, highly social and many situations require that we compute confidence not only in our own decisions but also in those of others. In competitive tasks, having a sense of who is more likely to succeed is useful for deciding whether to compete or opt out[14]. In cooperative tasks, having a sense of who is more likely to be correct is critical for assigning appropriate weight to others' advice[15,16], or deciding whose opinion to follow[17]. In addition, the confidence that others report is often biased[6], and can serve to manage social influence[12,18–20], and it is therefore useful to independently verify such reports. However, in contrast to the case of confidence in one's own decisions, little is known about the neurocomputational basis of confidence in decisions made by others.

Intuitively, the starting point for a social confidence computation is the same as in the individual case[7]. We should first assess the difficulty of the decision: the easier the decision is, the higher the probability that others will get it right. However, it is typically not sufficient to simply simulate what choice we would have made and how confident we would have felt about this choice. Instead, we should recognise that others' ability may be different from our own: what is hard for us may be easy for others and what is easy for us may be hard for others[7]. Previous studies have investigated how we track others' ability[15,21,22], but it remains unknown whether, or how, we integrate knowledge about ability with our own assessment of decision difficulty when computing confidence in others' choices. Critically, it is only by taking both of these factors into account that we can predict that a high-ability individual may get a difficult decision wrong and that a low-ability individual may get an easy decision right.

This requirement for integrating estimates of both decision difficulty and others' ability predicts that multiple brain systems support confidence in others' decisions. The first component, the estimation of decision difficulty, is computationally similar in both the individual and the social case and may therefore be supported by similar neural substrates. For example, in the context of perceptual decision-making, where decision difficulty is varied by changing the uncertainty associated with a sensory stimulus, we would expect areas that support sensory processing – such as visual motion area MT+ and lateral intraparietal sulcus (LIP)[8,23] – to be recruited when computing confidence in others' decisions. However, sensory representations on their own are unlikely to be sufficient because they do not reflect the characteristics of others' sensory systems. Instead, the second component to a social confidence computation, the estimation of others' ability, may be furnished by the so-called theory of mind (ToM) network – including temporoparietal junction (TPJ) and dorsomedial prefrontal cortex (dmPFC)[24–26] – believed to support the representation of others' attributes as distinct from one's own. In

support of this prediction, human fMRI has indicated that both TPJ and dmPFC are involved in the maintenance of running estimates of others' task performance[21,22,27].

Here, we develop a novel paradigm to disentangle the components of a social confidence computation, and we use behavioural analysis together with computational modelling to show that people take into account both decision difficulty and others' ability. More specifically, we asked subjects to place bets (post-decision wagers) on decisions on a random dot motion task[28] made by either themselves (self-trials) or one of three other players (other-trials) who differed in terms of their ability. To maximise reward on other-trials, subjects should combine an estimate of decision difficulty with distinct estimates of the other players' abilities. In this way, subjects can predict that a good player may get a difficult stimulus wrong and that a bad player may get an easy stimulus right. We refer to this account as the ToM model and we show that it provides a better fit to subjects' post-decision wagers than a set of alternative models (e.g., relying solely on self-projection or long-run estimates of others' task performance). In addition, analysis of fMRI data acquired during task behaviour indicated that the combination of distinct estimates of decision difficulty and others' ability was supported by an interaction between brain systems classically involved in decision-making (LIP) and ToM (TPJ and dmPFC). Overall, our results reveal a neurocomputational interplay between self- and other-related processes when forming confidence about others' decisions.

RESULTS

Experimental paradigm

Subjects ($N = 21$) placed post-decision wagers (PDWs) on either their own or another player's choices on a random dot motion task while undergoing fMRI (**Figure 1**). At the start of a trial, a screen indicated whether subjects had to perform a self-trial or an other-trial. On both trial types, subjects first viewed a field of moving dots inside a circular aperture. On each update of the motion stimulus, a fraction of dots moved coherently to the left or the right, whereas the remainder moved randomly. We varied the fraction of coherently moving dots in a parametric manner (range: .005-.5). On self-trials, subjects had to actively decide whether the average direction of dot motion was to the left or the right. On other-trials, subjects watched the motion stimulus while the other player decided; subjects were not informed which decision the other player made. After both trial types, a PDW screen appeared requiring subjects to choose between a safe and a risky option. The safe option delivered a small but certain reward. In contrast, the risky option delivered a greater reward if the self- or other-choice was correct and a corresponding loss if incorrect. We varied the PDW offers in a parametric manner (safe range: 5-20; risky range: 25-50). Finally, subjects received feedback about the PDW outcome and thereby the accuracy of the self- or other-choice. Subjects were paired with three players of varying average choice accuracy (low: .55; medium: .75; high: .95) in a block-wise manner (i.e., one player per scan run). The players were created using pilot data from subjects who had performed the random dot motion task without the social component. Subjects were told that the other players were not currently present and had made their decisions at another time.

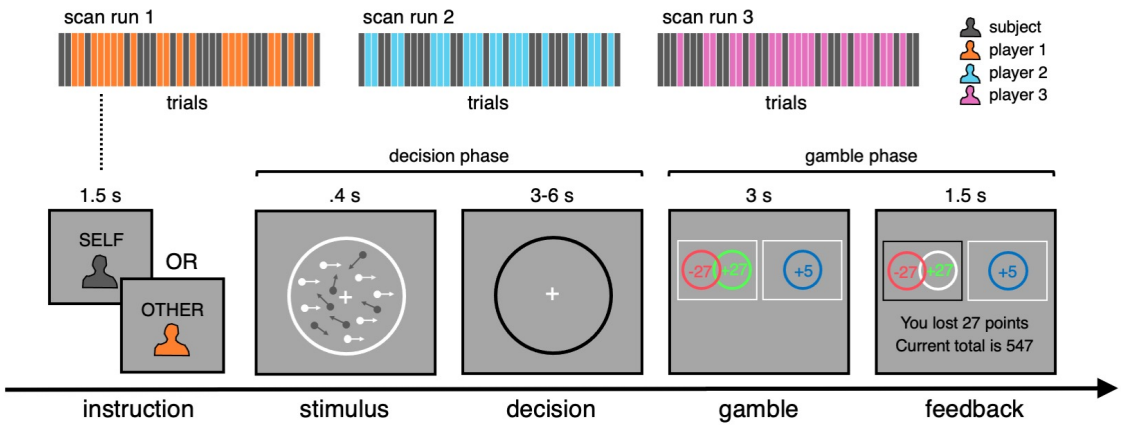

Figure 1. Social random dot motion task. On each trial, subjects placed a bet on a motion discrimination judgement (left or right) made by either themselves or one of three other players who varied in terms of their ability. The bet was made by choosing between a safe option, which yielded a small but guaranteed reward, and a risky option, which delivered a greater reward if the self- or other-choice was correct and a corresponding loss if incorrect. In this example, the subject chose the risky option (black box), but the self- or other-choice was incorrect. If subjects chose the safe option, then they were informed what the outcome would have been had they selected the risky option. Subjects were instructed how to calculate choice accuracy from the feedback screen. Subjects were paired with the three other players in a block-wise manner, and the trial type was signalled at the start of a trial. For fMRI analysis, we refer to the time window from the onset of the motion stimulus to 3 s after the onset of the decision screen as the decision phase and the time window from the onset of the gamble screen to the offset of the feedback screen as the gamble phase.

Post-decision wagering on self- and other-trials

We first sought to characterise subjects' PDWs behaviourally. Intuitively, the probability that subjects select the risky option should depend on the confidence that a self- or other-choice is correct and the difference between the rewards available under the risky and the safe option. The higher the confidence, the higher the probability that the risky option returns a reward, and the larger the reward difference, the higher the potential earnings from selecting the risky option. Importantly, self- and other-trials should differ in terms of the factors that enter into a confidence computation: on both trial types, confidence should take into account the difficulty of the decision, but on other-trials, confidence should also take into account the ability of the other player. To disentangle these contributions, we used a multiple logistic regression model to predict subjects' trial-by-trial PDWs, with predictors including trial type (other = -.5, self = .5), motion coherence, reward difference, others' ability (low = -.5, medium = 0, high = .5) and the interactions between trial type and the three other terms (see **Figure S1** for regression coefficients under full model and separate models for self- and other-trials).

The analysis showed that the probability that subjects selected the risky option increased with both motion coherence (**Figure 2A**; coherence: $t(20) = 7.78$, $p < .001$) and reward difference (**Figure 2B**; reward: $t(20) = 5.02$, $p < .001$), and that the impact of these factors did not depend on trial type (coherence x type: $t(20) = 0.18$, $p = .858$; reward x type: $t(20) = -0.44$, $p = .664$). The probability that subjects selected the risky option also increased with others' ability (ability: $t(20) = 2.85$, $p = .010$), but this effect did depend on trial type (type x ability: $t(20) = -4.60$, $p < .001$). Importantly, there was no effect of others' ability on self-trials (blue bars in **Figure 2C**; ability under separate model for self-trials: $t(20) = -1.46$, $p = .161$), whereas on other-trials, the higher the ability of the other player, the higher the probability that subjects selected the risky option (yellow bars in **Figure 2C**; ability under separate model for other-trials: $t(20) = 4.52$, p

< .001). Overall, subjects were more likely to select the risky option on self- than other-trials (type: $t(20) = 4.47$, $p < .001$) – with subjects on average selecting the risky option on 72% of self-trials and 61% of other-trials.

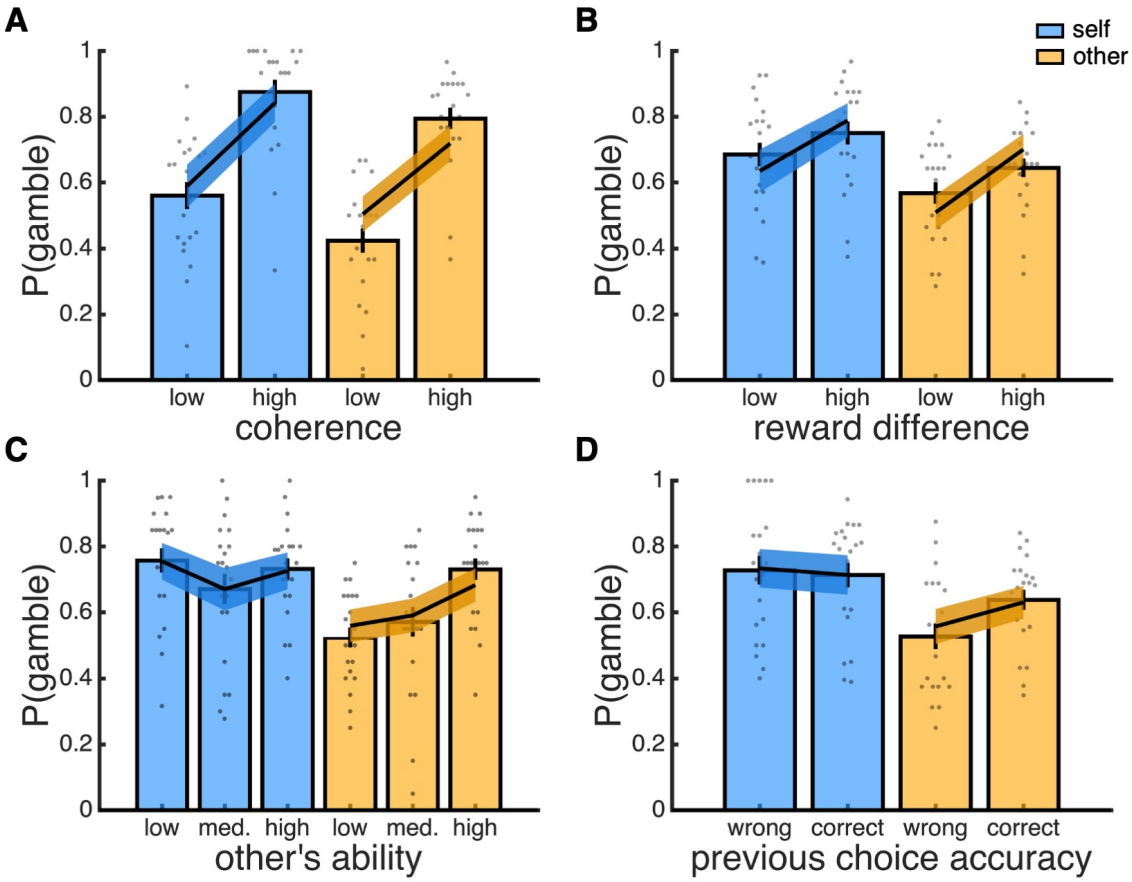

Figure 2. Behavioural results. In each panel, data are split into self-trials (blue) and other-trials (yellow). **(A)** Probability gamble as a function of coherence (median split). **(B)** Probability gamble as a function of the reward difference between the risky and the safe option (median split). **(C)** Probability gamble as a function of the ability of the other player. **(D)** Probability gamble as a function of choice accuracy on the previous trial of the same type. **(A-D)** Bar charts are empirical data, with individual subjects overlaid as dots. Lines are data simulated under the best-fitting model (ToM-model). Empirical and simulated data are represented as group mean ± 95% CI.

Critically, subjects were not informed about the ability of each player but had to learn it from experience. An obvious learning signal in our task is the feedback about choice accuracy available at the end of each trial. In line with this hypothesis, subjects were more likely to gamble on another player's choice when the player's previous choice was correct compared to when it was incorrect (yellow bars in **Figure 2D**; paired *t*-test: $t(20) = 2.97$, $p = .008$) – a history effect which was not present for self-trials (blue bars in **Figure 2D**: paired *t*-test, $t(21) = -0.35$, $p = .733$). Taken together, these results show (1) that subjects were equally sensitive to decision difficulty and reward characteristics on self- and other-trials and (2) that subjects tracked the abilities of the other players and used this knowledge to guide their PDWs in a player-dependent manner. We note that there was no effect of others' ability on subjects' task performance on self-trials, which we assessed using choice accuracy (logistic regression; coherence: $t(20) = 7.90$, $p < .001$; ability: $t(20) = -0.78$, $p = .444$) and reaction time (linear regression; coherence: $t(20) = -11.98$, $p < .001$; ability: $t(20) = 0.43$, $p = .674$).

Computational model of confidence in self and others

Qualitatively, the behavioural results fit with a ToM account in which the confidence in another player's decisions is based on both a self-related component, which reflects decision difficulty, and an other-related component, which reflects the ability of the other player. In order to test this interpretation formally, and to derive markers of the latent variables that underpin subjects' behaviour for fMRI analysis, we compared candidate models of subjects' PDWs.

In our task, subjects should decide whether to gamble on a choice by comparing the expected value of the risky and the safe option, $\Delta\mathrm{EV} = \mathrm{EV(risky)} - \mathrm{EV(safe)}$. We modelled the mapping between $\Delta\mathrm{EV}$ and gamble choices using a softmax function, $\mathrm{P(gamble)} = \mathrm{softmax}(\Delta\mathrm{EV})$, and we allowed the parameters of this function to vary between trial types to accommodate overall differences in the tendency to gamble on choices made by self and other. The expected value of the safe option is simply the amount offered, $\mathrm{EV(safe)} = \mathrm{V(safe)}$. By contrast, the expected value of the risky option must be computed by weighting the risky offer according to the confidence in the choice being correct, $\mathrm{EV(risky)} = \mathrm{P(correct)} \times \mathrm{V(risky)} - [1 - \mathrm{P(correct)}] \times \mathrm{V(risky)}$. We first consider the computation of confidence on self-trials, before turning to other-trials. All models are described in detail in the **Methods**.

Following earlier work on decision confidence[6,7], we modelled subjects' confidence on self-trials using Bayesian decision theory (**Figure 3**). We assumed that subjects represent the stimulus space as comprised of a set of distinct motion stimuli, each defined by a direction and a coherence. In the first version of the model, the representation consisted of linearly spaced motion stimuli. In a modified version, we allowed for more low- than high-coherence motion stimuli, in line with a Weber-Fechner law in which the resolution of sensory perception diminishes for stimuli of greater magnitude. On each trial, subjects receive noisy sensory evidence, modelled as a sample from a Gaussian distribution centred on the true motion stimulus. Subjects then compute a belief state over stimulus space – that is, the probability that each motion stimulus is shown on the current trial – given the sensory evidence and their own level of sensory noise. Finally, subjects use this belief state to generate both a decision and a confidence in this decision being correct.

There are different ways in which a ToM account of a social confidence computation could be implemented, but – for consistency with earlier work on decision confidence[6,7] – we considered an extension of the Bayesian framework to the social case (see Discussion for a heuristic ToM account). In brief, the ToM-model assumes that subjects compute confidence in the other player's decision by integrating their belief state over stimulus space with a distinct representation of the other player's expected accuracy for each motion stimulus (**Figure 3**). The ToM-model integrates these representations across the full stimulus space since subjects do not know whether the other player decided left or right. The ability representation, which is, in effect, a belief about the other player's psychometric function, is directly derived from an estimate of the other player's sensory noise. This estimate is updated at the end of each trial according to the difference between the observed and the predicted success of the other player's choice – a social prediction error. Inspired by previous work[7], we also considered a modified version of this model (ToM*-model), which assumed shared error variance between the subject's and the other player's sensory evidence – an assumption that subjects may make because they both viewed the same noisy instantiation of a motion stimulus.

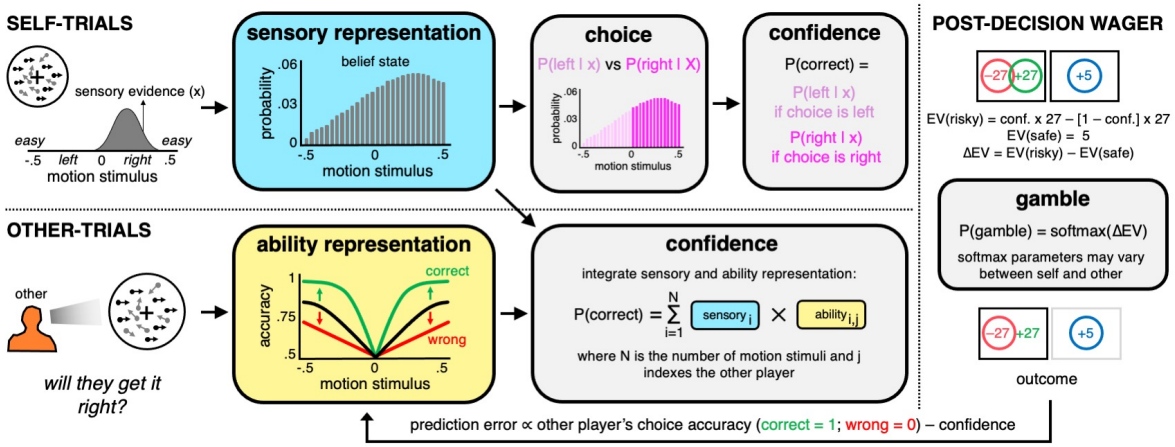

Figure 3. Illustration of self- and other-trials under the ToM-model. All models assumed that subjects' choices and confidence were computed on self-trials according to Bayesian decision theory. Subjects represent the stimulus space as comprised of a set of distinct motion stimuli, each defined by a direction (sign) and a coherence (absolute value). On each trial, subjects receive sensory evidence, sampled from a Gaussian distribution whose mean is given by the true motion stimulus and whose standard deviation is given by the subject's sensory noise. Subjects compute a belief state over the stimulus space given the sensory evidence and their sensory noise (blue box) and use this belief state to generate both a decision about the motion direction and their confidence in this decision being correct. The ToM-model assumes that subjects compute confidence in others' choices by combining their belief state over stimulus space with a representation of the other player's expected accuracy for each motion stimulus (yellow box). The latter representation – in effect, a psychometric function – is derived from an estimate of the other player's sensory noise. This estimate is updated at the end of each trial using a Rescorla-Wagner learning rule that takes into account the accuracy of the other player's choice and the confidence in this choice being correct. As a result of this update, the psychometric function becomes steeper after a correct choice (green) and shallower after an incorrect choice (red). On both trial types, confidence is used to calculate the difference in expected value between the risky and the safe options. Finally, the difference in expected value is entered into a softmax function to obtain the probability of selecting the risky option. The softmax parameters were allowed to vary between self- and other-trials.

Under a ToM account, confidence on other-trials is influenced by both decision difficulty and others' ability. For completeness, and to rule out extraneous explanations of the behavioural results, we also considered models that only took into account one of these components, or that did not adequately disentangle the two components. First, we considered a self-projection model (S-model), which assumes that subjects on other-trials ask themselves which choice they would have made and how confident they would have felt about this choice. Under the S-model, confidence on other-trials is influenced solely by decision difficulty and not by others' ability. Second, building on previous work[21], we considered a performance-tracking model (Q-model), which assumes that subjects use a Rescorla-Wagner rule to learn about the value of each player based on their historical choice accuracy and use this value estimate as a proxy for confidence in their decisions. Under the Q-model, confidence on other-trials is influenced solely by others' ability and not by decision difficulty.

Finally, inspired by a recent account of confidence in animals[29], we considered a modified performance-tracking model (Q*-model) under which confidence on other-trials is influenced by both decision difficulty and others' ability. Specifically, the Q*-model assumes that subjects use a Rescorla-Wagner rule to learn about the value of each motion coherence for each player and integrate these value estimates with a belief state over the space of motion coherence in order to compute confidence in the other player's decision. In other words, unlike the Q-model that tracks each player's historical choice accuracy overall, the Q*-model tracks each player's

historical choice accuracy for each motion coherence. However, in contrast to a ToM account where separate representations of decision difficulty and others' ability are maintained, the Q*-model makes counterintuitive predictions: for example, after experiencing a correct decision at low coherence, it may learn that a low-coherence stimulus is more likely to result in a correct choice than a high-coherence stimulus.

In line with the behavioural results, and in support of a ToM account of a social confidence computation, the ToM-model provided the best fit to subjects' PDWs – with the winning model containing a Weber-Fechner law and separate softmax parameters for self and other (see **Figure 2** for data simulated under ToM-model, **Figure S2** for model comparison and **Figure S3** for data simulated under best-fitting versions of all candidate models).

Neural basis of a social confidence computation

We next turned to the fMRI data acquired during task behaviour to identify neural substrates that may contribute to a social confidence computation. The behavioural and modelling results supported a ToM account in which subjects compute confidence in another player's decisions by integrating distinct estimates of decision difficulty and others' ability. The first component, a representation of decision difficulty, is directly linked to the motion stimulus and in particular motion coherence: the higher the coherence, the easier the decision. The second component, a representation of others' ability, must be tracked separately from one's own ability and be learned from feedback about the other player's choice accuracy. We would therefore expect brain systems traditionally involved in sensory and social processing to interact during a social confidence computation and in turn that the identified "social" areas carry learning signals relating to others' task performance.

To identify brain regions that were differentially activated by self- and other-related processing, we first used a whole-brain contrast between self- and other-trials during either the decision phase or the gamble phase (**Figure 1**). During the stimulus phase, self-related processing was associated with higher activity in classic perceptual decision-making regions, including extrastriate cortex, posterior parietal cortex and cingulate cortex (**Figure 4**). By contrast, other-related processing was associated with higher activity in the ToM network, including middle temporal gyrus, TPJ and dmPFC (**Figure 4**). During the gamble phase, self-related processing was not associated with differential activity, whereas other-related processing was again associated with higher activity in the ToM network (results not shown for the gamble phase – see **Table S2** and **Table S3** for all clusters surviving whole-brain correction for the decision and gamble phases).

We focused on a subset of these brain regions – using independent ROI masks (**Figure 5**) – in order to assess the neural evidence for a ToM account of a social confidence computation. First, we hypothesised that visual motion area MT+, which we identified in a separate localiser scan at the end of the scan session, and a human homologue of monkey LIP, encompassing posterior parts of the superior parietal lobule and the intraparietal sulcus[30], may provide a sensory representation of the motion stimulus that underlies confidence not only in one's own but also in others' choices. Both regions have consistently been shown to support perceptual decision-making on the random dot motion task in humans and non-human primates[8,31–34]. Second, we hypothesised that TPJ and dmPFC, often implicated in social inference[21,22,35],

may combine a sensory representation of the motion stimulus with a distinct representation of others' ability as required by a social confidence computation.

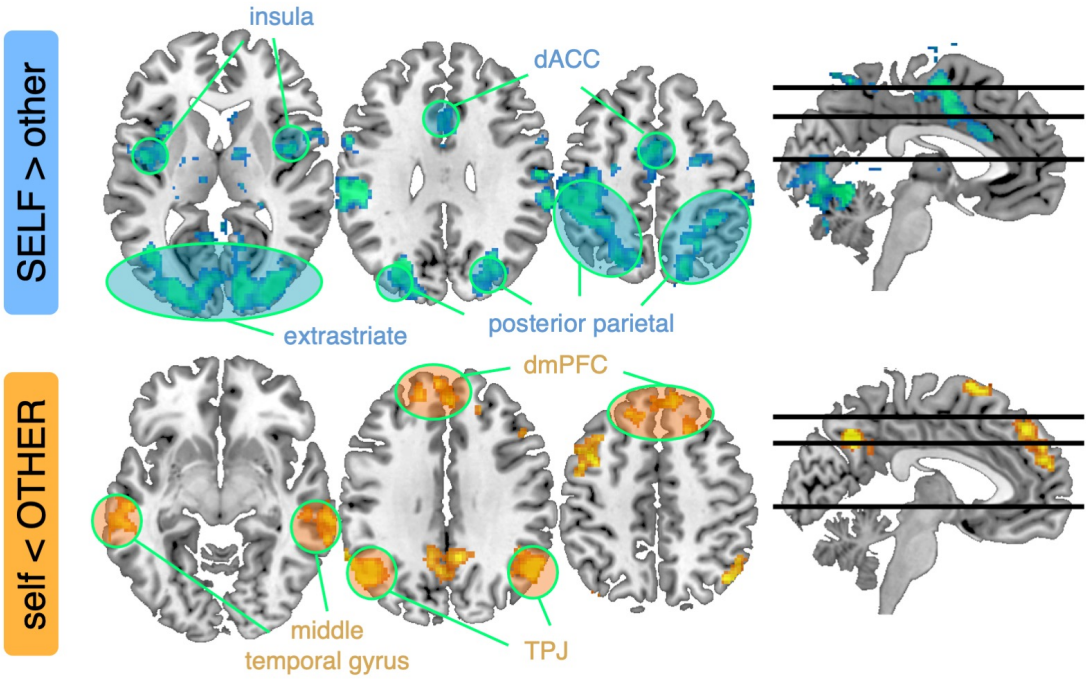

Figure 4. Whole-brain activations for self- and other-processing during decision phase. Images display clusters that survive whole-brain correction ($p < .05$, FWE-corrected for multiple comparisons at a cluster-defining threshold of $p < .001$, uncorrected) for contrast between self- and other-trials during the decision phase (cold: self > other; warm: self < other). Images are shown at $p < .001$, uncorrected. All clusters surviving whole-brain correction during the stimulus and gamble phases are detailed in **Table S2** and **Table S3**.

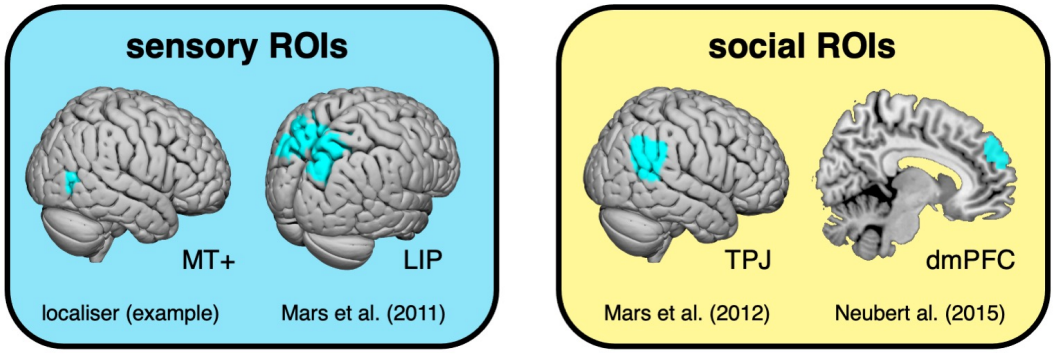

Figure 5. Anatomical masks for ROI analyses. The MT+ mask was specified for each subject using a localiser scan. The other masks were specified using published connectivity-based parcellation atlases.

The whole-brain analysis showed that classic perceptual decision-making regions – including MT+ and LIP – were more active on self- than other-trials during the decision phase. However, a ToM account predicts that regions supporting the formation of a sensory representation of the motion stimulus should also contribute to a social confidence computation. In other words, we would expect our sensory ROIs to carry information about motion coherence on both self- and other-trials. To test this hypothesis, we performed a sliding-window regression for each trial type in which we quantified whether MT+ and LIP activity time courses locked to stimulus

onset tracked motion coherence. As expected under a ToM account, the analysis showed that the sensory ROIs tracked motion coherence on both trial types (**Figure 6**). The MT+ response profile was the same on self- and other-trials: the higher the coherence, the higher the activity. In contrast, the LIP response profile differed between trial types: the higher the coherence, the lower the activity on self-trials, but the higher the activity on other-trials.

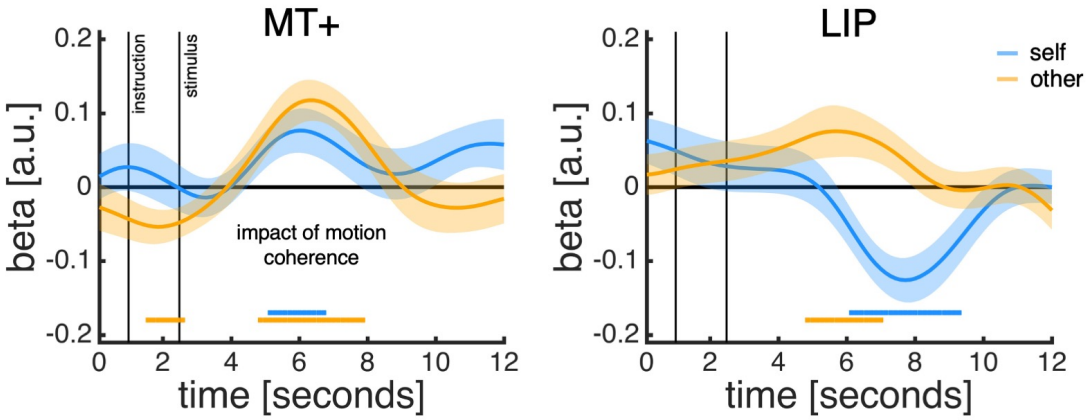

Figure 6. Encoding of motion coherence in MT+ and LIP. Coefficients from a regression analysis in which we predicted z-scored MT+ (*left*) and LIP (*right*) activity time courses separately for self-trials (blue) and other-trials (yellow) using z-scored motion coherence. We assessed statistical significance (squares) for each time point by asking whether coefficients across participants were different from zero ($P < .05$, one-sample *t*-test). Data are represented as group mean ± SEM.

The LIP response pattern is consistent with previous fMRI reports of posterior parietal activity during active choice (self-trials) and passive viewing (other-trials) conditions. When subjects are required to make a choice, posterior parietal activity has been reported to decrease with motion coherence[33]. In contrast, during passive viewing, posterior parietal activity has been found to be higher for coherent motion stimuli than dynamic noise[36]. One explanation of the inverse relationship between posterior parietal activity and motion coherence during active choice is that LIP neurons integrate sensory information into a belief state over stimulus space[37] – a process which usually terminates earlier for high-coherence motion stimuli and thereby may give rise to reduced bulk activity for high-coherence motion stimuli as measured with fMRI. Indeed, when motion coherence and reaction time are experimentally dissociated, posterior parietal activity increases with motion coherence[8].

Having found that MT+ and LIP track motion coherence on both trial types, we next turned to the social ROIs in order to assess the neural evidence for the second component of a social confidence computation. The whole-brain analysis showed that the ToM network – including TPJ and dmPFC – was more active on other- than self-trials during both trial phases. However, a ToM account predicts that the ToM network should not only discriminate between self and other but also carry information about the ability of the other players. To test this hypothesis, we performed a sliding-window regression in which we quantified whether TPJ and dmPFC activity time courses locked to stimulus onset tracked variation in confidence in others' choices – as computed under the ToM-model – that was driven by the dynamically evolving representation of others' ability (other's sensory noise) (**Figure 7A**). We isolated this aspect of a social confidence computation by orthogonalising the model-based confidence estimates with respect to motion coherence. Consistent with a ToM account, the analysis showed that

both TPJ and dmPFC tracked features of a social confidence computation that were directly linked to trial-by-trial learning about others' ability (**Figure 7B**).

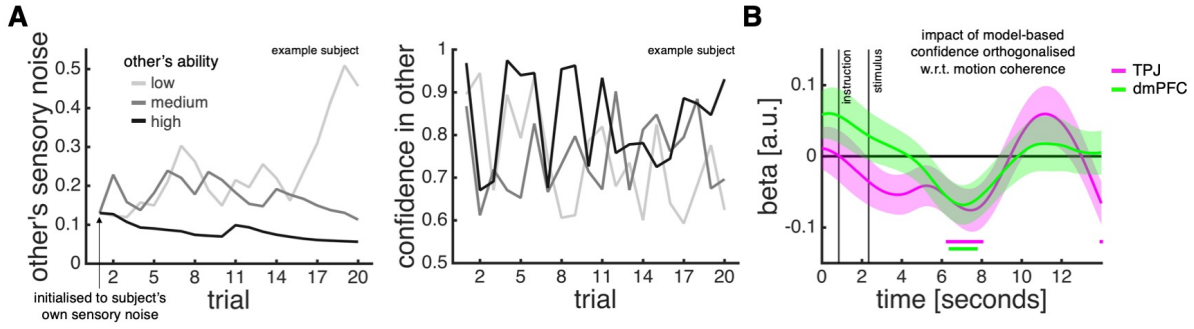

Figure 7. TPJ and dmPFC support a social confidence computation. **(A)** Trial-by-trial estimate of other player's sensory noise (*left*) and confidence (*right*) in their decisions as computed under the best-fitting ToM-model for an example subject. **(B)** Coefficients from a regression analysis in which we predicted z-scored TPJ (*left*) and dmPFC (*right*) activity time courses on other-trials using z-scored model-based confidence (first orthogonalised with respect to motion coherence). We assessed statistical significance (squares) for each time point by asking whether coefficients across participants were different from zero ($P < .05$, one-sample *t*-test). Data are represented as group mean ± SEM.

We reasoned that, if the ToM network receives a sensory representation of the motion stimulus from classic perceptual decision-making regions and in turn furnishes it with a representation of another player's ability during a social confidence computation, then functional connectivity between our sensory and social ROIs should be higher on other- than self-trials. To test this hypothesis, we performed a psychophysiological interaction (PPI) analysis using MT+ or LIP as seed region and trial type as psychological variable (**Figure 8A**). In line with a ToM account, the PPI analysis revealed that connectivity between LIP and TPJ/dmPFC was higher on other- compared to self-trials (**Figure 8B**; one-sample *t*-test: TPJ, $t(20) = 2.50$, $p = .021$; dmPFC, $t(20) = 2.56$, $p = .019$) – a pattern also evident in TPJ and dmPFC activity time courses locked to stimulus onset (**Figure 8C**). These relationships were not found for MT+ (one-sample *t*-test: TPJ, $t(20) = 0.51$, $p = .615$; dmPFC, $t(20) = -0.80$, $p = .432$).

Finally, having examined how sensory and social regions interact to compute confidence in others' choices, we analysed how the representation of another player's ability is updated with task experience. The behavioural and modelling results indicated that subjects revised their estimate of the other player's sensory noise based on the difference between the accuracy of the other player's choice and their confidence in this choice being correct (see schematic of learning mechanism in **Figure 3**). Consistent with a role in supporting this estimate, a sliding-window regression revealed that TPJ and dmPFC activity time courses locked to outcome onset tracked social prediction errors as computed under the ToM-model (**Figure 9**). We note that these response profiles remained after orthogonalising the social prediction error with respect to the outcome term (accuracy) or the prediction term (confidence).

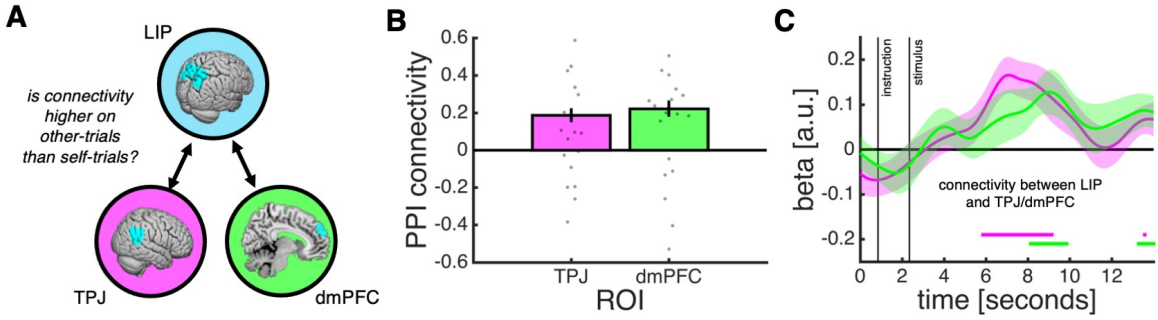

Figure 8. Functional coupling between sensory and social ROIs during social confidence computation. **(A)** Schematic of PPI analysis testing whether the correlation between LIP activity (seed region) and TPJ/dmPFC activity is higher on other- than self-trials. **(B)** Contrast estimates from PPI analysis. **(C)** Coefficients from a PPI regression analysis in which we predicted z-scored TPJ/dmPFC activity time courses using an interaction between z-scored LIP activity time courses and trial type (self = -.5; other = .5), while controlling for the main effect of each term. We highlight that the results remained the same when we controlled for main or interaction effects in relation to motion coherence. We assessed statistical significance (squares) for each time point by asking whether coefficients across participants were different from zero ($P < .05$, one-sample *t*-test). Data are represented as group mean ± SEM.

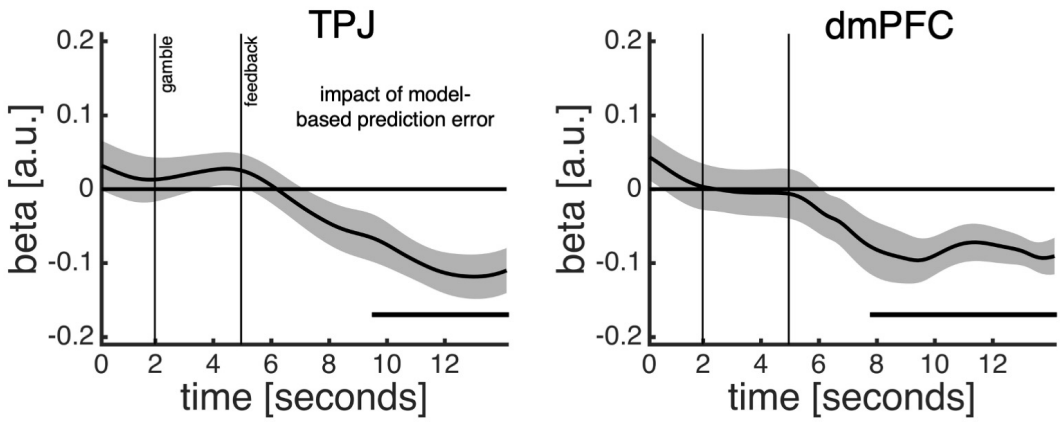

Figure 9. Encoding of social prediction errors in TPJ and dmPFC. Coefficients from a regression analysis in which we predicted z-scored TPJ (*left*) and dmPFC (*right*) activity time courses using z-scored social prediction errors as computed under the best-fitting ToM-model. We assessed statistical significance (squares) for each time point by asking whether coefficients across participants were different from zero ($P < .05$, one-sample *t*-test). Data are represented as group mean ± SEM.

DISCUSSION

Computing confidence in others' decisions is critical for success on a range of competitive and cooperative tasks. While previous research has examined how we track others' ability, little is known about whether, or how, we combine this knowledge with an assessment of the decision problem at hand to form a confidence estimate. In many situations, simply knowing someone else's ability is insufficient for an accurate confidence computation as the difficulty of a decision may vary substantially. For example, an excellent medical doctor may not be able to diagnose a complex case, whereas someone who is generally bad at mathematics may still be able to solve a simple arithmetical problem. In situations like these, distinct estimates of others' ability and decision difficulty are needed in order to predict whether someone else will be successful at a given moment in time.

Here we addressed how people solve this problem using a perceptual decision task in which subjects placed bets on choices about stimuli of varying difficulty made by either themselves

(self-trials) or one of three other players (other-trials) who differed in terms of their ability. By combining behavioural analysis and computational modelling, we found that subjects tracked the ability of each player and combined this knowledge with an estimate of decision difficulty in order to decide whether to gamble on the other player's choice being correct. In addition, we found, using fMRI, that self- and other-trials differentially activated brain regions typically involved in decision-making (e.g., LIP and dACC) and social cognition (e.g., TPJ and dmPFC). Critically, these brain regions also interacted during other-related processing (i.e., increased connectivity between LIP and TPJ/dmPFC on other-trials compared to self-trials), supporting the interpretation that a social confidence computation takes into account distinct estimates of both decision difficulty and others' ability.

Our winning model, the ToM model, is a natural extension of Bayesian models of confidence from the individual to the social case[7]. In this model, decision difficulty and others' ability are taken into account by integrating a belief state over stimulus space with a representation of the other players' expected accuracy for each stimulus. Our goal was to test whether a social confidence computation takes into account both decision difficulty and others' ability, rather than seeking to arbitrate between alternative algorithms that may support this computation. For example, people may solve the problem in a more model-free manner, by first calculating the confidence in the decision that they themselves would have made and then adjusting this estimate by a scalar value that reflects the performance of the other player (e.g., a combination of the S- and Q-models). This heuristic account and the ToM model, which both build on the hypothesis that a social confidence computation involves self- and other-related components, make subtly different predictions (e.g., the difference in confidence for two other players varies with motion coherence under the ToM model but is constant under the more model-free alternative once learning has completed). Future studies are needed in order to disambiguate these related accounts (e.g., extensive psychophysical mapping of the relationship between motion coherence and confidence in others' choices).

The question of how people infer what others think, feel or intend, and the relationship between inference about one's own and others' internal states, has attracted widespread attention in philosophy, psychology and neuroscience[26,38,39]. Our findings are in line with a view that self- and other-inference are distinct processes that involve distinct representations of oneself and others[24,40] – and go against a view that other-inference relies on self-projection[41,42] or that self-inference co-opts processes evolved for inference about others[39]. At a behavioural level, we found that a social confidence computation involves distinct self- and other-related components – people did not simply simulate which choice they would have made and how confident they would have felt about this choice. At a neural level, we found that inference about one's own and other's choices differentially activated brain systems often implicated in perceptual decision-making and ToM – indicating that confidence computations in the individual and the social case do not involve identical processes and/or representations. In support of this neural dissociation, a repetition suppression analysis, which indexes the information-selectivity of neuronal populations[43], revealed that activity in LIP (but not MT+) was highest on self-trials preceded by other-trials and lowest on other-trials preceded by other-trials, whereas activity in both TPJ and dmPFC was highest on other-trials preceded by self-trials and lowest on self-trials preceded by self-trials (**Figure S4**).

More broadly, our study extends our understanding of the neural basis of social cognition. TPJ and dmPFC have previously been shown to support learning of summary statistics that reflect others' ability on a variety of tasks, including the reliability of advice on a bandit task[35], recency-weighted performance on a reaction-time task[21], or prediction accuracy on a stock-market task[22]. Our results show that TPJ and dmPFC also support the integration of such knowledge about others' ability with information about the current state of a task – a process that facilitates accurate predictions about others' behaviour under varying task difficulty. However, we note that subtle differences between the self- and other-conditions of our task mean that we cannot conclude that TPJ and/or dmPFC are specialised for social inference[44]. First, unlike self-trials, subjects did not observe a choice on other-trials and had to consider that either choice could have been made (i.e., marginalise over choices). Second, unlike self-trials, subjects had to learn a new mapping between stimulus space and choice accuracy for each block of other-trials. It remains to be seen whether self-evaluation would co-opt putative social brain areas to the same degree as other-inference in situations where counterfactual thinking is required[45], or where one's own ability also varies across time[39].

In summary, our study shows (1) that people form confidence in others' decisions by flexibly combining knowledge about others' ability with their own assessment of the decision problem at hand and (2) that this computation is supported by an interaction between brain systems traditionally involved in decision-making and social cognition. We highlight, however, that our study only addresses one of the two sides of a social confidence computation. In many cases, adaptive control of social behaviour requires not only that we can predict the success of others' choices – as examined here – but also that we can infer how confident others themselves feel about these choices. For example, in a strategic game like chess, our response to a mistake made by our opponent may depend on whether we think that they themselves realise that they made a mistake. How our confidence in others interacts, in a potentially reciprocal manner, with inferences about others' confidence, awaits further study.

AUTHOR CONTRIBUTIONS

S.M.F. and N.D.D. conceived the study. S.M.F. collected the data. D.B., R.M., N.D.D and S.M.F developed the models. D.B. analysed the data. D.B., R.M., N.D.D. and S.M.F. interpreted the results. D.B. wrote the manuscript. R.M., N.D.D. and S.M.F. provided critical comments. All authors approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Sara Ershadmanesh and Matthew Rushworth for helpful comments on an earlier version of the manuscript. The Wellcome Centre for Human Neuroimaging was supported by core funding from the Wellcome Trust (203147/Z/16/Z). D.B. was supported by a Sir Henry Wellcome Postdoctoral Fellowship funded by the Wellcome Trust (213630/Z/18/Z). N.D.D. was supported by the John Templeton Foundation (61454). S.M.F. was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (206648/Z/17/Z). For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

SUBJECT DETAILS

Subjects performed a random dot motion task in a pre-scan and a scan session conducted on the same day. Twenty-two adults with no history of psychiatric or neurological disorders took part in the study. One subject was excluded due to poor performance in the pre-scan session, leaving twenty-one subjects for analysis (12 female, mean ± SD age = 22.6 ± 4.4 years). Subjects provided informed consent, and the study was approved by New York University's University Committee on Activities Involving Human Subjects. Subjects received a show-up fee of US$15, an additional US$25 for completing the MRI scan and were informed they could earn an additional performance-based bonus (in reality, all subjects received a US$9 bonus).

METHOD DETAILS

Experimental paradigm

Apparatus

The experiment was programmed in MATLAB 2014b using Psychtoolbox 3.0.12[46,47]. In the pre-scan session, stimuli were presented on an iMac desktop monitor viewed at a distance of approximately 45 cm. In the scan session, stimuli were presented via a projector at an approximate viewing distance of 58 cm.

Random dot kinematograms

Subjects viewed so-called random dot kinematograms (RDKs) contained in a circular aperture (7 degrees in diameter). Each RDK was made up of three independent sets of dots (each dot was 0.12 degrees in diameter) shown in consecutive frames. Each set of dots were shown for one frame (about 16 ms) and then replotted again three frames later (about 50 ms). Each time a set of dots was replotted, a subset of the dots, determined by the motion coherence, θ, was displaced in the direction of motion at a speed of 6 degrees s^{-1}, whereas the rest of the dots were displaced at random locations within the aperture. The motion direction, k, was to the left or to the right along the horizontal meridian. The use of three sets of dots means, for example, that dot positions in frame one are correlated with dot positions in frame four, seven and so forth. The dot density was fixed at 30 dots $degrees^{-2}$ s^{-1}.

Pre-scan session: familiarisation

Subjects were familiarised with the basic random dot motion task prior to the scan session. Each trial began with the presentation of a fixation cross at the centre of a circular aperture. After a fixed delay (.5 s), subjects viewed an RDK (.4 s). Once the RDK had terminated, subjects were asked to indicate whether the net direction of dot motion was to the left or to the right along the horizontal meridian. In particular, subjects were presented with a white box on the left and a white box on the right, corresponding to the two choice options. Once a choice had been made, the outline of the chosen option turned black (.5 s), and subjects received feedback about choice accuracy (low-pitched tone: error; high-pitched tone: correct). If subjects did not respond within 4 s, then the choice was scored as an error. The coherence, θ, was taken from the set, $\Theta \in \{.03, .06, .12, .24, .48, 1\}$, counterbalanced across trials. The motion direction, k, was sampled randomly from left or right. Subjects completed 240 trials.

Scan session: self-other

Subjects performed the same random dot motion task again, except that they were now also required to place post-decision wagers (PDWs) on the accuracy of either their own or another player's choices.

Each trial began with the presentation of a fixation cross at the centre of a circular aperture (1.5 s). Subjects then viewed a message instructing either a 'self-trial' or an 'other-trial' (1.5 s). On both trial types, after a brief delay (.5 s), subjects viewed an RDK (.4 s). After the RDK had terminated, subjects entered the choice screen – with the fixation cross and the circular aperture remaining visible. On self-trials, subjects actively decided about the net direction of dot motion. On other-trials, subjects passively viewed the screen while the other player made a decision about the motion stimulus – subjects were not informed whether these decisions were to the left or right. On both trial types, the circular aperture changed its colour to black once a choice had been made. The other player's choice reaction time was sampled uniformly from the range .5-1 s. On self-trials, if subjects did not respond within 3 s, then the trial was aborted – subjects had to wait for the duration that a full trial would have taken before proceeding to the next trial. The delay between the offset of the RDK and the offset of the choice screen was sampled uniformly from the range 3-6 s.

Subjects were then presented with a PDW screen (3 s) containing a safe and a risky option, indicated by a white box on each side of the screen (random assignment). The safe option, indicated by a blue circle, delivered a small but certain reward regardless of choice accuracy. The risky option delivered a larger reward, indicated by a green circle, if the self- or other-choice was correct, and a corresponding loss, indicated by a red circle, if the self- or other-choice was wrong. The safe value was drawn uniformly from the range 5-20 points, and the risky value was drawn uniformly from the range 25-50 points. Once subjects had made a PDW, they were presented with a feedback screen (1.5 s). The screen indicated the chosen PDW option (relevant box highlighted in black) and the accuracy of the choice (if correct, the green circle maintained its colour, whereas the red circle turned whited, and if incorrect, the red circle maintained its colour, whereas the green circle turned white). In this way, subjects were informed about the outcome on the current trial as well as the outcome that would have been obtained had they chosen the other PDW option. In addition, the screen also indicated the total earnings accumulated up to the current trial. Subjects started the scan session with 500 points and were told that their performance-based bonus depended on the total amount of points accumulated across the experiment. If subjects did not make a PDW within 3 s of the PDW onset, then they lost the points associated with the safe option.

Subjects performed 3 blocks (scan runs) consisting of 40 trials, with 20 self-trials and 20 other-trials randomly interleaved (120 trials in total). On each block, subjects were paired with a new player. All subjects faced the same three players. We created the three players using a pilot dataset of random dot motion decisions obtained from previous subjects who had visited the lab. We identified three pilot subjects who on average achieved about 60% (low ability), 75% (medium ability) and 90% (high ability) choice accuracy and then sub-selected 20 trials which covered a range of coherences (.001-.5), had an even split of leftwards and right motion and satisfied the above requirements for choice accuracy. On each block, the same set of motion stimuli were shown on self-trials and other-trials but in a randomised order. The order of the three players across blocks was randomised across subjects.

FMRI

Procedure

The scan session consisted of 5 scan runs. First, we acquired structural images. Second, subjects performed the self-other task over 3 runs (3 x 40 = 120 trials; 252 volumes per run). Lastly, we ran a localiser scan in which subjects viewed alternating displays (12 s) of static and coherently moving dots (2 x 10 = 20 displays; 215 volumes).

Acquisition

MRI data were acquired on a 3T Siemens Allegra scanner at New York University's Center for Brain Imaging. T1-weighted structural images were acquired using a 3D MPRAGE sequence: 1 x 1 x 1 mm resolution voxels; 176 sagittal slices. BOLD T2*-weighted functional images were acquired using a Siemens epi2d BOLD sequence: 3 x 3 x 3 mm resolution voxels; 42 transverse slices, 64 x 64 matrix; TR = 2.24s; TE = 30ms; slice tilt = -30 degrees T > C; slice thickness = 3 mm; interleaved slice acquisition). Local field maps were recorded for distortion correction of the acquired EPI data.

Pre-processing

MRI data were pre-processed using SPM12 (Wellcome Trust, London). The first 5 volumes of each functional run were discarded to allow for T1 equilibration. Functional images were slice-time corrected, re-aligned and un-warped using the collected field maps[48]. Structural T1-weighted images were co-registered to the mean functional image of each subject using the iterative mutual-information algorithm. Each subject's structural image was segmented into grey matter, white matter and cerebral spinal fluid using a nonlinear deformation field to map it onto a template tissue probability map[49]. These deformations were applied to both structural and functional images to create new images spatially normalised to the Montreal Neurological Institute (MNI) space and interpolated to 2 x 2 x 2 mm voxels. Normalized images were spatially smoothed using a Gaussian kernel with full-width half-maximum of 8 mm. Motion correction parameters estimated from the re-alignment procedure and their first temporal derivatives (12 regressors in total) were included as confounds in the first-level analysis for each subject.

Quantification and statistical analysis

Trial exclusion

We excluded trials when a decision was not made during the choice window or when a gamble was not made during the gamble window.

Behavioural analysis

We used multiple logistic regression to predict subjects' PDWs (**Figure S1**). We z-transformed all continuous predictors. We performed separate regressions for each subject. We tested group-level significance by comparing coefficients pooled across subjects to zero ($P < .05$, one-sample *t*-test).

Computational models of confidence

On each trial, a subject is asked to make a post-decision wager (PDW) about either their own choice or the choice of another player. On self-trials, the subject actively makes a choice about the motion stimulus. On other-trials, the subject is presented with the motion stimulus but not the other's choice. The subject has to decide between a risky and a safe option, by combining

the option values with an estimate of the probability that the choice is correct. We first consider models of PDWs on self-trials, before turning to other-trials. The models are all grounded in Bayesian decision theory[7,50].

Sensory sample

On each trial, a subject, s, receives a sensory sample, x, randomly sampled from a Gaussian distribution, $x \in N(k\theta_m, \sigma_s)$, where $m \in M = \{1,2,\dots,n\}$, $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ indicates the motion coherence, $k \in K = \{-1,1\}$ indicates the motion direction (-1: left; 1: right), n is the number of levels of coherence and σ_s indicates the subject's level of sensory noise.

Self-trials

The subject represents the task as comprised of a set of states (k, m), corresponding to the set of possible motion stimuli defined by direction and coherence, $k\theta_m$. The subject computes a belief distribution over states given their sensory sample, x. The probability of each state conditional on x is given by:

$$P(k,m|x) = \frac{P(x|m,k)P(m,k)}{\sum_{k'}\sum_{i=1}^{n} P(x|m_i,k')\,P(m_i,k')} = \frac{P(x|m,k)}{\sum_{k'}\sum_{i=1}^{n} P(x|m_i,k')} \qquad \text{Eq. 1}$$

Note that we relied on the fact that all states are equally probable by design (with probability $1/2n$). The likelihood of x conditional on a state is given by:

$$P(x|k,m) = \varphi(x; k\theta_m, \sigma_s) \qquad \text{Eq. 2}$$

where φ is the normal probability density function evaluated at x:

$$\varphi(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \text{Eq. 3}$$

The subject uses the belief distribution over states to compute the probability that the motion direction is leftwards ($k = -1$) and rightwards ($k = 1$):

$$P(k=-1|x) = \frac{\sum_{m\in M} P(x|m,k=-1)}{\sum_{k'}\sum_{m\in M} P(x|m,k')}$$

$$P(k=1|x) = \frac{\sum_{m\in M} P(x|m,k=1)}{\sum_{k'}\sum_{m\in M} P(x|m,k')} \qquad \textit{Eq. 4}$$

The subject makes their decision, r, by comparing the posteriors over motion direction:

$$r(x) = \begin{cases} -1 & \text{if } P(k=-1|x) > P(k=1|x) \\ 1 & \text{if } P(k=-1|x) < P(k=1|x) \end{cases} \qquad \textit{Eq. 5}$$

Finally, the probability that the choice is correct can be computed directly using the posteriors over motion direction relative to the chosen action:

$$P(\text{correct}|x) = P(k=r|x) = \begin{cases} P(k=-1|x) & \text{if r} = -1 \\ P(k=1|x) & \text{if r} = 1 \end{cases} \qquad \textit{Eq. 6}$$

Other-trials

Subjects were paired with three players of varying ability, $j \in J = \{1,2,3\}$. We consider below different models of how subjects compute confidence in player j's choices and how subjects learn about player j's ability.

Self-projection (S-model)

A simple way to compute confidence in another player's choice is to apply the same model to the other player as one applies to oneself. Under this model, the subject simulates the other's choices and confidence as if they were the other player.

The other player's posteriors over motion direction are calculated as:

$$P(k=-1|x,j) = P(k=-1|x,s)$$
$$P(k=1|x,j) = P(k=1|x,s) \qquad \text{Eq. 7}$$

The other player's choice is calculated as:

$$r(x) = \begin{cases} -1 & \text{if } P(k=-1|x,j) > P(k=1|x,j) \\ 1 & \text{if } P(k=-1|x,j) < P(k=1|x,j) \end{cases} \qquad \text{Eq. 8}$$

The subject's confidence in the other player's choice is calculated as:

$$P(\text{correct}|x,j) = P(k=r|x,j) = \begin{cases} P(k=-1|x,j) & \text{if r} = -1 \\ P(k=1|x,j) & \text{if r} = -1 \end{cases} \qquad \text{Eq. 9}$$

Note that the S-model does not learn about the other player's ability.

Simple value learner (Q-model)

An alternative strategy for estimating confidence in another player's choice is to keep a running estimate of the value (correctness) of their choices, $Q(j)$. This can be done compactly using a Rescorla-Wagner update rule:

$$Q(j)_{t+1} \coloneqq Q(j)_t + \alpha\delta_t \qquad \text{Eq. 10}$$

where t indicates the trial, α is a learning rate and δ_t is the difference between the observed outcome o_t (correct = 1; incorrect = 0) and the current estimate, $Q(j)_t$:

$$\delta_t = o_t - Q(j)_t \qquad \text{Eq. 11}$$

The subject's confidence in the other player's choice is then calculated as:

$$P_j(\text{correct}|j) = Q(j) \qquad \text{Eq. 12}$$

We set $Q(j)_{t=0} = .75$ for each player j. Note that, unlike the S-model, the Q-model is blind to the effect of coherence on the other player's chance of success.

Sophisticated value learner (Q*-model)

An alternative to the Q-model is to learn about the value of another player's choices for each coherence, $Q(m,j)$. Another version of this model would learn about $k\theta_m$ but here we ignore k as we did not manipulate the other players' preference for one or other motion direction (i.e., choice bias). In this model, the prediction about the other player's success is obtained by weighting the value of coherence by a belief distribution over coherence given the sensory sample:

$$P_j(\text{correct}|x) = \sum_{m \in M} Q(m,j)P(m|x) \qquad \text{Eq. 13}$$

Where $P(m|x)$ is calculated as:

$$P(m|x) = \sum_{k=\pm 1} P(m,k|x) \qquad \text{Eq. 14}$$

Similarly, the value of the other player's choice for each coherence is updated by weighting a prediction error by the belief distribution over coherence given the sensory sample:

$$Q(m,j)_{t+1} \coloneqq Q(m,j)_t + \alpha\delta_t P(m|x) \qquad \textit{Eq. 15}$$

Where the prediction error is calculated as:

$$\delta_t = o_t - P_j(\text{correct}|x) \qquad \textit{Eq. 16}$$

We set $Q(m,j)_{t=0} = .75$ for each player j. Note that, unlike the Q-model, the Q*-model will take into account the effect of coherence on the other player's chance of success. However, the Q*-model will in certain situations generalise experience in a counterintuitive manner: for example, after experiencing a correct decision at low coherence, the Q*-model may learn that a low-coherence stimulus is more valuable (i.e., more likely to lead to correct decisions) than a high-coherence stimulus.

Theory of mind (ToM-model)

A more sophisticated strategy is to maintain a running estimate of another player's sensory noise – that is, a model of their sensory system – and use this estimate to compute confidence in their choices. In the absence of shared error variance between the subject's and the other player's sensory sample (i.e., no noise correlation), the subject's prediction about the other player's success should take into account (1) the posterior probability that the world is in a given state, (k,m), given their own sensory sample, x, and (2) the other player's state-dependent performance (i.e., their expected accuracy in each state) given their estimated sensory noise, σ_j. The prediction about the other player's success is thus given by:

$$P_j(\text{correct}|x) = \sum_k \sum_{m\in M} P(m,k|x) P_j(\text{correct}|m,k) \qquad \textit{Eq. 17}$$

where P_j denotes the probability that the other agent j makes a correct response. The first term, the belief distribution over states, is calculated using σ_s, and the second term, the state-dependent performance of the other player, is calculated using σ_j:

$$P_j(\text{correct}|m,k) = \begin{cases} \phi\big(0; k\theta_m, \sigma_j\big) & \text{if } k\theta_m < 0 \\ 1 - \phi\big(0; k\theta_m, \sigma_j\big) & \text{if } k\theta_m > 0 \end{cases} \qquad \textit{Eq. 18}$$

where ϕ is the cumulative normal distribution:

$$\phi(y;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{y} e^{\frac{-(y-\mu)^2}{2\sigma^2}} dy \qquad \textit{Eq. 19}$$

The estimate of the other player's sensory noise is updated using an approximation to the Rescorla-Wagner rule:

$$\sigma_{j,t+1} = \sigma_{j,t} + \alpha\delta_t / D(x, \sigma_{j,t}) \qquad \textit{Eq. 20}$$

Where δ is a prediction error as in the Q*-model and $D(x,\sigma_j)$ is the following derivative:

$$D\big(x,\sigma_j\big) = \frac{dP_j(correct|x)}{d\sigma_j} \qquad \textit{Eq. 21}$$

In short, we want to define an update to the estimate of the other player's sensory noise, η_t, such that the prediction about the other player's success, $P_j(\text{correct}|x, \sigma_{j,t} + \eta_t)$, will be equal to $P_j(\text{correct}|x, \sigma_{j,t}) + \alpha\delta_t$:

$$P_j(\text{correct}|x, \sigma_{j,t} + \eta_t) = P_j(\text{correct}|x, \sigma_{j,t}) + \alpha\delta_t \qquad \text{Eq. 22}$$

We can approximate the left-hand term around $\sigma_{j,t}$ linearly. So, we obtain:

$$P_j(\text{correct}|x, \sigma_{j,t}) + D(x, \sigma_{j,t})\eta_t = P_j(\text{correct}|x, \sigma_{j,t}) + \alpha\delta_t \qquad \text{Eq. 23}$$

Or

$$\sigma_{j,t+1} = \sigma_{j,t} + \alpha\delta_t / D(x, \sigma_{j,t}) \qquad \text{Eq. 24}$$

The derivative of the prediction with respect to the parameter is calculated as:

$$\begin{aligned} D(x, \sigma_j) &= \sum\nolimits_{m \in M, k=\pm 1} \left(\frac{dP(m,k|x)}{d\sigma_j} P_j[\text{correct}|m,k] + P(m,k|x) \frac{dP_j[\text{correct}|m,k]}{d\sigma_j} \right) \\ &= \sum\nolimits_{m \in M, k=\pm 1} \left(P(m,k|x) \frac{dP_j[\text{correct}|m,k]}{d\sigma_j} \right) \end{aligned} \qquad \text{Eq. 25}$$

Because $P(m,k|x)$ is not a function of σ_j, its derivative is 0.

The term $\frac{dP_j[\text{correct}|m,k]}{d\sigma_j}$ is calculated as:

$$\frac{dP_j[\text{correct}|m,k]}{d\sigma_j} = -\frac{\theta_m}{\sigma_j^2\sqrt{2\pi}} \, e^{\frac{-(\theta_m)^2}{2\sigma_j^2}} = -\frac{\theta_m}{\sigma_j} \, \varphi(0; \theta_m, \sigma_j) \qquad \text{Eq. 26}$$

We thus have:

$$D(x, \sigma_j) = -\frac{1}{\sigma_j} \sum\nolimits_{m \in M, k=\pm 1} [P(m,k|x)\theta_m \, \varphi(0; \theta_m, \sigma_j)] \qquad \text{Eq. 27}$$

The linear approximation to the Rescorla-Wagner rule may result in numerical overflow in a small number of cases (e.g., when σ_j is extremely low and x is extremely high). To avoid this issue, we implemented a proposal mechanism under which we iteratively adjusted the update term, $\alpha\delta_t / D(x, \sigma_{j,t})$, until $|P_j(\text{correct}|x, \sigma_{j,t} + \eta_t) - (P_j(\text{correct}|x, \sigma_{j,t}) + \alpha\delta_t)| < .001$.

Sophisticated theory of mind (ToM*-model)

Finally, we considered a model in which subjects assume that there is shared error variance between their own and the other player's sensory sample. This assumption is often plausible, such as when two people are viewing the same noisy instantiation of a sensory stimulus. Under shared error variance, the subject's sensory sample, x_s, and the other player's sensory sample, x_j, are drawn from a multivariate Gaussian distribution:

$$\begin{bmatrix} x_s \\ x_j \end{bmatrix} | k, m \in N\left(\begin{pmatrix} k\theta_m \\ k\theta_m \end{pmatrix}, \mathbf{\Sigma} \right) \qquad \text{Eq. 28}$$

where the mean vector is given by the motion stimulus, $k\theta$, and the covariance matrix is given by:

$$\mathbf{\Sigma} \in \begin{bmatrix} \sigma_s^2 & \rho\sigma_s\sigma_j \\ \rho\sigma_s\sigma_j & \sigma_j^2 \end{bmatrix} \qquad \text{Eq. 29}$$

Here ρ indicates the subject's belief about the degree of shared error variance and is assumed to be fixed across partners.

To unpack the ToM*-model, we reframe the computation of the subject's prediction about the other player's success as:

$$P_j(\text{correct}|x_s) = P(r = k = -1|x_s) + P(r = k = 1|x_s) \qquad \text{Eq. 30}$$

The definition of conditional probability tells us that the joint probability of motion direction, k, and the other player's choice, r, given the subject's sensory sample, x_s, is:

$$P(r, k|x_s) = \frac{P(r, k, x_s)}{P(x_s)} \qquad \text{Eq. 31}$$

The numerator is calculated by summing over coherence:

$$P(r, k, x_s) = \sum_m P(r, k, x_s, m) \qquad \text{Eq. 32}$$

and using the chain rule:

$$P(r, k, x_s, m) = P(k)P(m|k)P(x_s|k, m)P(r|k, m, x_s) \qquad \text{Eq. 33}$$

where $P(k)$ is the probability of a motion direction, $P(m|k)$ is the probability of a coherence given a motion direction, $P(x_s|k, m)$ is the likelihood of a sensory sample given a motion direction and a coherence (calculated as on self-trials) and $P(r|k, m, x_s)$ is the probability of the other's choice given a motion direction, a coherence and a sensory sample.

The last term is given by:

$$P(r|k, m, x_s) = \int P(r|\, x_j)P\big(x_j\big|x_s, k, m\big)\, dx_j \qquad \text{Eq. 34}$$

In other words, the subject needs to infer the sensory sample x_j underlying the other's choice from their own sensory sample x_s and then marginalise over possible x_j's in order to estimate the probability of the other player's action.

This expression is a cumulative density function of the conditional density of a multivariate Gaussian distribution, which itself is a univariate Gaussian distribution with the following mean and standard deviation:

$$P\big(x_j\big|x_s, k, m\big) \sim N(\mu_{x_j|x_s}(k, m), \sigma_{x_j|x_s}) \qquad \text{Eq. 35}$$

where:

$$\mu_{x_j|x_s}(k, m) = k\theta_m + \frac{\sigma_j}{\sigma_s}\rho(x_s - k\theta_m)$$
$$\sigma_{x_j|x_s} = \sqrt{(1 - \rho^2)\sigma_j^2} \qquad \text{Eq. 36}$$

We can thus calculate the probability of the other player's choice given a motion direction, a coherence and a sensory sample as:

$$P(r|k, m, x_s) = \begin{cases} \phi\left(0; \mu_{x_j|x_s}(k, m), \sigma_{x_j|x_s}\right) & \text{if } r = -1 \\ 1 - \phi\left(0;\ \mu_{x_j|x_s}(k, m), \sigma_{x_j|x_s}\right) & \text{if } r = 1 \end{cases} \qquad \text{Eq. 37}$$

where ϕ is the cumulative normal distribution evaluated at 0 as described above.

By means of these equations, we can calculate the 2 × 2 set of joint probabilities over motion direction ($k = -1$ or $k = 1$) and the other player's choice ($r = -1$ or $r = 1$) conditional on the subject's sensory sample and in turn calculate the subject's prediction about the other player's success – that is, the confidence in the other player's decision.

The estimate of the other player's sensory noise is updated as in the ToM-model:

$$\sigma_{j,t+1} = \sigma_{j,t} + \alpha\delta_t / D(x, \sigma_{j,t}) \qquad \textit{Eq. 38}$$

The derivative of the prediction with respect to the parameter is calculated as:

$$D\big(x, \sigma_{j,t}\big) = \frac{dP_j(correct|x_s)}{d\sigma_j} = \frac{d\big(P(r = k = -1|x_s) + P(r = k = 1|x_s)\big)}{d\sigma_j}$$

$$= \frac{dP(r = k = -1|x_s)}{d\sigma_j} + \frac{dP(r = k = 1|x_s)}{d\sigma_j}$$

$$= \frac{d(\sum_m [\frac{P(k=-1)P(m|k=-1)P(x_s|k=-1,m)P(r=-1|k=-1,m,x_s)}{P(x_s)}])}{d\sigma_j} + \frac{d(\sum_m [\frac{P(k=1)P(m|k=1)P(x_s|k=1,m)P(r=1|k=1,m,x_s)}{P(x_s)}])}{d\sigma_j} \qquad \textit{Eq. 39}$$

$$= \frac{P(k=-1)}{P(x_s)} \sum_m \left(P(m|k=-1)P(x_s|k=-1,m) \frac{dP(r=-1|k=-1,m,x_s)}{d\sigma_j} \right) + \frac{P(k=1)}{P(x_s)} \sum_m \left((m|k=1)P(x_s|k=1,m) \frac{dP(r=1|k=1,m,x_s)}{d\sigma_j} \right)$$

Where

$$\frac{dP(r=-1|k=-1,m,x_s)}{d\sigma_j} = \frac{-\theta_m e^{\frac{-\left(\theta_m - \frac{\sigma_j}{\sigma_s}\rho(x_s+\theta_m)\right)^2}{2(1-\rho^2){\sigma_j}^2}}}{\sqrt{2\pi(1-\rho^2){\sigma_j}^2}}$$

$$\frac{dP(r=1|k=1,m,x_s)}{d\sigma_j} = \frac{-\theta_m e^{\frac{-(\theta_m + \frac{\sigma_j}{\sigma_s}\rho(x_s-\theta_m))^2}{2(1-\rho^2){\sigma_j}^2}}}{\sqrt{2\pi(1-\rho^2){\sigma_j}^2}} \qquad \textit{Eq. 40}$$

Post-decision wagering

In all models, confidence is transformed into PDWs by computing the expected values of the risky and the safe options:

$$EV_{risky} = P(\text{correct}|x,a)V_{risky} - [1 - P(\text{correct}|x,a)]V_{risky}$$
$$EV_{safe} = V_{safe} \qquad \textit{Eq. 41}$$

where a indicates the agent (subject or other) and V indicates the option value.

Gamble probability

To account for individual variance in preferences for risk and sensitivity to changes in expected value, we modelled the probability of selecting the risky option using a softmax function:

$$P(risky) = \frac{1}{1 + e^{-[\beta_0 + \beta_1 (EV_{risky} - EV_{safe})]}} \qquad \text{Eq. 42}$$

Model space

We considered versions of the above models where the softmax bias, β_0, and/or temperature, β_1, could vary between self- and other-trials and where the representation of the stimulus space, M, was rescaled according to a power law (using a factor p). In brief, we rescaled each state as $m' = m^p$ and normalised the resulting state space, M', using the minimum and maximum values of the original stimulus space, M. The rescaling allowed for a higher density of motion stimuli at low ($p > 1$) or high ($p < 1$) stimulus intensities (where $p = 1$ preserves the original stimulus space). There were thus 40 models: 5 (S, Q, Q*, ToM or ToM*) x 2 (same or different bias, β_0, for self and other) × 2 (same or different temperature, β_1, for self and other) × 2 (linearly-spaced or rescaled stimulus space).

Model fitting and comparison

We used variational Bayesian inference as implemented in STAN[51] to fit the models to the behavioural data. We used a hierarchical fitting procedure where group-level parameters constrain subject-level parameters. We summarise the parameter space in **Table S1**. We fitted each model using the following specifications: maximum number of iterations = 4,000; number of samples for Monte Carlo estimation of objective function = 200; number of iterations between evaluation of objective function = 100; convergence tolerance on the relative norm of the objective = .0001. In order to generate trial-by-trial predictions about each subjects' behaviour, and estimates of their latent model variables, we drew 500 samples from the posterior distributions over fitted parameters (using the 'generated quantities' block in STAN), calculated the trial-by-trial values for each set of samples and then averaged across these 500 iterations. To assess the ability of each model to account for a subject's data, we computed the log-likelihood of their trial-by-trial PDWs under each draw of parameters and applied leave-one-out cross-validation using Pareto-smoothed importance sampling (PSIS-LOO) and the widely applicable information criterion (WAIC) – both methods for estimating pointwise out-of-sample prediction accuracy under a Bayesian model[52]. We performed these steps 4 times for each model using different random seeds and averaged outputs across the 4 runs. The results of the model comparison are shown in **Figure S2**.

FMRI

Whole-brain general linear models

The whole-brain analysis shown in **Figure 4** was based on a single event-related general linear model (GLM1). We separated trials into self-trials and other-trials and specified separate 'condition' regressors for the decision phase (i.e., boxcar from RDK onset to 3 s after decision onset) and the gamble phase (i.e., boxcar from gamble onset until feedback offset) for each trial type (yielding 4 condition regressors in total). We included motion parameters as 'nuisance' regressors. Regressors were convolved with a canonical hemodynamic response function. Regressors were modelled separately for each scan run and constants were included to account for between-run differences in mean activation and scanner drifts. A high-pass filter

(128s cut-off) was applied to remove low-frequency drifts. Group-level significance was assessed by applying one-sample *t*-tests against zero to the first-level contrast images. We report clusters significant at $p < .05$, FWE-corrected for multiple comparisons, with a cluster extent of 10 voxels or more and a cluster-defining threshold of $p < .001$, uncorrected. Numerical simulations and tests of empirical data collected under the null hypothesis show that this combination of cluster-defining threshold and random field theory produces appropriate control of false positives[53].

Regions of interest

ROI masks for MT+ were created using the localiser scan: we created a group mask using the second-level contrast between dynamic and static motion, and then, for each subject, created a MT+ mask (8-mm sphere) around their peak activity inside the group mask. ROI masks for LIP, TPJ and dmPFC were created using published connectivity-based parcellation atlases: LIP was defined as the union of areas SPLD and SPLE in the atlas developed by Mars et al.[30]; TPJ was defined as area TPJp in the atlas developed by Mars et al.[54]; and dmPFC was defined as area 9 in the atlas developed by Neubert et al.[55]. All ROI masks were bilateral.

Psychophysiological interaction analysis

To assess changes in connectivity between visual and social ROIs as a function of trial type, we carried out a psychophysiological interaction (PPI) analysis. PPI is a measure of context-dependent connectivity, explaining the regional activity of other brain regions in terms of the interaction between responses in a seed region and a cognitive or sensory process. We used GLM1 as basis for PPI analysis. We carried out the PPI analysis using the Generalised PPI toolbox for SPM (gPPI; http://www.nitrc.org/projects/gppi/). gPPI creates a new GLM in which the deconvolved activity of the seed region is assigned to separate regressors dependent on the status of the original psychological variable (self versus other during the decision phase), and re-convolved with the hemodynamic response function. We extracted average contrast estimates within an ROI for each subject and then used these estimates for group-level testing.

Single-trial activity time courses

We transformed each ROI mask from MNI to native space and extracted preprocessed BOLD time courses as the average of voxels within the mask. For each scan run, we regressed out variation due to head motion as specified above, applied a high-pass filter (128 s cut-off) to remove low-frequency drifts, and oversampled the BOLD time course by a factor of ~23 (time resolution of .144 s). For each trial, we extracted activity estimates in a 14 s window time-locked to an event of interest.

Single-trial activity estimates

We estimated single-trial ROI activity estimates as a beta time series. This was achieved by means of an event-related model with a separate regressor for each trial. Regressors were boxcars time-locked to the onset of the motion display and spanning until choice[56]. Each regressor was convolved with a canonical hemodynamic response function. We included motion parameters as 'nuisance' regressors. Regressors were modelled separately for each scan run and constants were included to account for between-run differences in mean activation and scanner drifts. A high-pass filter (128 s cut-off) was applied to remove low-frequency drifts. One important consideration in using single-trial activity estimates is that a

beta for a given trial can be strongly affected by acquisition artefacts that occur together with that trial (e.g., scanner pulse artefacts). Therefore, for each subject, we computed the grand-mean beta estimate across both voxels and trials and excluded any trial whose mean beta estimate across voxels was 3 SDs below or above this grand mean[57]. Finally, we used the ROI masks to extract ROI single-trial activity estimates.

DATA AND SOFTWARE AVAILABILITY

[*To be completed upon acceptance*] Anonymized behavioural data and code supporting our main analyses are available on GitHub (link). Unthresholded group-level statistical maps are available on NeuroVault (link).

REFERENCES

1. Fleming, S. M., Van Der Putten, E. J. & Daw, N. D. Neural mediators of changes of mind about perceptual decisions. *Nat. Neurosci.* **21**, 617–624 (2018).
2. Van Den Berg, R. *et al.* A common mechanism underlies changes of mind about decisions and confidence. *Elife* **5**, 1–21 (2016).
3. Ma, W. J. & Jazayeri, M. Neural Coding of Uncertainty and Probability. *Annu. Rev. Neurosci.* **37**, 205–220 (2014).
4. Pouget, A., Drugowitsch, J. & Kepecs, A. Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* **19**, 366–374 (2016).
5. Meyniel, F., Sigman, M. & Mainen, Z. F. Confidence as Bayesian probability: from neural origins to behavior. *Neuron* **88**, 78–92 (2015).
6. Aitchison, L., Bang, D., Bahrami, B. & Latham, P. E. Doubly Bayesian analysis of confidence in perceptual decision-making. *PLoS Comput. Biol.* **11**, (2015).
7. Fleming, S. M. & Daw, N. D. Self-evaluation of decision-making: A general bayesian framework for metacognitive computation. *Psychol. Rev.* **124**, 91–114 (2017).
8. Bang, D. & Fleming, S. M. Distinct encoding of decision confidence in human medial prefrontal cortex. *Proc. Natl. Acad. Sci.* **115**, 6082–6087 (2018).
9. Gherman, S. & Philiastides, M. G. Human VMPFC encodes early signatures of confidence in perceptual decisions. *bioRxiv* 224337 (2017) doi:10.1101/224337.
10. Pereira, M. *et al.* Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 8382–8390 (2020).
11. Fleming, S. M., Huijgen, J. & Dolan, R. J. Prefrontal contributions to metacognition in perceptual decision making. *J. Neurosci.* **32**, 6117–6125 (2012).
12. Bang, D., Ershadmanesh, S., Nili, H. & Fleming, S. M. Private–public mappings in human prefrontal cortex. *Elife* **9**, e56477 (2020).
13. Rouault, M. & Fleming, S. M. Formation of global self-beliefs in the human brain. *Pnas* **in press**, 1–9 (2020).
14. Johnson, D. D. P. & Fowler, J. H. The evolution of overconfidence. *Nature* **477**, 317–320 (2011).
15. Mahmoodi, A. *et al.* Equality bias impairs collective decision-making across cultures. *Proc. Natl. Acad. Sci.* **112**, 201421692 (2015).
16. Yaniv, I. Receiving other people's advice: Influence and benefit. *Organ. Behav. Hum. Decis. Process.* **93**, 1–13 (2004).
17. Bang, D. *et al.* Does interaction matter? Testing whether a confidence heuristic can replace interaction in collective decision-making. *Conscious. Cogn.* **26**, 13–23 (2014).
18. Hertz, U. *et al.* Neural computations underpinning the strategic management of influence in advice giving. *Nat. Commun.* **8**, 1–12 (2017).
19. Mahmoodi, A., Bang, D., Ahmadabadi, M. N. & Bahrami, B. Learning to make collective decisions: the impact of confidence escalation. *PLoS One* **8**, (2013).
20. Bang, D. *et al.* Confidence matching in group decision-making. *Nat. Hum. Behav.* **1**, (2017).
21. Wittmann, M. K. *et al.* Self-other mergence in the frontal cortex during cooperation and competition. *Neuron* **91**, 482–493 (2016).
22. Boorman, E. D., O'Doherty, J. P., Adolphs, R. & Rangel, A. The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron* **80**, 1558–1571 (2013).
23. Shadlen, M. N. & Kiani, R. Decision making as a window on cognition. *Neuron* **80**, 791–806 (2013).
24. Saxe, R. Uniquely human social cognition. *Curr. Opin. Neurobiol.* **16**, 235–239 (2006).
25. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
26. Frith, C. D. & Frith, U. Mechanisms of social cognition. *Annu. Rev. Psychol.* **63**, 287–313 (2012).
27. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).
28. Newsome, W. T., Britten, K. H. & Movshon, J. A. Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989).
29. Lak, A., Nomoto, K., Keramati, M., Sakagami, M. & Kepecs, A. Midbrain dopamine neurons signal

belief in choice accuracy during a perceptual decision. *Curr. Biol.* **27**, 821–832 (2017).
30. Mars, R. B. *et al.* Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and Macaque resting-state functional connectivity. *J. Neurosci.* **31**, 4087–4100 (2011).
31. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
32. Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
33. Hebart, M. N., Donner, T. H. & Haynes, J. D. Human visual and parietal cortex encode visual choices independent of motor plans. *Neuroimage* **63**, 1393–1403 (2012).
34. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* **86**, 1916–1936 (2001).
35. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–9 (2008).
36. Braddick, O. J. *et al.* Brain areas sensitive to coherent visual motion. *Perception* **30**, 61–72 (2001).
37. Beck, J. M. *et al.* Probabilistic population codes for Bayesian decision making. *Neuron* **60**, 1142–1152 (2008).
38. Adolphs, R. The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* **60**, 693–716 (2009).
39. Carruthers, P. How we know our own minds: The relationship between mindreading and metacognition. *Behav. Brain Sci.* **32**, 121–182 (2009).
40. Saxe, R. Against simulation: The argument from error. *Trends Cogn. Sci.* **9**, 174–179 (2005).
41. Mitchell, J. P. Inferences about mental states. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1309–1316 (2009).
42. Buckner, R. L. & Carroll, D. C. Self-projection and the brain. *Trends Cogn. Sci.* **11**, 49–57 (2007).
43. Barron, H. C., Garvert, M. M. & Behrens, T. E. J. Repetition suppression: a means to index neural representations using BOLD? *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150355 (2016).
44. Lockwood, P. L., Apps, M. A. J. & Chang, S. W. C. Is There a 'Social' Brain? Implementations and Algorithms. *Trends Cogn. Sci.* **24**, 802–813 (2020).
45. Mazor, M., Friston, K. J. & Fleming, S. M. Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. *Elife* **9**, 1–34 (2020).
46. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
47. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision* vol. 10 437–442 (1997).
48. Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R. & Friston, K. Modeling geometric deformations in EPI time series. *Neuroimage* **13**, 903–919 (2001).
49. Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
50. Green, D. & Swets, J. *Signal Detection Theory and Psychophysics*. (Wiley, 1966).
51. Kucukelbir, A., Ranganath, R., Gelman, A. & Blei, D. M. Automatic variational inference in Stan. in *Advances in Neural Information Processing Systems 28* 568–576 (2015).
52. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2017).
53. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci.* **113**, 7900–7905 (2016).
54. Mars, R. B. *et al.* Connectivity-based subdivisions of the human right 'temporoparietal junction area': evidence for different areas participating in different cortical networks. *Cereb. Cortex* **22**, 1894–1903 (2012).
55. Neubert, F.-X., Mars, R. B., Sallet, J. & Rushworth, M. F. S. Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc. Natl. Acad. Sci.* **112**, E2695–E2704 (2015).
56. Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P. & Hirsch, J. Detection of time-varying signals in event-related fMRI designs. *Neuroimage* **43**, 509–520 (2008).

57. Atlas, L. Y., Bolger, N., Lindquist, M. A. & Wager, T. D. Brain mediators of predictive cue effects on perceived pain. *J. Neurosci.* **30**, 12964–12977 (2010).

Supplementary Materials

Neurocomputational mechanisms of confidence in others

Dan Bang*, Rani Moran, Nathaniel Daw and Stephen M. Fleming*

*Correspondence: danbang.db@gmail.com and stephen.fleming@ucl.ac.uk

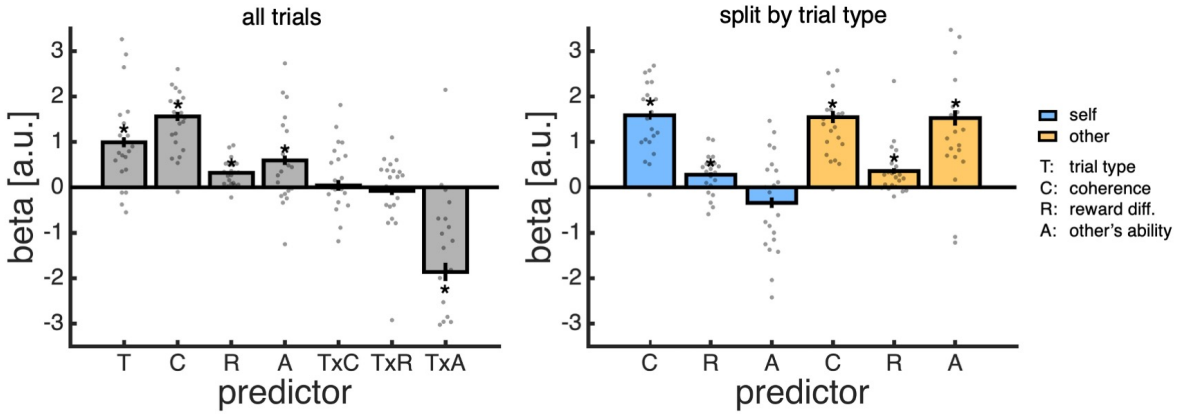

Figure S1. Multiple logistic regression analysis of post-decision wagers. **(A)** We first predicted gamble choices (i.e., selecting the risky option over the safe option) on all trials using trial type (T; other = -.5; self = .5), coherence (C; z-score), reward difference (R; z-score), other's ability (A; low = -.5, medium = 0, high = .5) and the interactions between trial type and the three other terms (TxC, TxR and TxA). **(B)** We then predicted gamble choices separately for self- (yellow) and other-trials (blue) using coherence, reward difference and other's ability. **(A-B)** We performed separate regressions for each subject. We tested group-level significance (asterisk) by comparing coefficients pooled across subjects to zero (one-sample *t*-test). Data are represented as group mean ± SEM, with individual subjects overlaid as dots.

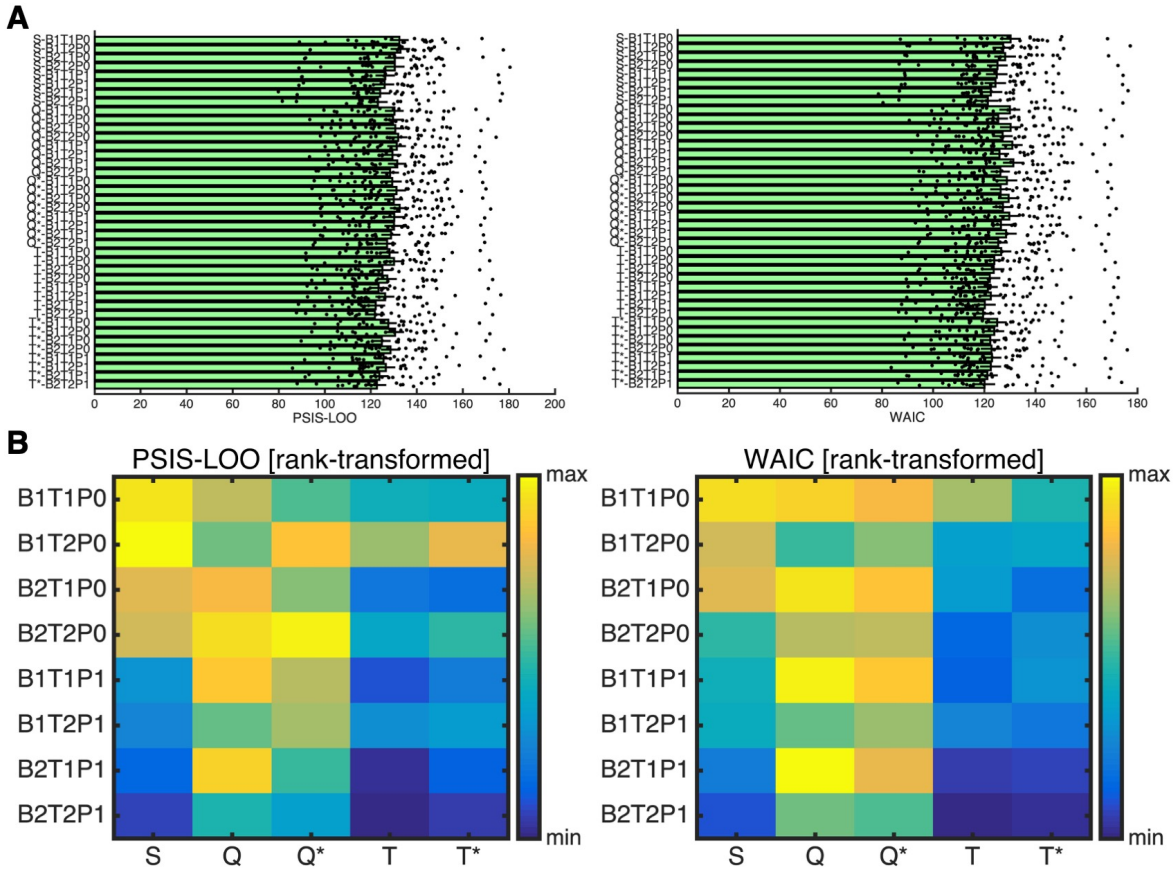

Figure S2. Model comparison based on leave-one-out cross-validation using (*left*) Pareto-smoothed importance sampling (PSIS-LOO) or (*right*) the widely applicable information criterion (WAIC). Here smaller values indicate better fit. **(A)** Metrics for each model represented as group mean ± SEM, with individual subjects overlaid as dots. **(B)** Rank-transformed average model metrics. **(A-B)** Model names are abbreviated as follows: S, S-model; Q, Q-model; Q*, Q*-model; T, ToM-model; T*, ToM*-model. Model versions are abbreviated as follows: B, softmax bias (1: same bias for self and other; 2: different bias for self and other); T, softmax temperature (1: same temperature for self and other; 2: different temperature for self and other); P, rescaling of stimulus space (0: linear spacing; 1: re-scaled according to fitted factor). See **Table S1** for overview of model parameters.

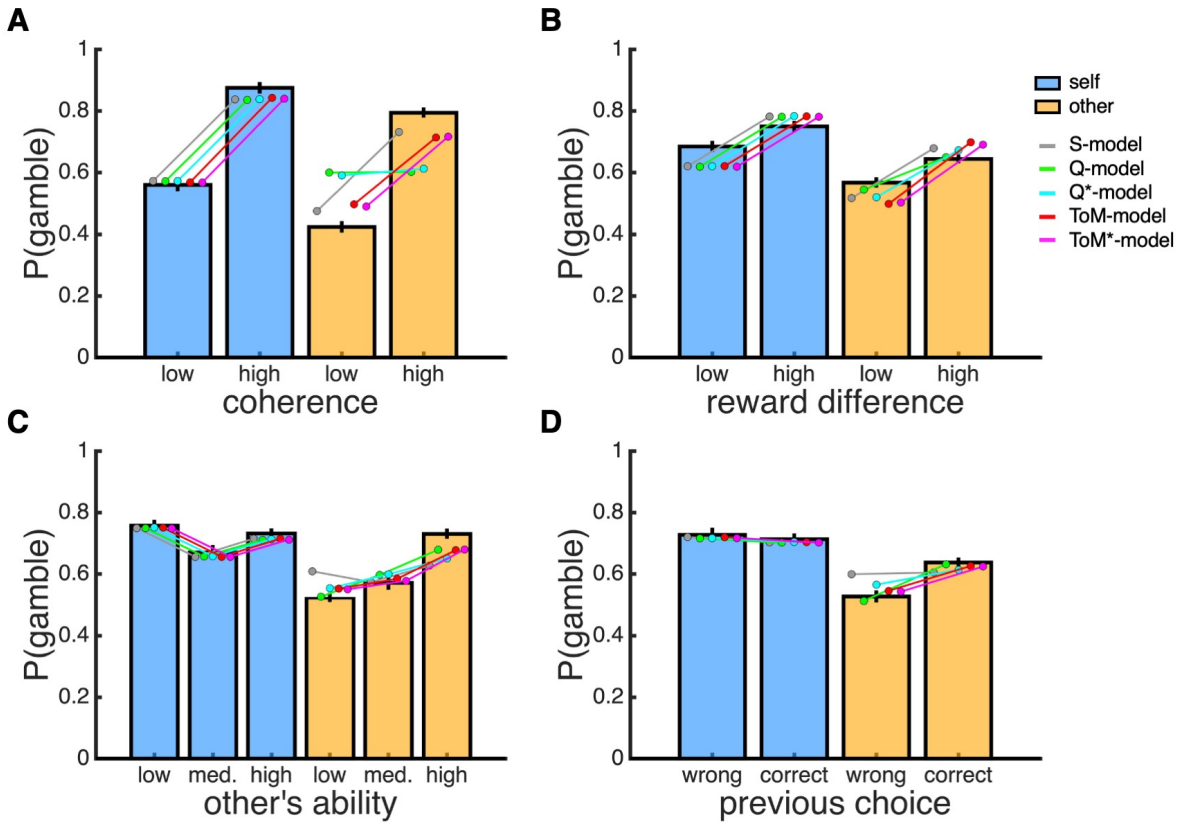

Figure S3. Behavioural results and model predictions. In each panel, data are split into self-trials (blue) and other-trials (yellow). **(A**) Probability gamble as a function of coherence (median split). **(B)** Probability gamble as a function of the reward difference between the risky and the safe option (median split). **(C)** Probability gamble as a function of the ability of the other player. **(D)** Probability gamble as a function of choice accuracy on the previous trial of the same type. **(A-D)** Bar charts are empirical data. Dots show data simulated under the best-fitting version of each candidate model (**Figure S2**). Models are denoted by colour and dots displayed in the following order: S-model, Q-model, Q*-model, ToM-model and ToM*-model. While all models predict the observed pattern of results on self-trials, only the ToM-model and the ToM*-model predict the observed pattern of results on other-trials. Empirical data are represented as group mean ± SEM and simulated data are represented as group mean.

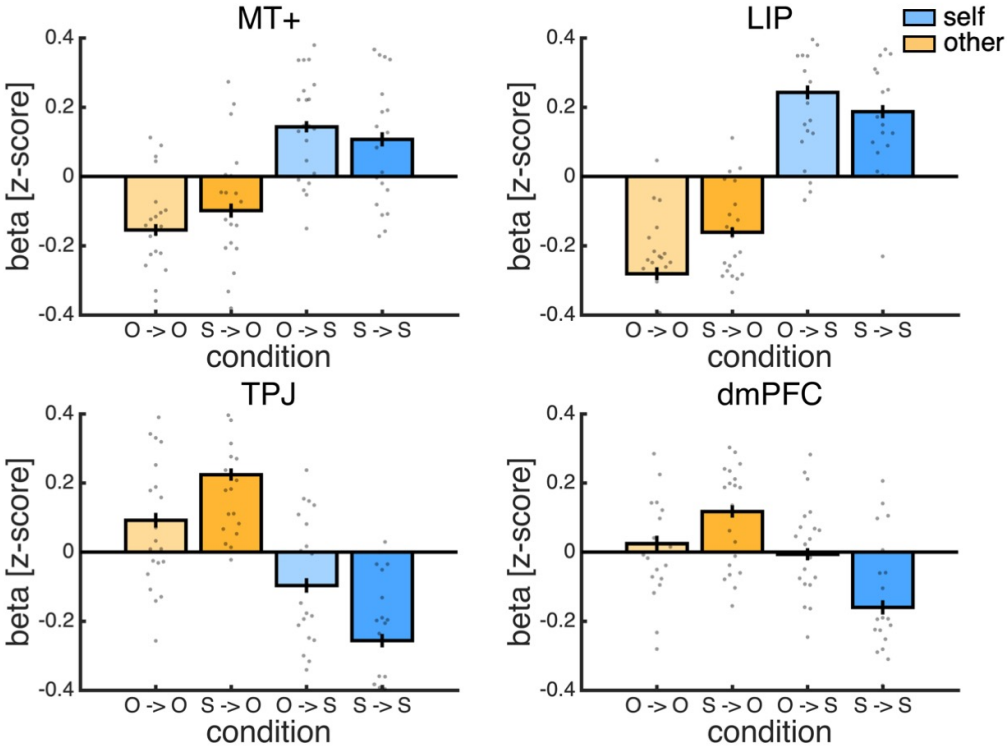

Figure S4. Repetition suppression analysis of agent-specific effects in sensory and social ROIs. In each panel, data are split according to the current trial type (yellow: other, O; blue: self, S) and whether the previous trial type was the same (light: same; dark: different). In MT+, there was only an effect of the current trial type (linear regression: current, $t(20) = 5.27$, $p < .001$; previous, $t(20) = 0.26$, $p = .801$; interaction, $t(20) = -1.12$, $p = .276$). However, in addition to an effect of the current trial type, there was an interaction between the previous and the current trial type (i.e. repetition suppression) in LIP (linear regression: current, $t(20) = 9.23$, $p < .001$; previous, $t(20) = 0.72$, $p = .480$; interaction, $t(20) = -2.85$, $p = .010$), TPJ (linear regression: current, $t(20) = -6.18$, $p < .001$; previous, $t(20) = -0.35$, $p = .730$; interaction, $t(20) = -3.44$, $p = .003$) and dmPFC (linear regression: current, $t(20) = -3.51$, $p = .002$; previous, $t(20) = -0.64$, $p = .530$; interaction, $t(20) = -2.77$, $p = .011$). Under the repetition suppression framework, if a region is primarily involved in self-inference, then activity should be highest on self-trials preceded by other-trials, and lowest on other-trials preceded by other-trials – a pattern observed for LIP. By contrast, if a region is primarily involved in other-inference, then activity should be highest on other-trials preceded by self-trials, and lowest on self-trials preceded by self-trials – a pattern observed for TPJ/dmPFC. Statistical analyses were performed on single-trial data. Data are single-trial activity estimates and represented as group mean ± SEM. Each dot is a subject.

parameter	range	sampling	note
sensation (all)			
noise, σ	[.01,2]	$\sigma \sim normal(a,b)$ where a and b are obtained from a posterior distribution over sensory noise as fitted in the prescan session	
sensory sample, x	[-1,1]	$x \sim normal(a,b)$ where a is the motion stimulus, $k\theta$, and b is noise, σ	
softmax (all)			
bias, β_0	[-10,10]	$\beta_0 \sim normal(a,b)$, where a and b are group-level parameters, $a \sim normal(0,10)$ and a: $[-10,10]$, $b \sim uniform(.01,10)$	same (B1) or different (B2) bias for self and other
temperature, β_1	[.01,2]	$\beta_1 \sim normal(a,b)$, where a and b are group-level parameters, $a \sim normal(.2,2)$ and a: $[.01,2]$, $b \sim uniform(.01,2)$	same (B1) or different (B2) temperature for self and
stimulus space			
factor, p	[.01,10]	$p \sim normal(a,b)$, where a and b are group-level parameters, $a \sim normal(2,4)$ and a: $[.01,10]$, $b \sim uniform(.01,4)$	linearly-spaced (P0) or re-scaled (P1) stimulus space
learning (Q, Q*, T, T*)			
alpha, α	[0,1]	$\alpha \sim beta(a,b)$, where a and b are group-level parameters, $a = x * y^{-2}$, $x \sim normal(.2,1)$ and x: $[0,1]$, $y \sim uniform(0,2)$, $b = y^{-2} - a$	
covariance (only T*)			
rho, ρ	[-.95,.95]	$\rho \sim normal(a,b)$, where $a \sim normal(0,.5)$ and $b \sim normal(0,1)$	

Table S1. Overview of model parameters and priors as implemented in STAN. Model names are abbreviated as follows: S, S-model; Q, Q-model; Q*: Q*-model; T, ToM-model; T*, ToM*-model. See online repository for code implementation. Here '∼' indicates 'is distributed as' and ':' indicates 'range'.

contrast	label	voxels at $p < .001$	peak z-score	p (cluster FWE corrected)	peak voxel MNI coordinates			
self > other	extrastriate, parietal, motor, cingulate	23632	6.25	<.001	-34	-16	68	LR
	thalamus	252	5.95	<.001	-22	-26	-6	L
	thalamus	145	5.49	<.001	22	-26	4	R
	inferior frontal gyrus	116	5.22	.002	-58	6	26	L
	insula	656	5.04	<.001	42	0	8	R
	striatum	222	4.40	<.001	12	18	-8	R
	striatum	192	4.36	<.001	18	-10	22	R
other > self	dorsomedial frontal	1941	5.30	<.001	14	30	52	LR
	temporoparietal	892	5.30	<.001	-48	-56	26	L
	temporoparietal	741	5.23	<.001	50	-62	36	R
	temporal	848	4.91	<.001	-58	-38	0	L
	precuneus	607	4.85	<.001	2	-60	38	LR
	middle frontal gyrus	177	4.75	<.001	48	24	36	R
	temporal	525	4.42	<.001	66	-28	-8	R
	middle frontal gyrus	467	4.41	<.001	-36	6	46	L
	temporal	82	3.85	.014	58	-8	-30	R

Table S2. Summary of significant activations for GLM1: decision phase. All reported activations are significant at $p < .05$, FWE-corrected for multiple comparisons. Cluster-defining threshold: $p < .001$, uncorrected. FWE: familywise error. MNI: Montreal Neurological Institute. L: left. R: right.

contrast	label	voxels at $p < .001$	peak z-score	p (cluster FWE corrected)	peak voxel MNI coordinates			
self > other	posterior cingulate	87	4.16	.009	-12	-52	10	LR
	ventromedial prefrontal	66	3.95	.037	2	38	-4	LR
other > self	middle frontal gyrus	258	5.44	<.001	-42	10	48	L
	temporoparietal	719	5.30	<.001	-48	-58	24	L
	temporal	503	5.30	<.001	64	-24	-12	R
	temporoparietal	550	5.23	<.001	54	-60	30	R
	dorsomedial frontal	1751	5.04	<.001	0	32	52	LR
	temporal	467	5.01	<.001	-64	-46	-2	L
	precuneus	176	4.96	<.001	6	-66	42	LR
	middle frontal gyrus	316	4.74	<.001	44	22	38	R
	temporal	414	4.68	<.001	-50	12	-28	L
	orbitofrontal	140	4.63	<.001	42	52	-12	R
	orbitofrontal	529	4.48	<.001	-34	54	-10	L
	orbitofrontal	97	4.44	.004	52	22	-6	R
	cerebellum	107	4.43	.002	-8	-80	-28	LR
	orbitofrontal	177	4.36	<.001	-28	22	-8	L
	precuneus	83	4.26	.011	6	-52	34	LR
	orbitofrontal	89	4.21	.007	-10	48	-24	LR
	temporal	84	4.17	.010	58	-38	2	R
	cerebellum	116	4.16	.001	-30	-84	-30	L
	orbitofrontal	75	3.98	.019	30	18	-6	R

Table S3. Summary of significant activations for GLM1: gamble phase. All reported activations are significant at $p < .05$, FWE-corrected for multiple comparisons. Cluster-defining threshold: $p < .001$, uncorrected. FWE: familywise error. MNI: Montreal Neurological Institute. L: left. R: right.