



Published in final edited form as:

Nat Neurosci. 2021 January ; 24(1): 116–128. doi:10.1038/s41593-020-00746-9.

Neuronal Correlates of Strategic Cooperation in Monkeys

Wei Song Ong¹, Seth Madlon-Kay², Michael L Platt^{1,3,4}

¹Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania

²Department of Biostatistics and Bioinformatics, Duke University School of Medicine

³Department of Psychology, School of Arts and Sciences, University of Pennsylvania

⁴Marketing Department, Wharton School of Business, University of Pennsylvania

Abstract

We recorded neural activity in male monkeys playing a variant of the game “chicken” in which they made decisions to cooperate or not cooperate to obtain rewards of different sizes. Neurons in mid superior temporal sulcus (mSTS), previously implicated in social perception, signaled strategic information, including payoffs, other player’s intentions, reward outcomes, and predictions about the other player; a subpopulation of mSTS neurons selectively signaled cooperatively obtained rewards. Neurons in anterior cingulate gyrus (ACCg), previously implicated in vicarious reinforcement and empathy, carried less information about strategic variables, especially cooperative reward. Strategic signals were not reducible to perceptual information about the other player or motor contingencies. These findings suggest the capacity to compute models of other agents has deep roots in the strategic social behavior of primates and that ACCg and mSTS support these computations.

Introduction

The evolutionary, economic, and biological origins of human cooperation remain hotly debated ^{1,2}. Both emotional and cognitive mechanisms shape the decisions people make when they interact with others, especially whether to cooperate. Vicarious feelings of reward or pain experienced by another (empathy) can provoke prosocial actions³. Strategic reasoning about the beliefs, desires, and goals of another individual (mentalizing or theory-of-mind) guides the decision to cooperate with or betray a partner⁴. These two processes interact; manipulations that increase empathy enhance cooperation ^{5,6}.

Two separate but interacting brain systems appear to support empathy and mentalizing during social decisions ⁷. In humans, empathy and vicarious experience evoke hemodynamic

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding author, Wei Song Ong.

Author contributions

WSO conceived and carried out the experiments and analyzed the data. SM-K designed the strategic-learning model. WSO and MLP wrote the manuscript with input from SM-K.

Ethics declaration:

The authors declare no competing interests

activity in anterior cingulate gyrus (ACCg), anterior insula, and amygdala, and neurons in primate ACCg and amygdala signal experienced by other monkeys^{8,9}. By contrast, thinking about the beliefs, desires, or goals of others evokes hemodynamic activity in the dorsomedial prefrontal cortex (dmPFC) and temporo-parietal junction (TPJ) in humans^{7,10,11}. The neuronal mechanisms underlying such mentalizing-related activity, however, remain poorly understood, in part due to difficulty in eliciting strategic social behavior in primates or other animals in which neuronal activity can be studied directly¹² as well as lack of neurophysiological or histological evidence for a TPJ homolog in nonhuman primates¹³.

Recent studies in macaques proposed that portions of the middle superior temporal sulcus (mSTS), an area of extrastriate visual cortex known to encode perceptual social information like faces and bodies,^{14,15} may be the primate homolog of TPJ based on patterns of correlated hemodynamic activity revealed by resting-state functional MRI. An fMRI study in macaques found multiple patches along and within the STS are selectively activated by viewing social interactions compared with viewing interactions between objects¹⁶. Sallet et al. (2011)¹⁷ reported the size of an area within the ventral bank and fundus of the mSTS varied with social network size in macaques, suggesting a functional relationship between portions of the mSTS and social interactions. These findings invite the possibility that inferences about the mental states of others and predictions of their behavior, which are associated with activation of human dmPFC and TPJ, emerge from integration of perceptual information about the actions and behavioral states of others, which in macaques is encoded by neurons in mSTS and other areas, with information about reward outcomes and prior interactions. Whether neurons in macaque mSTS encode information supporting inferences about the mental states of others during social interactions remains unknown.

Mentalizing in humans is typically studied in two ways. One, participants are asked to verbally report what they think another individual believes, wants, or will do¹⁸. Examples of this approach include false belief tasks¹⁸, reading the mind in the eyes test (RMET)¹⁹, or psychometric questionnaires such as the Social Responsiveness Scale (SRS)²⁰. Two, participants engage in strategic gameplay and their choices are modeled to infer the underlying computational processes²¹. This approach derives from game theory in economics, and includes classic tasks such as Prisoner's Dilemma, Dictator Game, Stag Hunt, and others. Importantly, the game theory approach can be applied to the behavior of animals, making it well-suited to uncovering the computational processes and neurobiological mechanisms underlying strategic behavior^{12,22}.

Here we trained monkeys to play a novel variant of the "chicken" game²³. Metaphorically, this game pits two cars driving towards each other. Deviating yields a small "chicken" reward, whereas going straight, without colliding, yields a larger reward. If both players go straight, they crash and no one receives reward. Chicken is uniquely suited to contrasting selfish and prosocial preferences²⁴. When played iteratively, chicken offers players the opportunity to learn about the motivations and tendencies of the other player, and to cooperate by taking turns choosing the large and chicken rewards. Iterated chicken games take longer to converge on an equilibrium, or steady state, than other games like prisoner's dilemma,²⁵ consistent with a greater depth of reasoning and extended learning. Iterated play

of chicken thus provides a powerful paradigm to expose the computations underlying strategic social decisions in animals.

The classic chicken game is an anti-coordination game in which the best strategies and outcomes are achieved by players making different moves.²⁴ Typically, choices are made simultaneously by both players. These features of the classic chicken game make it difficult to develop simultaneous cooperation.^{24,26} We modified the classic chicken game in three important ways to promote the development of cooperation between monkeys and to uncover the contributions of perceptual and cognitive processes on strategic behavior. First, we added a “cooperation bar” that allowed monkeys to either cooperate with each other to obtain equal rewards. This approach is explicitly modeled on rope-pulling tasks used in prior studies to elicit cooperation in chimpanzees²⁷, bonobos²⁸, capuchin monkeys²⁹, hyenas³⁰, and children³¹. Second, we added a visible cue that on half of trials was correlated with a monkey’s currently favored option, which the other monkey could use to help guide his decision and potentially coordinate their actions; on the other half of trials this cue was uninformative. Third, we varied social context: monkeys played against a live monkey, a computer, or a decoy monkey, allowing us to determine the impact of social perceptual cues that might inform strategic decisions. Our approach thus permitted us to study the computational processes and neuronal mechanisms mediating cooperation on single trials, as the product of current payoffs, prior history of rewards, previous choices, prior behavior of the other player, and social context. To formally guide our analyses, we developed an iterated series of mathematical models that assumed increasingly sophisticated computations underlying decision-making

Our task was physically realized as follows (see supplementary material: task video). In each session, a monkey played against another monkey (‘live’), against a computer (‘computer’), or against a computer with a decoy monkey present (‘decoy’). In live and decoy conditions, two monkeys faced each other across a horizontal screen (Figure 1A). Two colored rings (hereafter ‘cars’), 6 arrays of tokens, and a white “cooperation bar” were presented (Figure 1B & C). Each car, indicated by color and position, was controlled by a monkey player. Monkeys used a joystick to choose whether their car would go straight or deviate. Token arrays, also identified for each monkey by color, indicated amount of juice reward available for choosing straight, for deviating alone (‘chicken reward’), and for cooperating. If one monkey chose straight and the other deviated, each would receive an amount of juice proportional to the tokens acquired. If both monkeys chose straight, the cars “crashed” into each other, and no reward was delivered. If both monkeys chose to deviate, they each received the associated rewards plus bonus tokens exposed by moving the cooperation bar (Figure 1C & D). Number of tokens, and thus available rewards, varied randomly from trial to trial.

Each car enclosed white dots that moved smoothly in the direction in which the joystick was currently held, providing a cue to the other monkey that could either be visually clear (90% correlated dots) or ambiguous (0% correlated dots). Dot motion coherence was randomized across trials. Holding the joystick in one direction for 0.5 secs committed the player to a choice, whereupon the dots changed color from white to the players’ color (Figure 1, M1=red and M2=blue).

Results

Monkeys use perceptual information and strategic inference when deciding whether to cooperate

Overall, monkeys' choices aligned with the Nash equilibrium, a steady state in which no player can gain an advantage by unilaterally adopting another strategy³²(Figure 2A, example pair; Yellow outlines = Nash equilibrium). Figure 2A plots frequency of choices for an example pair of monkeys, segregated by dot correlation and payoffs. When joystick movements were clearly signaled (high signal, 100% dot correlation), the pair was optimal in both payoff conditions, which favored either a mixed-strategy Nash equilibrium ($V_{str} > V_{coop} > V_{safe}$, Figure 2A, top left) or a pure-strategy Nash equilibrium ($V_{coop} > V_{str} > V_{safe}$, Figure 2A, bottom left). When joystick movements were ambiguous (low signal, 0% dot correlation), monkeys' choices again aligned with the Nash equilibria, but with more departures (Figure 2A, right column). This was true for all 4 pairs: 91% of choices aligned with the predicted Nash equilibria in the high signal condition and 69% in the low signal condition (Extended Figure 1B). The fact that monkeys largely avoided crashing even in the low signal condition suggests they understood the rules of the game, the payoffs available to each player, and the likelihood the other player would choose each option. It is possible that monkeys also used additional information to guide their choices, including the eye gaze, facial movements, and body posture of the other monkey.

To test these ideas, we first examined monkeys' patterns of gaze during gameplay. After a brief fixation period to start each trial, gaze was unconstrained. Monkeys spent most of the token onset period (500ms) looking at the tokens in front and to the side of the screen, regardless of social context (Figure 1C). Across all social contexts, monkeys spent 29.5% (± 1.4) of the first 500ms of the choice period looking at the other player's car (Figure 2Bi, top panel). However, they spent less time looking at the other player's car in the live context (26.7% $\pm 1.9\%$) compared to either the decoy (32.5% $\pm 2.6\%$) or computer (30.9% $\pm 3.0\%$) conditions. In the live context, monkeys spent more time (2.1% $\pm 0.5\%$) looking at the other monkey's face than they did in either the decoy (1.0% $\pm 0.2\%$) or computer (1.3% $\pm 0.2\%$) conditions. However, monkeys looked more frequently at the other monkey's face during the middle of the choice period (2-2.5s) in the decoy condition than they did earlier in the trial (3.1% $\pm 0.3\%$, compared to 2.4% $\pm 0.2\%$ in the computer condition), similar to their frequency of looking at the face of the other player in the live condition (3.1 $\pm 0.2\%$). Thus, monkeys looked at key sources of information during the trial and their gaze patterns were different in social contexts that were perceptually similar (decoy vs. live).

Gaze also varied with availability of joystick direction signals conveyed by moving dots within the cars. Monkeys looked more frequently at the other player's car when correlated dot motion was high (22.9% $\pm 1.8\%$) than when it was low (12.6% $\pm 1.0\%$) during the initial 0.5s-1.5s of the choice period (Numbers shown are for live players but were similar across all social contexts; Extended Figure 1C, lower panel). During the first 0.5s of the reward delivery period, monkeys were more likely to look at a live monkey (4.1% $\pm 0.7\%$) than towards a decoy monkey (2.3% $\pm 0.4\%$), and were more likely to look at another

monkey (live or decoy) than the dripping juice tube (computer condition, 2.0% \pm 0.3%, Extended Figure 1Cii). Finally, monkeys were much less likely to look at a monkey with whom they had cooperated (0.9% \pm 0.2%) than one who had acted selfishly (5.8% \pm 1.1%), even after controlling for reward size ($p < 1 \times 10^{-40}$, Supplementary Figure 1A. Example sessions Figure 4 & 5, top panel). Thus, monkeys adaptively sampled visual information about payoffs, the visible cues available about the other player's joystick movements, and the face of the other monkey. Finally, differences in gaze patterns in the live and decoy conditions, which were perceptually similar, suggested monkeys developed a sense of the agency of the other player.

These data invite the hypothesis that monkeys sampled multiple sources of information to compute utility functions to guide their decisions. The utility function describes the subjective value, or desirability, of each option, as inferred from observed choices³³. We assumed monkeys acted to maximize their own utility. To infer how these utility functions were computed and updated across trials, we compared behavior to a series of decision models of increasing sophistication (Figure 3A, for model specifications and equations, see Methods). Each model assumes monkeys calculated the expected value of each option based on a prediction of the other player's actions (p , figure 3A). If the other player was predicted to go straight, the monkey should deviate to secure the small but safe 'chicken' reward instead of risking a crash and obtaining no reward. In the least sophisticated model, the other player chooses with a fixed probability that is not influenced by visible payoffs but is influenced by reward history ('*naïve RL*'). In the next more sophisticated models, monkeys compute their own utility functions with the inclusion of a logistic function based on the visually cued rewards available on the current trial ('*logistic*') and the visually cued reward available on the current trial combined with reward history ('*combined logistic-RL*'). In the more sophisticated models, monkeys compute the utility functions for both themselves and the other player (X_t , '*intention*'), and choose accordingly, and then learn adaptively about the other player's strategies based on experience ('*SPE-learning*'). The SPE-learning models update beliefs about other player's strategies using a strategic prediction error (SPE), which is calculated as the difference between the other player's predicted strategy and his actual choice.

The best-fitting model was the most sophisticated, and included both representation of the other player's utility functions and SPE-driven learning (full model, in comparison to the combined logistic-RL sub-model resulted in a mean decrease in $AIC = 783$, or 0.0313 AIC /trial; cf. ³⁴ Figure 3B). By contrast, monkeys' choices did not follow tit-for-tat³⁵ or win-stay-lose-shift³⁶ strategies (Supplementary Figure 2B). We found no evidence that simple reinforcement-learning³⁷ could account for the behavior of the monkeys in the game (absolute beta q -values for all conditions < 0.01 , in comparison to weights for the expected value component of the model (beta v), which ranged from 1-3 across monkeys. For more details see Methods). Our modeling results indicate monkeys behaved as if they computed, at least implicitly, a model of the other players' utility functions and strategies in order to predict the other player's choices (Figure 3D). Furthermore, it captures subtleties in the monkey's own behavior in response to the visible payoffs and signal conditions (Figure 3C). Additionally, our results indicate monkeys can update their model of other players based on

new information by learning via the predicted strategy of the other player and his actual choice (SPE).

Whether a monkey learned from the other player's actions depended on his social rank (Extended Figure 2D). We compared the most sophisticated model, in which the monkey both calculates the utility function of the other player and updates that model using SPE, to the less sophisticated model in which the monkey calculates the other player's utility function but does not update his own beliefs about the other monkey. The improvement in fit with the model that includes strategic learning was greater for subordinate monkeys than for dominant monkeys. This was true for every pair of monkeys (Extended Figure 2D). These findings suggest subordinate monkeys were more sensitive to evidence of change in the utility functions and strategies of dominant monkeys, consistent with prior reports that subordinate monkeys pay more attention to dominant monkeys, who themselves attend selectively to other dominant monkeys^{38,39}. When the same mid-ranking monkey played against monkeys who were of higher or lower rank, (brown and purple in Extended Figure 2C,D), strategies evident in play were more consistent with relative dominance than individual identity.

In summary, our behavioral and eye-tracking data demonstrate monkeys are sensitive to payoffs for themselves and other players, visible information about joystick movements, and reward outcomes, as well as social information reflecting identity and dominance status. Monkeys used this information to compute predictions about the other player's actions, including how likely he was to behave cooperatively, as well as updated their computations to improve future play. Neural circuits supporting the behavior of monkeys playing our chicken game would, at minimum, be expected to encode this information in some way.

Neurons in ACCg and mSTS signal strategic information guiding decisions

We recorded spiking activity of 448 mSTS neurons and 528 ACCg neurons (Extended Figure 3). Firing rates of neurons in both areas were sensitive to payoffs early in the trial, the availability of visible joystick direction signals within the cars, and the amount of reward received. The clarity and prevalence of these signals varied between brain areas, across social contexts, and as a function of time during each trial (Extended Figure 3B, Supplementary Figure 4).

Figure 4 (lower panels) shows an example mSTS neuron recorded in a monkey playing the game with a live monkey. The firing rate of this neuron remained steady during the illumination of payoff tokens, cars and moving dots, as well as joystick manipulation. This neuron responded phasically during reward delivery, but fired more strongly for rewards received through cooperation than for equivalent rewards received for non-cooperative actions (Figure 4, lower panels). By contrast the firing rate of this neuron was not modulated by the absolute amount of reward received, regardless of which choice, cooperative or not, was made (Supplementary Figure 3Aii). Both of these findings held true even after removal of trials in which the recorded monkey looked at the other monkey's face during the reward epoch. This neuron also fired more on trials when our model predicted the other monkey was less likely to deviate compared with when the model predicted he was more likely to do so (p' ; Supplementary Figure 3Aiii).

Figure 5 (lower panels) plots the activity of an example ACCg neuron in a monkey playing the game with a live monkey. When moving dots appeared in the cars and the players manipulated their joysticks, firing rate during the decision period differed for “chicken” and “cooperate” choices, which were actuated by the same movement (deviate; Figure 5, lower panels). After juice delivery, this neuron fired less for cooperatively obtained rewards compared with rewards obtained through non-cooperative actions (Figure 5, lower panels). Analyzing only those trials on which the recorded monkey did not look at the other monkey’s face, this neuron responded differentially to delivery of cooperatively obtained and non-cooperatively obtained rewards (Figure 5, bottom panel: no-look trials).

We used linear models (LMs, 20-fold cross-validated) to quantify neuronal sensitivity to payoffs, availability of joystick direction signals, reward outcomes, cooperation, opponent’s predicted strategy, strategy prediction error (SPE), and gaze toward the other player’s face. Across the population, we found that activity of 33% of mSTS neurons differed between payoff conditions during the period when tokens were presented (0-500ms from token onset), but only 14% of ACCg neurons responded differentially to the presentation of different payoff conditions (Supplementary Figure 4A). In this same time period, a smaller proportion of neurons in both areas encoded SPE from the previous trial, as estimated by our model, which is a signal that could potentially influence behavior and neural activity on the current trial (10% in mSTS, 8% in ACCg, Supplementary figure 4A).

Following dot onset, the firing rates of many neurons in both brain areas were modulated by the presence of coherently moving dots in the cars during the decision period (34-35%, Supplementary Table 5 & 6). Firing rates of very few neurons (<8-15%) were modulated by looking at the other player’s face. Firing rates of approximately 13-19% of neurons varied as a function of potential payouts and 16-18% of neurons forecast cooperation. A further 14-18% of neurons carried information about the other player’s predicted strategy (p'). There were no significant differences between ACCg and mSTS in the prevalence of these signals and nor did they differ significantly as a function of social context.

Firing rates of 19-26% of neurons in both brain areas were significantly modulated by impending reward payouts in the post-decision period preceding juice delivery, whereas firing rates of between 29% and 47% neurons were significantly modulated by the amount of obtained reward in all three social contexts during reward delivery (250-1250ms post juice delivery, 2-way ANOVA, $F=4.92$, $p=0.007$, Figure 6A). Remarkably, firing rates of 38% of mSTS neurons were modulated by whether rewards were obtained through cooperation, compared to only 20% in ACCg (2-way ANOVA, $F=40.24$, $p<10^{-7}$). Firing rates of roughly 20% of neurons in both areas carried information about the other player’s predicted strategy (p' ; Figure 6A). Activity of only a small percentage of neurons was modulated by whether the monkey looked at the other monkey during the reward period (ACCg=7-11%; mSTS=10-16%).

Excitation and suppression of neuronal activity as a function of cooperation varied over time during the reward period. Early (250-750ms post-reward), 20-23% of mSTS neurons responded to cooperation by increasing firing rate, while 7-15% did so by decreasing firing rate. Later (750-1250ms post-reward), 25-29% of neurons decreased firing rate for

cooperation while 7-15% of neurons increased firing rate. Overall, we found that firing rates of 50-60% of mSTS neurons were modulated in one or both reward epochs by whether or not the monkeys had cooperated (36-45% for either epoch alone, 15-16% for both epochs), compared to a much smaller percentage of ACCg neurons (16-20% for either epoch alone, 5-6% in both epochs).

We used ROC analysis to determine if we could decode whether monkeys had cooperated based on firing rates during the reward period—the task epoch in which firing rates showed the clearest and most prevalent modulation by cooperation. The area under the ROC curves for both brain areas and each agency condition was significantly different from that computed from the same firing rate data randomly shuffled across trials (Figure 6C, $p < 10^{-20}$, 20-fold cross validation, see methods for more details). More neurons in mSTS discriminated between cooperative and non-cooperative rewards than neurons in ACCg did (36/448 and 4/528 cells, respectively; 2-way ANOVA, $F = 104$, $p < 10^{-20}$; AUROC > 0.7).

Prior neurophysiological studies of mSTS in monkeys identified neurons in this area that respond selectively to the sight of faces, bodies, and direction of gaze⁴⁰⁻⁴². Such perceptual sensitivity could potentially confound the strategic signals we identified in the firing rates of mSTS neurons, and potentially ACCg neurons as well. We addressed this possibility in two ways.

First, we compared neuronal activity on trials where the recorded monkey looked at the other monkey to trials on which he did not, during live and decoy conditions. We found that activity of only a small percentage of neurons was modulated by whether gaze was directed at the other monkey or not (10-16% in mSTS, 7-11% in ACC, Figure 6, also supplementary table 4), using the same linear models (LMs, 20-fold cross-validated) we used to quantify neuronal sensitivity to payoffs, availability of joystick direction signals, reward outcomes, cooperation, opponent's predicted strategy, and gaze toward the other player's face. We also applied the linear model analyses solely to those trials on which the monkey did not look at the other monkey during the juice delivery period, and found the same pattern of results (Supplementary Figure 5C). Together, these analyses suggest strategic signaling by neurons in mSTS and ACCg is not reducible solely to perceptual information associated with looking at the face of another monkey—at least within the context of playing our chicken game.

Second, we characterized visual responsiveness to images of faces, bodies, and other complex objects, as well as perceptual selectivity for faces and bodies, in a subset ($n = 227$) of mSTS neurons subsequently studied in the chicken game. mSTS neurons showing modulations in firing rate during passive viewing of images ($n = 118/227$, 52%, example neuron, Figure 7Ai) tended to show modulations in firing rate during illumination of payoff tokens in the chicken game ($n = 73/118$, 62%. Same example neuron does not show modulation, Figure 7Aii), consistent with visual responsiveness to the onset of payoff tokens in the full population of mSTS neurons studied in the chicken task (256/448, 62.5%). By contrast, fewer ACCg neurons showed modulations in firing rate during illumination of payoff tokens in the chicken game ($n = 157$ out of 528, 29.7%; note we did not study the responses of ACCg neurons during passive viewing). Overall, ACCg neurons were less likely to be visually responsive to payoff token onset in the chicken game than were mSTS

neurons regardless of whether the other player was a live monkey, decoy, or computer (2-way ANOVA, $p < 10^{-20}$, $F = 89.5$).

In mSTS, 51% (n=118) of neurons responded differentially to distinct types of visual images during passive viewing. Of these, we identified a subpopulation of neurons (n=53, 23%) that responded selectively to faces (n=15, 6%) or bodies (n=17, 7%) of other monkeys or both (n=21, 9%), but not to other complex objects (Figure 7Ci). These neurons were no more likely than any other subpopulation of visually selective neurons to signal strategic information in the chicken game (Figure 7B, Supplementary Table 2). Importantly, the subpopulation of mSTS neurons that selectively signaled strategic information in the chicken game (cooperation, n=87; (p' , n=22) were not more likely to show feature selectivity or selectivity for monkey faces or bodies compared to the overall population (cooperation, t-test, $p=0.78$, t-stat =0.2853; (p' , t-test, $p=0.93$, t-stat = -0.087; Figure 7Cii, Supplementary Table 2). These findings demonstrate that the subpopulation of neurons encoding strategic social information, such as cooperation and predicted strategy, are not over- nor under-represented within the population of neurons that signal perceptual social information. These subpopulations were not segregated spatially but rather were intermingled within the mSTS. Together, these analyses indicate that strategic social signaling by mSTS neurons does not simply reflect visual sensitivity to faces and bodies, or other visual features of the chicken game task.

Our findings demonstrate that neurons in ACCg and, even more so, in mSTS signal strategic information, like available payoffs, availability of joystick direction cues, reward outcomes, and whether rewards were achieved cooperatively, independent of the physical social context or attention to species-specific social perceptual information. This is strong evidence that these two brain areas encode important components of the decision model that best accounted for monkeys' decisions in the chicken game.

Discussion:

Here we show that monkeys adroitly navigated a strategic game and tended to cooperate when payoffs favored working together. Monkeys' decisions were best explained by a model that included other player's intentions, goals, and strategies. Monkeys did not use simple strategies such as tit-for-tat³⁵ or win-stay-lose-shift³⁶ to play the game (Supplementary Figure 2A), nor was their behavior explained by simple reinforcement learning. Monkeys attended to payoffs available for both themselves and the other player, as well as joystick direction cues indicating the other player's currently favored choice, and readily distinguished the agency of decoy and live players via their eye gaze patterns and early choice behavior in each session. These findings suggest monkeys implemented a sophisticated model of the other player in the game, and the complexity of this model varied with relative social status. Like humans⁴³, low status monkeys deployed greater expertise when interacting strategically with higher status monkeys⁴⁴, who themselves were more likely to simply choose selfishly when playing a lower status monkey (Extended figure 2C).

Brain imaging studies in humans indicate two interacting systems, one associated with empathy and social emotions and the other linked to mentalizing and strategic reasoning,

support social interactions²². Our findings show that neurons in ACCg and the mSTS, previously implicated in vicarious reinforcement and social perception, respectively, encode abstract information associated with strategic gameplay. Notably, strategic signals were clearer and more prevalent in mSTS than in ACCg, and were sensitive to whether the other player was a live monkey or a decoy. Prior neurophysiological studies in STS revealed patches of neurons that selectively respond to the sight of faces^{40,41}, facial expressions⁴², and the direction of another individual's gaze⁴⁰. Comparison of firing rates during passive viewing of faces, bodies, and objects with firing rates during strategic gameplay revealed that overlapping populations of mSTS neurons encoded perceptual and abstract social information. Importantly, neurons encoding perceptual social information comprised only a small subset of mSTS neurons encoding strategic information.

Our findings provide neurophysiological evidence for the representation of strategic information, in addition to perceptual social information, in primate mSTS. These results align with a recent model derived from analyses of BOLD fMRI signals in macaques viewing interactions of monkeys with other monkeys or objects, as well as purely physical interactions between objects. In that model, face patches in STS, including mSTS, provide an entry point for perceptual information about the interactions of others to a “social interaction network”—including multiple areas in temporal, parietal, and prefrontal cortex—which supports inferences about the social interactions of others. By contrast, the model suggests body patches in STS provide a gateway to the “mirror neuron system”—including premotor areas in parietal and prefrontal cortex—which supports inferences about goals and intentions of others interacting with both objects and other individuals. Our findings indicate mSTS maintains a surprisingly rich representation of strategic information supporting social decision making, which had not been uncovered previously.

Notably, neurons that selectively responded to faces during passive viewing were no more likely to encode strategic information than neurons that did not respond selectively to faces. This was also true for neurons that responded selectively to body parts. The observation that these neurons were intermingled within mSTS could mask visibility of these signals to fMRI, which relies on hemodynamic responses reflecting aggregate activity of many thousands of neurons⁴⁵. This same conjecture would apply to using fMRI to image subpopulations of mSTS neurons that encoded strategic information via excitation versus suppression. Future studies using a combination of brain imaging and neurophysiological recordings in macaques playing strategic games would provide crucial information for further testing the relationships between encoding of perceptual social information and strategic information in mSTS, as well as the larger networks encoding social interaction¹⁶.

Although we found evidence that neurons in both ACCg and mSTS signal strategic social information, representation of this information was more abundant in mSTS—particularly the frequency with which neurons encoded cooperatively-obtained rewards. This finding aligns with the hypothesis that portions of mSTS may share functions with the human TPJ, which has been implicated in the representation of strategic social information¹⁰ and strongly linked to mentalizing, and theory-of-mind⁷. ACCg, by contrast, has been more closely linked with tracking the rewards⁸ and emotional experiences of others⁴⁶ to support psychological processes including empathy⁴⁶ and social learning, in both monkeys and

humans^{8,47} Nevertheless, our observation that some neurons in ACCg encode strategic social information monkeys used to play chicken aligns with prior studies showing neurons in this area signal reward outcomes and predicted social outcomes, including cooperation, in monkeys playing Prisoner's dilemma¹² and during social learning. Correspondence between the location of neurons encoding strategic social information in our study, concentrated on the lower bank of STS, and the locations most closely linked to the functional connectivity profile of human TPJ⁴⁸ concentrated on the upper bank of STS, remains uncertain and requires further study.

Surprisingly, there were only small differences in encoding of strategic information between live, decoy, and computer conditions in both mSTS and ACCg, despite the fact that these conditions differed perceptually and socially. Although some neurons in our population, particularly in mSTS, were sensitive to where the monkey being recorded from was looking, and that these effects varied as a function of social context, multiple analyses indicate the strategic signals we uncovered cannot be accounted for solely by perceptual sensitivity. In light of these findings, we speculate that the minimal differences in neural encoding of strategic information across social contexts in both mSTS and ACCg reflects the fact that the strategic behavior of monkeys in the chicken game was also similar across social contexts. Our model showed that the behavior of monkeys was similarly complex and sensitive to the utilities, strategies, and strategic prediction errors inferred for the other player—whether live, decoy or computer. This suggests monkeys learned and deployed an internal model of the other player that was abstract yet only weakly sensitive to the presence and agency of another monkey player, as evidence by differences in cooperation and “crashing” between live and decoy conditions early in sessions that rapidly subsided—often within 10 trials—as monkeys played the game. Based on this evidence, we would expect that neurons encoding strategic information would be relatively insensitive to the physical presence or absence of another monkey, if these neurons were indeed engaged to inform the decision process. We also note that people often attribute intentions, beliefs, and desires to computers and other objects and that “animism” is widespread in traditional societies.

Though long thought to be uniquely human, the ability to strategically navigate mixed-motive games likely characterizes many social animals, particularly primates, that form differentiated relationships, including alliances and friendships, to deal with the complexities of group life⁴⁹. For long-lived, social primates like macaques, success depends on deft deployment of cooperation and competition, which leverages individual identification, memory for previous interactions, investment of biological capital, learning, knowledge of others' social relationship, and sensitivity to the quality of potential allies⁵⁰. Both prosocial behavior and cooperation in humans also depend on these factors, strongly suggesting the underlying mechanisms are conserved⁴⁹. Our findings confirm this prediction by demonstrating neurons in mSTS and ACCg encode a wealth of strategic information, including payoffs, visible cues to the other player's intentions, reward outcomes, whether rewards were obtained through cooperation, strategic tendencies of other players and deviations from these patterns. Our findings thus support the hypothesis that human societies, with all the complexity that attends cooperation and selfishness, arise from biological mechanisms that evolved in the primate clade to support social interactions.

Methods:

All experimental methods were approved by the Duke University Institutional Animal Care and Use Committee (protocol registry number: A295-14-12) and were conducted in accordance with the Public Health Service *Guide to the Care and Use of Laboratory Animals*.

Subjects:

Five male rhesus macaques (8.6-13.2kg, 9-15 YO) were implanted with head-restraint prostheses (Crist) and neurophysiological recording cylinders (Crist) using standard sterile techniques as described previously⁸. Animals were initially anesthetized with ketamine hydrochloride and maintained with isofluorene (0.5-5% mg/kg). Enrofloxacin or other veterinarian-prescribed broad-spectrum antibiotics and buprenorphine for pain management were administered after surgical procedures. The animals were visually monitored continuously for at least 2 hours after surgery. The post-operative recovery period was 4 weeks, during which the animal was given ad libitum access to fluids and no training or testing was carried out. The recording chambers were cleaned at least 3x/week, treated with antibiotics and sealed with sterile caps. During testing, animals were given access to at least 20mL/kg/day of fluids and supplemented with fruit and vegetables. On testing days, all animals earned 87-118% of their daily minimum fluid intake (20mL/kg/day) in the form of tangerine or berry juice during the experiment. The balance in daily fluid intake was made up with water in the evenings, at least 2 hours after the end of testing. Dominance relationships between pairs of monkeys were determined by controlled confrontation (Deaner et al 2005).

Behavioral task:

Monkeys sat in primate chairs (Crist) facing each other at a distance of 30 inches with a 27-inch LCD screen placed horizontally between them (Figure 1a). Their heads were tilted slightly downward, at an approximate angle of 20°, allowing them to view both the other player's face and the screen.

Eye position and pupil diameter for one monkey was sampled at 1kHz using an infrared eye tracker (SR Research Eyelink) mounted on the primate chair. At the start of each trial, the eye-tracker sent timestamps to the experimental software (Matlab), which collated them with timestamps from the neurophysiological recording system (Plexon) and task events (PsychToolBox). The monkeys manipulated joysticks (60Hz) placed within the primate chairs. The front of the primate chair, including the neck plate, was painted black to obscure the shoulders, hands, and joysticks of both monkeys.

The task was presented on a shared horizontal screen between the two animals. To initiate a trial, the monkey whose eye position was being monitored was required to fixate a central white dot on a black background (200ms). The fixation point was then extinguished, and two colored annuli (hereafter "cars") appeared, one above and one below the extinguished fixation point. Each monkey controlled the car located closer to him, which was also cued by color (e.g., blue for M1 and red for M2; Figure 1b, Supplementary Figure 1a is an image

with the task stimuli to scale). To continue, the joysticks of both animals were required to be in the neutral position. After a variable delay (≤ 500 ms), the cooperation bar and four sets of tokens appeared, two for each player, cued by color and position. The number of tokens was proportional to the volume of juice available for achieving that option (0.02ml/token). Five hundred ms later, a moving dot kinematogram appeared within each car. Monkeys committed to a choice by holding their joystick towards a specific token array for 500ms, at which point the white dots in the kinematogram changed to the player's color (e.g., blue for M1 and red for M2); if did not commit to a choice within 4 seconds, the last joystick direction was implemented as his choice. Monkeys were permitted to make multiple joystick movements freely as they deliberated, which were immediately translated into the direction the dots moving within the car ('joystick direction'). After the choice period, the dots disappeared and the cars moved in the chosen direction. Juice rewards were delivered to the monkeys via a tube controlled by a solenoid valve. If the monkey achieved the token array straight ahead or the cooperative outcome, the solenoid opened twice, once for the smaller constant payoff and again for the second set of tokens in the same location.

The total number of tokens presented on the screen was always 41 for each player, divided between two locations, straight ahead and on the side of the screen. At each location, there was a small constant payoff of 3 tokens; the remaining 35 tokens were divided between the two locations in multiples of 5. The payoff for both animals was always symmetrical.

On 75% of trials, the larger reward was opposite the controlling monkey, behind the other player's car; smaller rewards were to the side. To obtain the larger reward, M1 must choose straight, but if M2 also chose straight the cars collided and neither monkey received reward ('*crash*', Figure 1Di). If M1 (M2) went straight and M2 (M1) deviated, M1 (M2) received 28 tokens and M2 (M1) received 3 tokens—the '*chicken*' reward (Figure 1Dii & iii). On the remaining 25% of trials, the smaller reward was opposite the controlling monkey, behind the other player's car, and the larger reward was to the side, with all but 3 tokens behind the cooperation bar. To obtain the larger reward on these trials, both monkeys had to coordinate their movements and drive their cars to push the cooperation bar ('*cooperate*'; Figure 1Div). If only one monkey moved his car to the side and encountered the bar, it did not move and he only received the 3 tokens in front of the bar ('*chicken*' outcome, blue player in Figure 1Dii, red player in Figure 1Diii).

On half of trials, the coherence of the moving dot kinematogram was 0% to obscure intention signals indicating the current directions in which the monkeys were holding their joysticks. Within session controls were trials on which only one monkey's car and tokens were displayed. On control trials (10% total, randomly interleaved), monkeys should always choose straight regardless of the displayed payouts since that would return at least 8 tokens while deviating would only yield 3 tokens. Control trials were excluded from all behavioral and neural analyses unless explicitly mentioned.

Seven monkeys were chosen based on availability and trained to play the task by first playing against a computer player that made straight/deviate choices randomly. We varied joystick direction signal strength to ensure that the monkeys were attending to the projected future motion of the computer player's car. Two monkeys did not reach criterion and were

removed from further studies. The remaining 5 monkeys were deemed to have reached criterion when they were able to successfully avoid crashes 95% of the time when the intention signals were 100%. These monkeys also had choice thresholds, where their probability of crashing was 50%, that ranged between 15-25% dot motion coherence. Only two joystick direction signal strengths were used in the final experiment: high, 90% coherence; and ambiguous, 0% coherence). All trials were randomly interleaved. The monkeys played against different players on consecutive days and no agency condition was repeated on consecutive days (pseudo-randomized on a weekly basis).

Agency conditions

‘Live’: Two monkeys were present in the experimental setup and both of them actively played the chicken game against each other (4 pairs, $n = 75$, 630 trials).

‘Computer’: One monkey was placed into the experimental setup opposite an empty primate chair. Joystick movements and choices from a randomly chosen prior live monkey behavior session were played back as the other player to the current monkey, and any juice rewards obtained by the computer were delivered to the empty primate chair (4 players, $n = 38$, 938 trials).

‘Decoy’: Two monkeys were present in the experimental setup, but only one of the animals was the designated active player. The ‘decoy’ animal sat in the primate chair and drank the juice rewards delivered, but the joystick movements and choices from a prior live monkey behavior session were played back to the active player as the other player (4 players, $n = 49$, 691 trials).

Electrophysiological recordings

We acquired structural magnetic resonance images (3T, 1-mm slices) of each monkey’s brain. We made a mask consisting of a 3mm sphere around a seed at the fundus of the STS ($X = 18.75$, $Y = -10.00$, $Z = -2.25$) according to the Montreal Neurological Institute (MNI) atlas). This location in mid-STS was selected based on research indicating this region exhibits a functional connectivity profile similar to the human temporo-parietal junction (TPJ; Mars et al., 2013). The mask was then converted into the individual monkey’s native-space structural scan to identify our target recording location (‘mSTS’) using FSL’s FMRI Expert Analysis Tool (FEAT) Version 6.0.0. For both mSTS and ACCg (Brodmann areas 24a and 24b), detailed localizations were made using Osirix (<http://www.osirix-viewer.com>) or Horos (<https://horosproject.org>) data viewer.

All neurophysiological recordings were made using single tungsten microelectrodes (FHC). In each recording session, a sterilized single electrode was secured onto the recording chamber (Crist Instrument) via an X-Y stage (Crist Instrument) and an adapter (Crist Instrument). The dura was penetrated using a sterilized guide tube (22 gauge, stainless steel, custom made), and the electrode was lowered through the guide tube via a hydraulic microdrive (Kopf Instruments). Signals were filtered and recorded using an 8-channel recording system (Plexon Inc). In addition to being guided by stereotaxic coordinates and MRI localization, each day we confirmed the recording site by listening to multiunit changes

corresponding to gray and white matter transitions while lowering the electrode. The neuronal locations were chosen based on availability.

For the mSTS recordings, we further verified the recording site by listening for multiunit activity that was visually responsive to a set of 200 images (consisting of human and nonhuman primate faces, body parts, and objects). The neurons selected for recording in mSTS were within 150 μ m of visually responsive cortex. Beyond that, neurons in both ACCg and mSTS were selected for recording based strictly on location, stability, and quality of isolation. Only neurons with a minimum of 300 completed trials in the chicken game were included in the analysis. A total of 528 ACCg and 448 mSTS neurons were recorded and analyzed in 4 monkeys, in 3 social contexts. Live: 256 in ACCg, 208 in mSTS); decoy (142 in ACCg 151 in mSTS); and computer (130 in ACCg, 89 in mSTS).

Analysis of behavioral data

All behavioral data, including joystick movements and eye-tracking data, was collected and analyzed with custom code (MATLAB). Data collection and analysis were not performed blind to the conditions of the experiments.

To visualize event outcomes across payout conditions (difference in the number of tokens available straight ahead, Vstr, and cooperate, Vcoop), trials were sorted into high joystick direction signal trials (90% dot motion coherence) and ambiguous joystick direction signal trials (0% dot motion coherence, low signal).

To analyze eye position, we drew boxes around areas of interest (figure 2B) and quantified the instances in which eye position fell into those areas. The face region of the recipient was determined empirically prior to the experiments and defined as the area between the neck plate and the top bar, and the side panels of the primate chair. We used a large window to capture gaze shifts that were brief in duration and large in magnitude and often directed at varying depths (e.g., eyes, mouth). In the computer condition, this would be the space where the other player's face would have been⁵¹. Eye positions were plotted in 1ms bins, and shown on the figures with standard errors of the mean calculated between behavioral sessions. Trials were also sorted into high and low joystick direction signal trials.

Statistical tests were conducted as two-tailed ANOVAs with multiple comparisons (Tukey's HSD test) unless otherwise specified. All figures are shown with standard errors of the mean unless otherwise noted.

The hybrid reinforcement learning-strategic learning model was estimated by maximum likelihood using the software package Stan⁵² via the MATLAB interface. The choices of each individual player was fit separately from the choices of their opponent, and separately for each social context. The model predicts the choices of one monkey conditioned on the other player's choices, so the data from each monkey pair was used twice; once with a given monkey being the agent whose choices we fit and a second time with that same monkey being the other player, whose choices are conditioned on and treated as fixed. All model comparisons were performed using the Akaike Information Criterion (AIC)⁵³ and Akaike weights⁵⁴. To compare the predictive performance of the model across different subsets of

the choice data, we used the log-likelihood per trial (Figure 3B, Supplementary figure 2), because the total log-likelihoods of a data set scales linearly with the number of trials in the data set.

Model specification and fitting

a) Hybrid reinforcement learning and strategic learning model—We model each monkey's choices using a model that combines a simple reinforcement learning (RL) system with an expected value model. The RL system takes into account only which action, straight or deviate, the monkey has taken and what reward they received, while the expected value system prospectively takes into account which potential reward outcomes are available on each trial (as indicated by the token symbols on the game board), and chooses according to how likely each outcome is.

The probability of the monkey deviating on trial t is determined by the difference in utility between the deviate and straight choices, U_t , according to the equations

$$p(c_t = \text{yield}) = \text{logit}^{-1}(U_t)$$

$$U_t = \beta_0 \cdot s_t + \beta_V \cdot s_t \left(E[V_t^{\text{yield}}] - E[V_t^{\text{straight}}] \right) + \beta_Q \cdot s_t \left(Q_t^{\text{yield}} - Q_t^{\text{straight}} \right) + \beta_\kappa \cdot s_t \left(\kappa_t^{\text{yield}} - \kappa_t^{\text{straight}} \right)$$

where $c_t \in \{\text{yield}, \text{straight}\}$ is the animal's choice on trial t . The utility difference is a linear combination of the output of three valuation sources, denoted by the Q , κ and $E[V]$ values respectively, each weighted by a temperature parameter. We will discuss each of these sources of valuation in turn. Note that the temperature parameter for each system differs depending on the joystick direction signal strength s used on the trial t , which can be either high or ambiguous, for a total of six temperature parameters.

The Q values for deviate and straight are learned through a simple RL system using reward prediction error update equations,

$$Q_t^c = \begin{cases} Q_{t-1}^c + \alpha(r_{t-1} - Q_{t-1}^c) & \text{for } c = c_{t-1} \\ Q_{t-1}^c & \text{for } c \neq c_{t-1} \end{cases}$$

Here the Q value of the previously choice is incremented according to the learning rate α towards the reward received. At the beginning of a session both Q values are initialized to the value Q_0 , which is fit as a free parameter bounded between zero and the largest possible payoff on any trial.

Second, the κ values capture autocorrelations in the monkey's choices, such as a tendency to either repeat or avoid the choice that has been taken recently. Similar to the Q values, the κ values are updated each trial using a Rescorla-Wagner rule,

$$\kappa_t^c = \begin{cases} \kappa_{t-1}^c + \tau(1 - \kappa_{t-1}^c) & \text{for } c = c_{t-1} \\ (1 - \tau)\kappa_{t-1}^c & \text{for } c \neq c_{t-1} \end{cases}$$

where the parameter $\tau \in [0,1]$ determines how rapidly the influence of past choices decays. Each κ value is initialized to zero at the beginning of a session.

Finally, the expected reward values $E[V_t^{\text{yield}}]$ and $E[V_t^{\text{straight}}]$ are estimated by the monkey on each trial based on the potential reward values indicated on the game screen as well as the animal's beliefs about the other player's strategy. On each trial the monkey will obtain one of four possible reward values, V_t^{coop} if both the animal and the other player deviate, V_t^{straight} if the monkey chooses straight while the other player deviates, V^{safe} if the monkey swerves while the other player goes straight, or V_t^{crash} if both players go straight. Note that the safe and crash values do not change from trial to trial and are fixed at three tokens and zero tokens, respectively. Accordingly, the monkey calculates the expected values of the actions straight and deviate given his belief regarding how likely the other player is to deviate. This likelihood is denoted by $p'_t = p(c_t = \text{yield})$ where c'_t is the opponent's choice on trial t . The formulae for the expected values are given by

$$\begin{aligned} E[V_t^{\text{yield}}] &= p'_t V_t^{\text{coop}} + (1 - p'_t) V^{\text{safe}} \\ E[V_t^{\text{straight}}] &= p'_t V_t^{\text{straight}} + (1 - p'_t) V^{\text{crash}} = p'_t V_t^{\text{straight}} \end{aligned}$$

This prompts the question of how the monkey obtains his beliefs regarding the other player's strategy -- that is, where does p' come from? In our model, the monkey's representation of the other player's strategy takes the form of a logistic regression which maps the characteristics of each trial onto a probability of deviating. The monkey learns this logistic regression using online-updating as he observes the other player's actions. Formally, this is given by

$$p'_t = \text{logit}^{-1}(X_t^T \beta'_{t,s})$$

where $X_t = [1, V_t^{\text{coop}} - V_t^{\text{straight}}]^T$ is a vector of regressors consisting of an intercept and the difference between the cooperative and straight reward values, and $\beta'_{t,s}$ is a vector of regression coefficients. Note that the monkey's belief regarding the other player's strategy differs between the high and ambiguous joystick direction signal conditions, as $\beta'_{t,s}$ differs depending on the signal strength s .

The monkey updates his beliefs about the other player's strategy on each trial using stochastic gradient descent updates given by

$$\beta'_{t,s} = \beta'_{t-1,s} + \eta_{s,t-1} (\mathbf{1}_{\text{yield}}(c'_t - 1) - p'_t) X_{t-1}$$

The logic of this update is very similar to that of reward prediction error (RPE) updating used in RL models. In an RL model, the predicted reward value Q is updated such that it will be closer to the reward received on the previous trial. Analogously, here the regression coefficients are updated such that the prediction of the logistic regression p' will be closer to the outcome observed on the previous trial. The size of the step taken towards the previously observed value is governed by a learning rate, here denoted η .

As in an RL model, the critical quantity for trial-by-trial learning in our strategic learning model is the error term that captures how predictions differed from the true outcome. Here this error is the term $\mathbf{I}_{yield}(c'_i - 1) - p'_i$, where $\mathbf{I}_{yield}(c')$ is the indicator function that returns one if the other player chose deviate and zero if he chose straight. We refer to this quantity as the strategic prediction error (SPE) by analogy to the RPE of RL systems.

Intuitively, beliefs about the other player's strategy on ambiguous joystick direction signal trials may be less affected by trials from the high signal condition, and vice versa. Therefore, the regression coefficients for the high and ambiguous joystick direction signal conditions, $\beta'_{i, high}$ and $\beta'_{i, low}$, are updated differentially depending on which signal condition of the previous trial. Specifically, each set of regression coefficients has different learning rates depending on whether the trial that is being learned from was high or ambiguous joystick direction signal condition, such that $\beta'_{i, high}$ is updated using $\eta_{high, high}$ if the previous trial was high signal condition, and using $\eta_{high, low}$ otherwise. The same is true for $\beta'_{i, low}$ for a total of four different learning rates.

Beliefs about the other player's strategy at the beginning of a session are determined by the initial values $\beta'_{1, s}$, which are fit as free parameters.

b) Sub-model comparisons—In order to determine the level of sophistication of each monkey's decision process, we compared a number of sub-models of the model presented above, as well as the full model. The full model assumes that the monkey ascribes intention to the other player (by desiring to obtain higher juice payouts, and having the ability to choose his own actions) and represents the other player's strategy in the form of a logistic regression and updates it online via the strategic prediction error (SPE). Each sub-model is equivalent to the full model with a subset of those features turned off, which we accomplish by fixing certain parameters at zero. We describe the sub-models in order of (approximately) increasing sophistication, and in the same order that the sub-models are shown, from left to right, in the x-axes of figure 3B.

1. *Naive RL*: The least sophisticated sub-model is a naive RL model in which all β parameters other than β_Q and β_k are fixed at zero. This model estimates the values of the actions deviate and straight based only on reward history and does not incorporate the visual information presented on each trial about the payoffs available.
2. *Logistic*: The next sub-model is a logistic regression on the payoffs available on the current trial. This model does not use RL or SPE and instead chooses based only on the visual presented about the payoffs on each trial. β_Q and all learning

rates η are fixed at zero. Also, the second elements of the parameter vectors $\beta'_{i,s}$ are fixed at zero, which leads to beliefs about the other player's strategy being invariant to payoff condition; i.e. the monkey does not consider that the other player has his own intentionality and cares about the payoff condition.

3. *Combined logistic-RL*: A combined logistic-RL model, equivalent to the second model described above with β_Q not fixed at zero.
4. *SPE-learning*: A model that incorporates SPE learning, but without representing the other player's intentionality. This is equivalent to the full model with the second elements of the parameter vectors $\beta'_{i,s}$ fixed at zero.
5. *Intention*: A 'static' model in which the other player is assumed to have intentionality and cares about obtaining higher payoffs, but there is no SPE learning and so beliefs about the other monkey's strategy do not adjust over time. This is equivalent to the full model with all learning rates η fixed at zero.

Using the expected utility calculations, payoffs and the other monkey's predicted choices, we were able to compute and predict monkey's choices as well as predict the other player's choices (Figure 3C, for an example pair, sub-model 5 (Combined logistic-RL)).

Analysis of electrophysiology data

Single-unit firing rates were isolated using a combination of principle component analysis (PCA), the Template Matching algorithm, and hand-sorting in Offline Sorter (Plexon Inc). All subsequent data analyses were accomplished with custom MATLAB scripts. The peristimulus time histograms (PSTHs) shown are rendered in 1ms steps with Gaussian smoothing of 10ms on both sides. For population PSTHs, firing rates were normalized to the pre-fixation firing rate (200ms time window immediately before the onset of the fixation cue). Using different time windows and an alternative normalizing methods of a) normalizing to whole trial firing rates, and b) z-scoring of firing rates to the whole trial did not significantly change any main results reported. Statistical tests were conducted as two-tailed ANOVAs with multiple comparisons (Tukey's HSD test) unless otherwise specified. Sessions with fewer than 200 trials were completely excluded from analysis based on the calculation that each trial type and outcome would have occurred on average <4 times in that session.

Epoch-based analyses were conducted for 3 distinct time windows: payoff presentation (0-500ms after the onset of the tokens on the screen), post-decision/cars move (0-500ms after the end of the 4s decision period and start of car movement), and juice delivery (250-1250ms after the juice is delivered).

Responses (neuronal firing rate in the epoch of interest as described above) from non-control trials were fit with linear models (LM) and cross validated 20-fold, holding out 1/20 of the trials from each neuron each time. These were compared to the same trial sets where the outcomes were randomly shuffled. All continuous variables (including neural responses) were z-scored by the trials that made up each neuron's data structure. The models were individually fit to each neuron and ANOVAs were used to classify responses. The variables

used in each model are as follows: For payoff presentation, the differences between the straight and cooperative token amount (V_{diff}), the predicted strategy of the other player for the current trial (p'), and the strategy prediction error for the trial immediately prior (SPE1), all of which are continuous variables. For the post-decision/cars move and juice realization epoch, we include categorical the variables: cooperate, joystick direction signal strength (indicating availability of explicit information about intentions), and gaze (considered '1' if the animal makes a fixation for 150ms or longer within the defined 'face'/target boundaries of the other monkey during the 1.5s window from juice delivery, boundaries schematically illustrated in Extended Figure 2), and the continuous variable reward amount, which is the number of tokens the player received in juice. The other player's predicted strategy (p'), a continuous variable, was orthogonalized against 'cooperate' and signal strength to avoid collinearity of the model, and is shown on its own as well as an interaction term with cooperate.

The terms used to categorize outcomes in Figures 5 and 6 are '*cooperate*', where both players moved pushed the cooperation bar and received the V_{coop} payout; '*selfish*', where one player chose straight (the *selfish* player), and the other deviated ('*chicken*'); and the controls, where only one player's car and payoffs were present.

Evaluation of image preferences of mSTS neurons

Passive fixation task—The passive fixation task was presented to monkeys immediately before and/or after the chicken task. During this task, there was only one monkey in the room. The monkey was seated in the same position as the chicken task with the stimulus presentation screen positioned parallel to the ground. The trials were initiated by the monkey aligning gaze with (5 ± 2.5 degrees) a small central red square (20×20 pixels). After maintaining gaze aligned with the fixation point for 250 ms, a sequence of 5 luminance balanced images were presented for 400ms each at central fixation, and the monkey received a small juice reward for maintaining gaze on the fixation point. Images were randomly drawn from a set of 974 images. The inter-trial interval varied randomly between 300-800ms.

We recorded 407 distinct mSTS neurons, of which 227 were recorded in the chicken task. Calculations for the visual response index and image preference index are given below, using the time epochs of baseline activity (200-0ms prior to first image onset while the monkey's gaze was aligned with the central square), and image responsiveness (150-350ms from the image onset at the fovea). Visual responsiveness index is the ratio of the difference in firing rate of the neuron to the first image at the fovea and baseline activity over the sum of that activity.

To calculate the feature preference index, we divided the images shown into four main categories (monkey face / human face / monkey perinea/ objects / and scrambled, which served as a luminance control). We then examined the responses of the neuron to the images in each category and identified the one eliciting maximal response and the one eliciting minimal response (Figure 7A). The feature selectivity index was calculated as the ratio of the difference in the maximum and minimum response to feature categories over the sum of that activity. Results were not significantly different when we combine the two separate face

categories (monkey and human). Data distribution was assumed to be normal but this was not formally tested.

Statistics: The number of animals and the number of neurons recorded from was based on previous experience. No statistical method was used to predetermine sample size. Statistical analyses were conducted in Matlab (2013, 2016a). 20-fold, cross validated linear models were used to determine the influence of variables on the firing rate, and the difference in proportion of neurons between agency and brain areas were determined with 2-way ANOVAs followed by post-hoc tests (Tukey's HSD). For characterization of mSTS neurons in the passive fixation and chicken tasks, two-tailed, unpaired student's t-test were performed. Significance was set at $\alpha = 0.05$ for all statistical tests. Data distribution was assumed to be normal but this was not formally tested.

Further information is available in the Life Sciences Reporting Summary linked to this article.

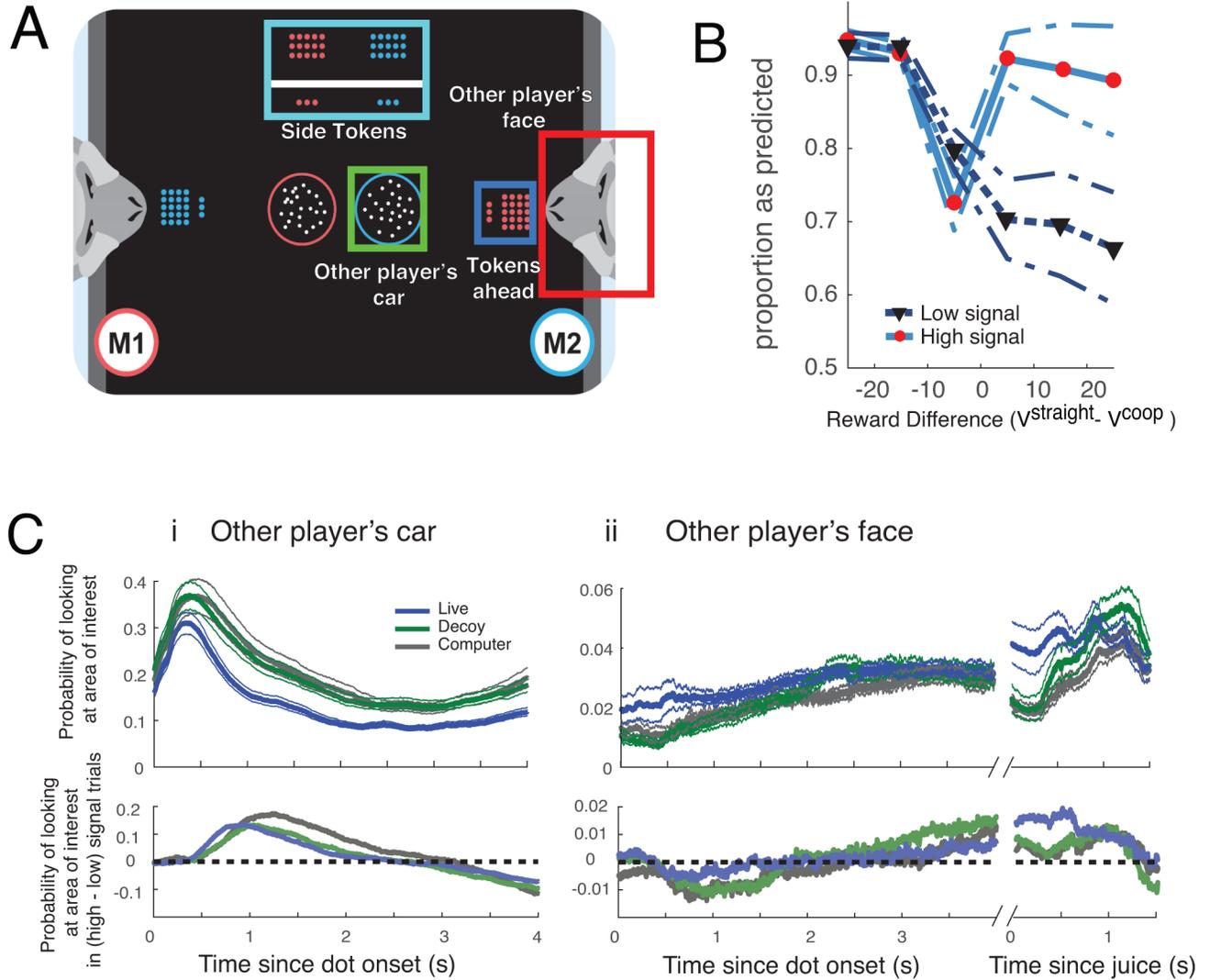
Data availability:

The data forming the basis of this study are available from the corresponding author upon reasonable request.

Code availability:

The custom analysis code used in this study are available from the corresponding author upon reasonable request.

Extended Data



Extended Data Fig. 1. Monitoring of gaze behaviour and trial outcomes

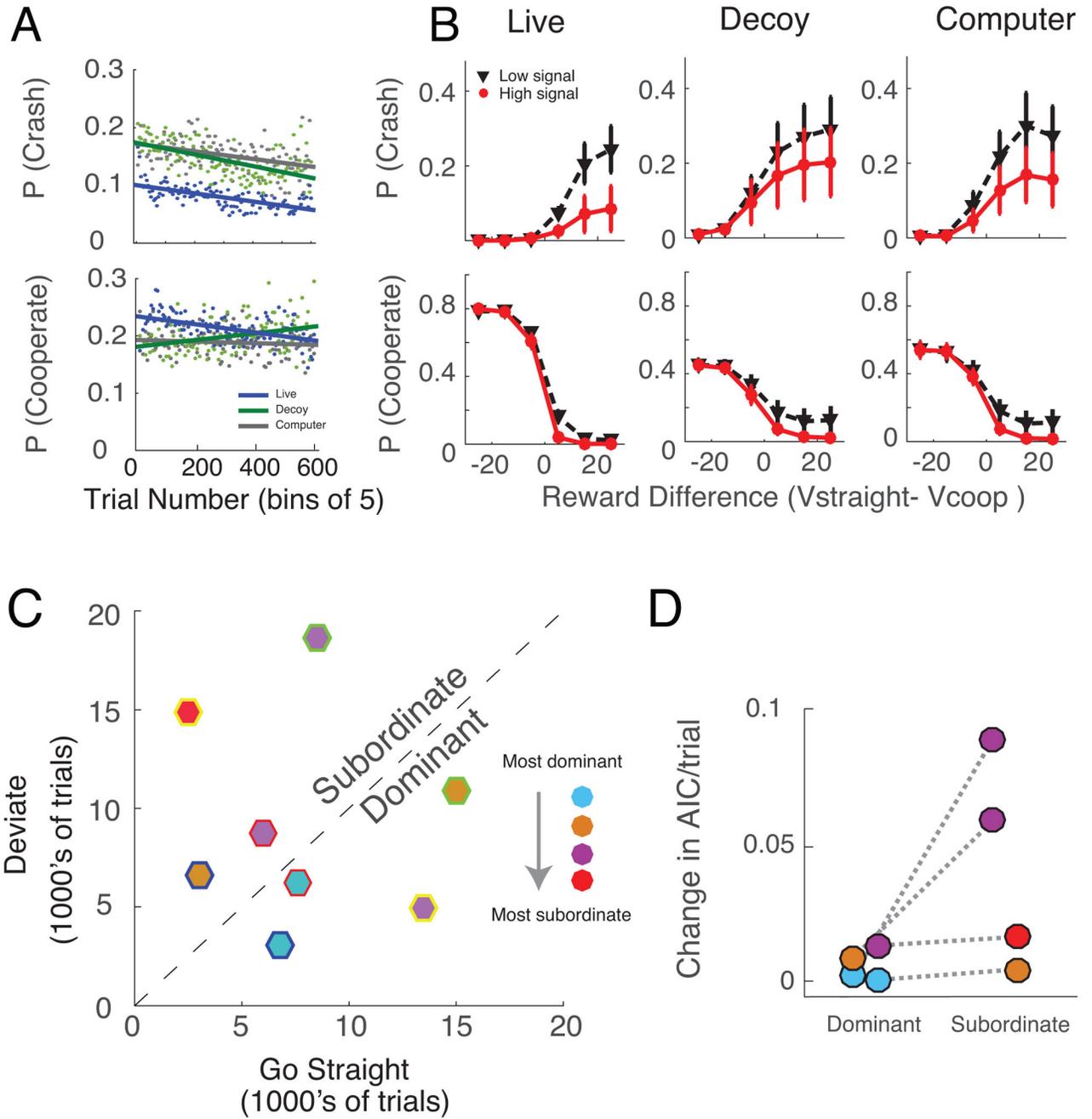
A) ROIs for M1's gaze (red) indicated by rectangles drawn around other monkey's face, other player's car, and the payout tokens ahead and to the side. All other gaze points directed to the screen are labeled 'Screen'.

B) Proportion of trials resulting in mixed strategy Nash equilibrium when $V_{\text{str}} > V_{\text{coop}}$ and pure Nash equilibrium when $V_{\text{coop}} > V_{\text{str}}$. Solid light blue line, high joystick direction signal (90% dot motion coherence) trials; dotted dark blue line, ambiguous joystick direction signal (0% dot motion coherence) trials. Fine lines indicate \pm SEM for 4 live monkey pairs ($n=75,630$ trials).

C) Monkeys look at the most informative stimuli, calibrated by social context.

Top panel: Probability M1 looked towards other player's car (i) or face/face space (computer condition) (ii), aligned to moving dots onset, in live (blue), decoy (green), and computer (black) conditions. Bottom panel: Difference in probability of gaze towards other player's

car (i) or face (ii) on high and ambiguous joystick direction signal (dot motion coherence) trials. All data calculated in 1ms bins.



Extended Data Fig. 2. Choice outcomes over time and between agency conditions, and the effect of dominance relationships on model fits.

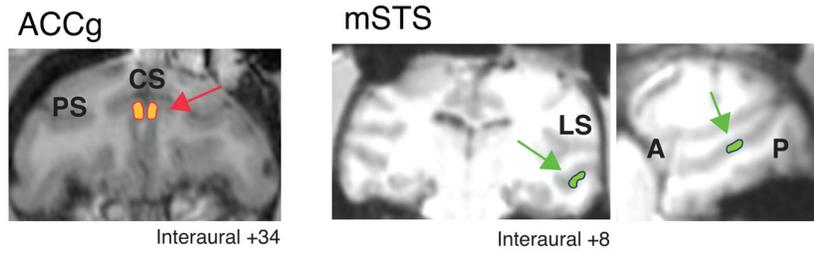
A) Probability of a crash (top panel) or cooperate (lower panel) event over time. Each point represents a bin of 5 trials from all sessions as a function of social context (blue, live; green, decoy; grey, computer. $n = 164,259$ trials).

B) Probability of a crash (top panel) or cooperate (lower panel) event as a function of payout conditions (difference in the number of tokens available straight ahead, V_{str} , and cooperate, V_{coop}). Red circles/solid lines, high joystick direction signal (90% dot motion coherence) trials; black triangles/dashed lines, ambiguous joystick direction signal (0% dot motion coherence) trials.

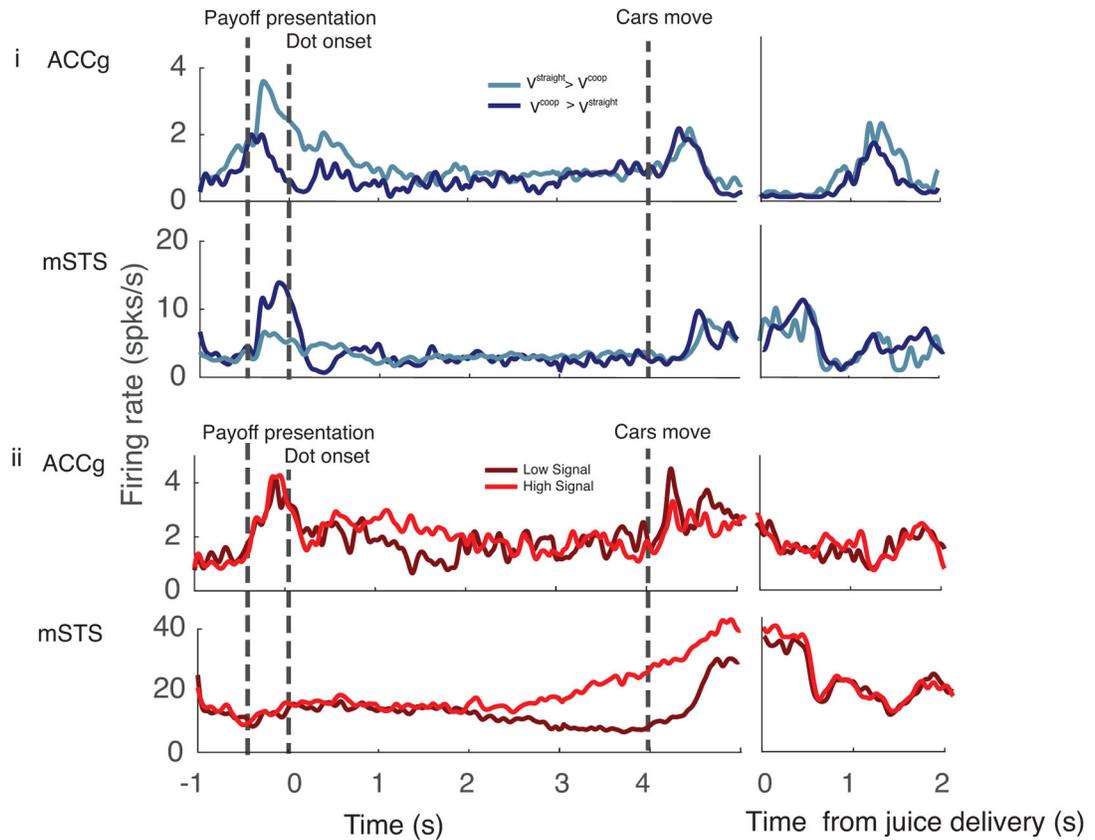
C) Frequency of deviating as a function of frequency of choosing straight for 4 monkey pairs, segregated by social dominance. For a given pair, number of times they chose deviate (y-axis) or straight (x-axis) are plotted. Symbol fill colors indicate monkey identity; symbol outline colors indicate monkey dyads. Note that mid-ranking monkeys (purple and brown) fall on both sides of the diagonal, suggesting that strategy depends more on relative status than identity.

D) Improvements in model fits depend on dominance relationship between monkeys. Models were fit to each monkey's choices playing specific other monkeys. Y-axis shows increase in AIC/trial when *intention* was added to model. Within each monkey pair, one was always dominant to the other, and AIC/trials are segregated by relative status. Grey dotted lines connect monkeys playing each other; colors indicate monkey identity.

A. Location of recordings



B. Four example neurons



Extended Data Fig. 3. Location of neural recordings (ACCg & mSTS) and example neurons.

A) Recording locations in ACCg (orange, coronal section) and mSTS (green, coronal and para-sagittal sections). CS, cingulate sulcus; LS, lateral sulcus; A, anterior; P, posterior. A/P distance from inter-aural 0 indicated for coronal sections.

B) PSTHs for 2 example ACCg neurons and 2 example mSTS neurons in monkeys playing a live monkey, aligned to moving dots onset (left) and juice delivery (right). Payoff token onset occurs 500 ms before dots onset, cars move 4 s, later and juice delivery occurs 900 +/- 100ms after cars move. Top two rows, example neurons sensitive to payouts (light blue: $V_{\text{straight}} > V_{\text{coop}}$; dark blue: $V_{\text{coop}} > V_{\text{straight}}$); bottom two rows, neurons

sensitive to joystick direction signal strength (dot motion coherence); dark red, 0% coherence; light red, 90% coherence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank the members of Duke University's Division of Laboratory Animal Resources for their superb animal care.

Grants:

This work was supported by a grant from SFARI (304935, MLP), R01 MH095894, R01 NS088674, R37 MH109728, and R01 MH108627.

References:

1. Warneken F & Tomasello M The Developmental and Evolutionary Origins of Human Helping and Sharing. in *The Oxford Handbook of Prosocial Behavior* (2015). doi:10.1093/oxfordhb/9780195399813.013.007
2. Efferson C & Fehr E Simple moral code supports cooperation. *Nature* 555, 169–170 (2018).
3. Decety J, Bartal IB-A, Uzefovsky F & Knafo-Noam A Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* 371, 20150077 (2016). [PubMed: 26644596]
4. Ridinger G & McBride M Theory of Mind Ability and Cooperation in the Prisoners Dilemma. (2016).
5. Rumble AC, Van Lange PAM & Parks CD The benefits of empathy: When empathy may sustain cooperation in social dilemmas. *Eur. J. Soc. Psychol* 40, 856–866 (2009).
6. Batson CD & Moran T Empathy-induced altruism in a prisoner's dilemma. *Eur. J. Soc. Psychol* 29, 909–924 (1999).
7. Saxe R & Kanwisher N People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *Neuroimage* 19, 1835–1842 (2003). [PubMed: 12948738]
8. Chang SW, Gariépy J-F & Platt ML Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci* 16, (2013).
9. Chang SWC et al. Neural mechanisms of social decision-making in the primate amygdala. *Proc. Natl. Acad. Sci* 112, 16012–7 (2015). [PubMed: 26668400]
10. Carter RM, Bowling DL, Reeck C & Huettel SA A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science* 337, 109–111 (2012). [PubMed: 22767930]
11. Yoshida W, Seymour B, Friston KJ & Dolan RJ Neural mechanisms of belief inference during cooperative games. *J. Neurosci* 30, 10744–10751 (2010). [PubMed: 20702705]
12. Haroush K & Williams ZM Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates. *Cell* (2015). doi:10.1016/j.cell.2015.01.045
13. Rushworth MF, Mars RB & Sallet J Are there specialized circuits for social cognition and are they unique to humans? *Curr. Opin. Neurobiol* 23, 436–442 (2013). [PubMed: 23290767]
14. Tsao DY, Freiwald W. a, Knutsen TA, Mandeville JB & Tootell RBH Faces and objects in macaque cerebral cortex. *Nat. Neurosci* 6, 989–995 (2003). [PubMed: 12925854]
15. Perrett DI, Rolls ET & Caan W Visual neurones responsive to faces in the monkey temporal cortex. *Exp. brain Res* 47, 329–42 (1982). [PubMed: 7128705]
16. Sliwa J & Freiwald WA A dedicated network for social interaction processing in the primate brain. *Science* (80-.). 356, 745–749 (2017).
17. Sallet J et al. Social network size affects neural circuits in macaques. *Science* 334, 697–700 (2011). [PubMed: 22053054]

18. Baron-Cohen S, Leslie AM & Frith U Does the autistic child have a 'theory of mind'? *Cognition* 21, (1985).
19. Baron-Cohen S, Wheelwright S, Hill J, Raste Y & Plumb I The 'Reading the Mind in the Eyes' Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry Allied Discip* 42, 241–251 (2001).
20. Constantino JN & Gruber C Social responsiveness scale. (2005).
21. Camerer C Behavioral game theory : experiments in strategic interaction. (Russell Sage Foundation, 2003).
22. Lee D & Seo H Neural Basis of Strategic Decision Making. *Trends Neurosci.* 39, 40–48 (2016). [PubMed: 26688301]
23. Maynard Smith J Evolution and the theory of games. (Cambridge University Press, 1982).
24. Camerer CF Progress in Behavioral Game Theory. *The Journal of Economic Perspectives* 11, (1997).
25. Carlsson B & Jönsson KI Differences between the iterated prisoner's dilemma and the chicken game under noisy conditions. in Proceedings of the 2002 ACM symposium on Applied computing - SAC '02 42 (ACM Press, 2002). doi:10.1145/508791.508802
26. Carlsson B & Johansson S An iterated hawk-and-dove game. in 179–192 (Springer Berlin Heidelberg, 1998). doi:10.1007/BFb0055028
27. Crawford MP The cooperative solving of problems by young chimpanzees. (Johns Hopkins Press, 1938).
28. Hare B, Melis AP, Woods V, Hastings S & Wrangham R Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Curr. Biol* 17, 619–23 (2007). [PubMed: 17346970]
29. Mendres KA & M De Waal FB Capuchins do cooperate: the advantage of an intuitive task. *Anim. Behav* 60, 523–529 (2000). [PubMed: 11032655]
30. Drea CM & Carter AN Cooperative problem solving in a social carnivore. *Anim. Behav* 78, 967–977 (2009).
31. Siposova B, Tomasello M & Carpenter M Communicative eye contact signals a commitment to cooperate for young children. *Cognition* 179, 192–201 (2018). [PubMed: 29960903]
32. Osborne MJ & Rubinstein A A course in game theory. (MIT Press, 1994).
33. Samuelson PA Probability, Utility, and the Independence Axiom. *Econometrica* 20, 670 (1952).
34. Camerer C & Ho TH Experience-weighted Attraction Learning in Normal Form Games. *Econometrica* 67, 827–874 (1999).
35. Axelrod R & Hamilton WD The Evolution of Cooperation. *Science* 211, 1390–6 (1981). [PubMed: 7466396]
36. Nowak M & Sigmund K A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 363, (1993).
37. Sutton RS & Barto AG Reinforcement Learning: An Introduction. (MIT Press, 1998).
38. Chance MRA Attention Structure as the Basis of Primate Rank Orders. *Man* 2, 503–518 (1967).
39. Shepherd SV, Deaner RO & Platt ML Social status gates social attention in monkeys. *Curr. Biol* 16, R119–20 (2006). [PubMed: 16488858]
40. Perrett DI et al. Social Signals Analyzed at the Single Cell Level: Someone is Looking at Me, Something Moved! *Int. J. Comp. Psychol* 4, (1990).
41. Tsao DY, Freiwald WA, Tootell RBH & Livingstone MS A cortical region consisting entirely of face-selective cells. *Science* 311, 670–4 (2006). [PubMed: 16456083]
42. Hasselmo ME, Rolls ET & Baylis GC The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behav. Brain Res* 32, 203–218 (1989). [PubMed: 2713076]
43. Parkinson C, Kleinbaum AM & Wheatley T Spontaneous neural encoding of social network position. *Nat. Hum. Behav* 1, 72 (2017).
44. Drea CM & Wallen K Low-status monkeys 'play dumb' when learning in mixed social groups. *Proc. Natl. Acad. Sci. U. S. A* 96, 12965–9 (1999). [PubMed: 10536031]

45. Logothetis NK, Pauls J, Augath M, Trinath T & Oeltermann A Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–7 (2001). [PubMed: 11449264]
46. Singer T et al. Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–9 (2006). [PubMed: 16421576]
47. Apps MAJJ & Sallet J Social Learning in the Medial Prefrontal Cortex. *Trends Cogn. Sci* 21, 151–152 (2017). [PubMed: 28173997]
48. Mars RB, Sallet J, Neubert F-XF-X & Rushworth MFS Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex. *Proc. Natl. Acad. Sci* 110, 10806–10811 (2013). [PubMed: 23754406]
49. Platt ML, Seyfarth RM & Cheney DL Adaptations for social cognition in the primate brain. *Philos. Trans. R. Soc. B Biol. Sci* 371, (2016).
50. Cheney DL & Seyfarth RM *How monkeys see the world : inside the mind of another species.* (University of Chicago Press, 1990).

Reference (Methods only):

51. Chang SWC, Winecoff AA & Platt ML Vicarious reinforcement in rhesus macaques (*macaca mulatta*). *Front. Neurosci* 5, 27 (2011). [PubMed: 21516263]
52. Carpenter B et al. Stan : A Probabilistic Programming Language. *J. Stat. Softw* 76, 1–32 (2017).
53. Akaike H Information Theory and an Extension of the Maximum Likelihood Principle. In Petrov BN & Caski F (Eds). *Proceedings of the Second International Symposium on Information Theory.* 267–281 (1973)
54. Wagenmakers E-J & Farrell S AIC model selection using Akaike weights AIC model selection using Akaike weights. *Psychon. Bull. Rev* 11(1), 192–196 (2004). [PubMed: 15117008]

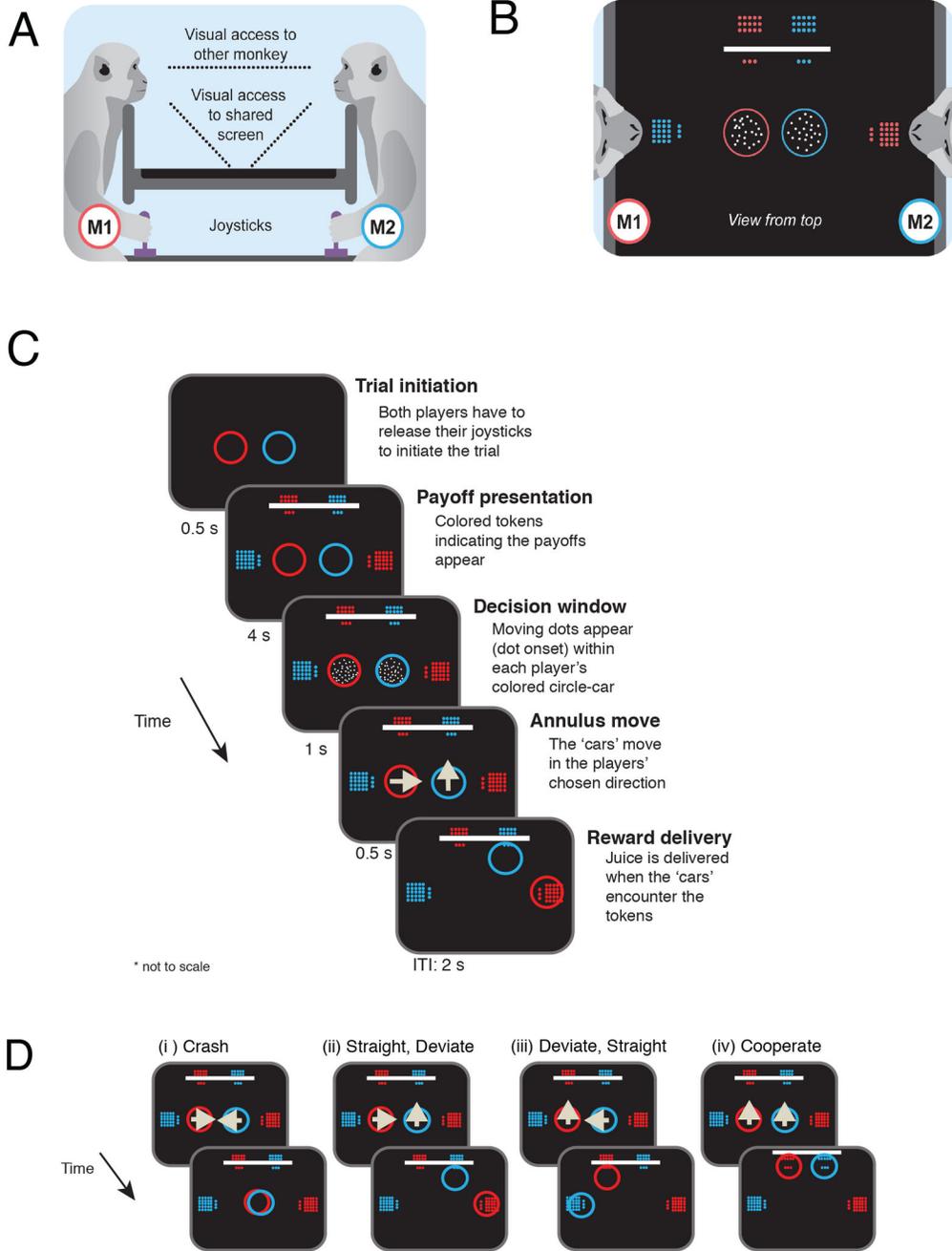


Figure 1: Modified “chicken” game with the option to cooperate.

A) General setup. Two male monkeys (M1 & M2) sat opposite each other across a shared horizontal screen. Each monkey had visual access to the other monkey and the screen, but could not see the bodies, hands or joysticks of either monkey.

B) Overhead view of the setup. M1 (red) controlled the red annulus (hereafter “car”) with a joystick. He could choose either the 23 red tokens straight ahead, behind M2’s (blue) car, or ‘deviate’ and choose the 18 tokens on the left. If he chose the latter, he received only 3 tokens if M2 chose straight, but he would receive all 18 if M2 also chose the same option and they “pushed” the coordination bar together, displacing it and allowing access to the

tokens behind it. If both chose straight the cars collided and neither monkey received reward. Each token = 20ul of juice. For more details see Supplementary Information.

C) Task sequence. After two cars appeared on screen, both monkeys released their joysticks to begin the trial. Colored tokens denoting available reward amounts appeared, followed by onset of moving dots within each car 500ms later. On a randomly selected 50% of trials, dot motion was 90% correlated with joystick direction; on the remaining trials dot motion was random. The moving dots stayed illuminated for 4s, while monkeys moved their joysticks to point toward one of the two token arrays. The monkeys were allowed to move the joysticks freely as they deliberated. To commit to a choice, monkeys held the joystick in a single direction for 500ms, at which point the moving dots changed from white to colored. After commitment, further movement of the joystick had no impact on task events or dot movement. At the end of the 4s decision period, the cars translated in the chosen direction, and juice was delivered after the cars touched the tokens, approximately 1s later.

D) Example choice outcomes. Each monkey could choose straight or deviate. (i) Both go straight and crash, neither receives juice. ii (iii) M1 (M2) goes straight, receiving 28 tokens and M2 (M1) deviates, receiving 3 tokens or the 'chicken reward.' (iv) Both deviate/cooperate and push the bar, releasing the 10 tokens behind it for a total of 13 tokens each. Note that the white bar pushes back to release the tokens behind the bar in the (iv) cooperate outcome but in none of the others.

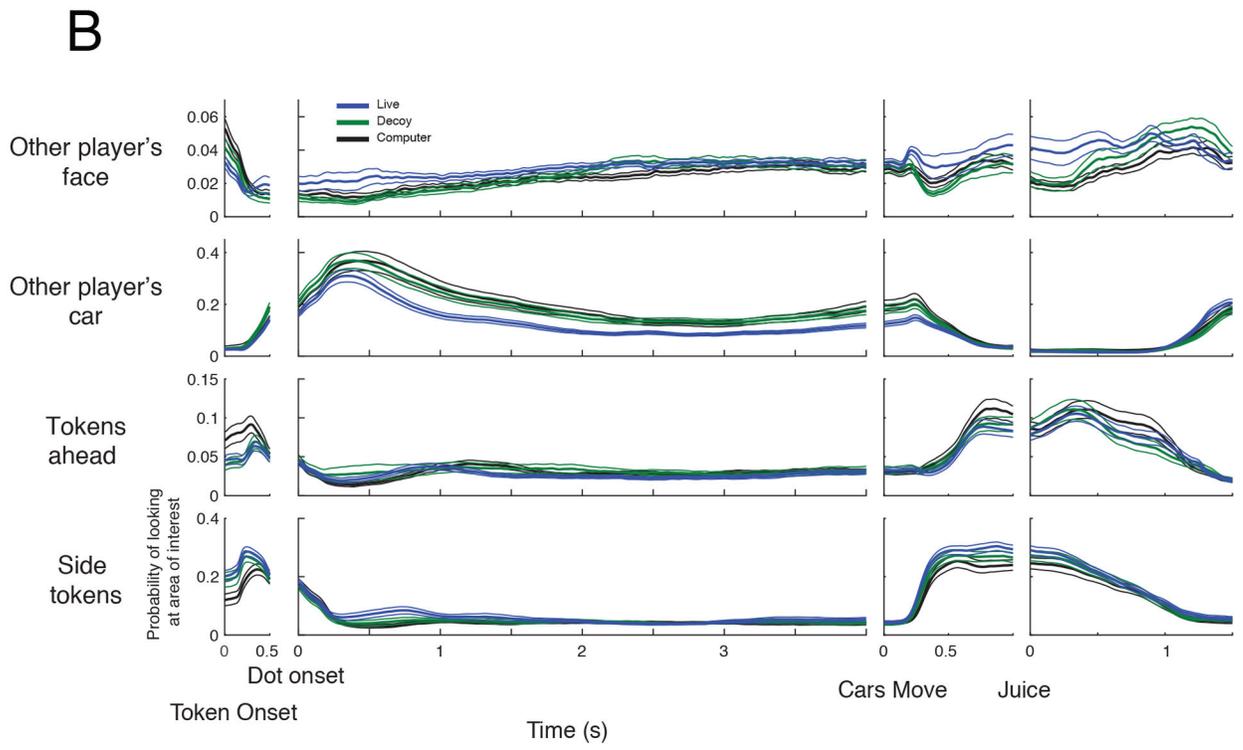
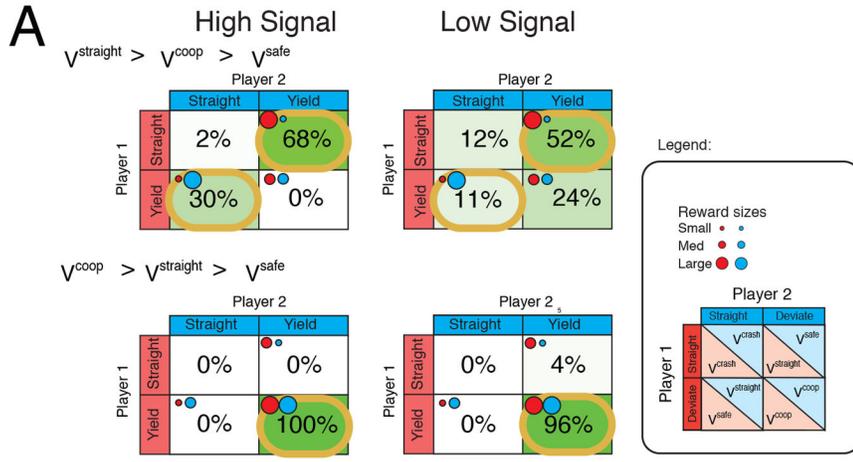


Figure 2: Monkeys understand the task and discriminate social context.
 A) Outcomes for monkey B and monkey C across all sessions playing together (n=29,147 trials) plotted in a payout matrix. Top row, outcomes when payout for straight (V_{str}) was larger than payout for cooperation (V_{coop}); bottom row, outcomes when payout for cooperation was larger than payout for straight. V_{safe} was a small, constant reward of 3 tokens. Right column, high joystick direction signal trials (90% coherent dot motion); left column, low joystick direction signal trials (0% coherent dot motion). Yellow bordered cells indicate mixed strategy Nash equilibrium when $V_{\text{str}} > V_{\text{coop}}$ and pure Nash equilibrium when $V_{\text{coop}} > V_{\text{str}}$. Payoff matrix. M1 (red) = row player; M2 (blue) = column player.

B) Probability of gaze on regions of interest aligned to token onset, dot onset (decision period), cars move (outcomes revealed), and juice delivery (reward) when monkey played against a live monkey (blue), a decoy monkey (green) or computer (black). First row: other player's face/face space (computer condition). Second row: other player's car. Third row: tokens straight ahead. Bottom row: cooperative tokens.

*curves do not sum to 1 as we only show gaze behaviour to specific areas of interests (refer Extended Data Figure 2A)

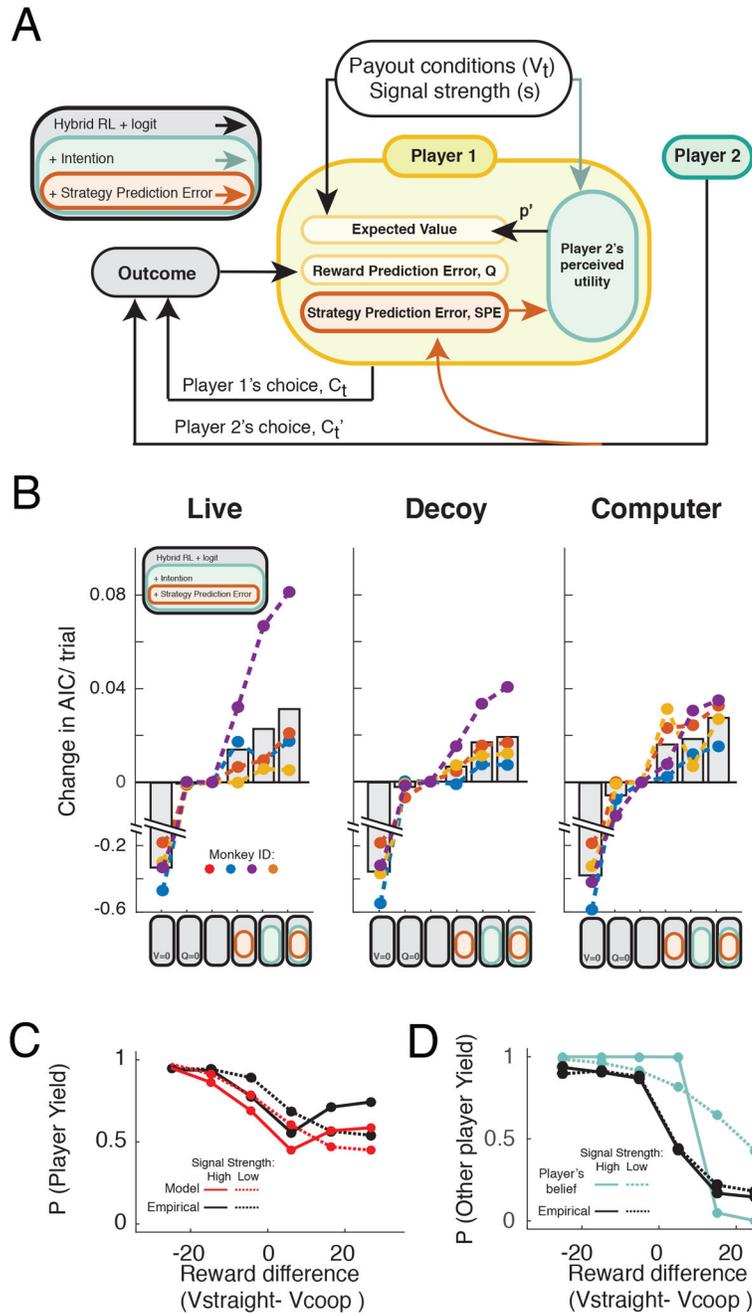


Figure 3: Monkey choice behavior is best explained by a sophisticated, recursive model including information about payoffs for self and other, other player's predicted strategy, and strategy prediction error.

A) Schematic model of the decision-making process for a monkey player. Models of increasing sophistication were iteratively fit to choices and evaluated by AIC. Yellow box represents the player monkey and all variables hypothesized to be calculated internally; everything outside box is externally observable to both players. Payout conditions (V_t) and joystick direction signal strength (dot motion coherence, s) are independent variables selected randomly across trials. The monkey and the other player make choices (C_t and C_t'

respectively), resulting in outcomes observed by both players. In the logit model, we assume the only factor contributing to the monkey's choice is payout (V_t). In the hybrid RL-logit model, we included calculation of an expected value for each option from the payout conditions to include information about prior outcomes via reinforcement learning. We also included an autocorrelation factor (not shown). The intention model (green boxes and arrows) assumes the monkey player (yellow) forms a belief about the other player's propensity to yield (p), which itself is based on available payouts. Prediction of the other player's strategy (p') can be improved by incorporating a strategic prediction error (SPE, red box and arrows), which allows the monkey to update his prediction of the other player's actions based on the difference between his predicted and observed choices rather than trial outcomes alone.

B) Difference in AIC per trial compared to hybrid RL-logit model in the live, decoy, and computer conditions. Grey bars, average AIC/trial decrease compared to the *combined logistic-RL* model (third bar location on the x-axis, empty grey box). First bar shows decrease in AIC when payout conditions (V_t) were not included in model ($V=0$, *naïve RL*). Second bar shows AIC when reward prediction error (Q) was not included ($Q=0$, '*logistic*'). 4th-6th bars all show model improvements by addition of the other player's utility function/intentionality ('*intention*', green box) and strategy prediction error learning ('*SPE-learning*', red boxes) to the model. The full model is represented by the 6th bar. Model fit results for individual monkeys denoted by colored circles.

C) Comparison of intention model predictions and observed choices for an example pair of monkeys. Red lines show predicted choices for monkey C given payoffs and monkey B's predicted choices. Black lines show observed choices for monkey C when playing monkey B. Solid lines, high joystick direction signal (90% dot motion coherence) trials; dotted lines, ambiguous joystick direction signal (0% dot motion coherence) trials.

D) Comparison of intention model predictions (green lines) of monkey B's choices from the perspective of monkey C (player's belief) compared to monkey B's observed choices (black). Solid lines, high joystick direction signal (90% dot motion coherence) trials; dotted lines, ambiguous joystick direction signal (0% dot motion coherence) trials.

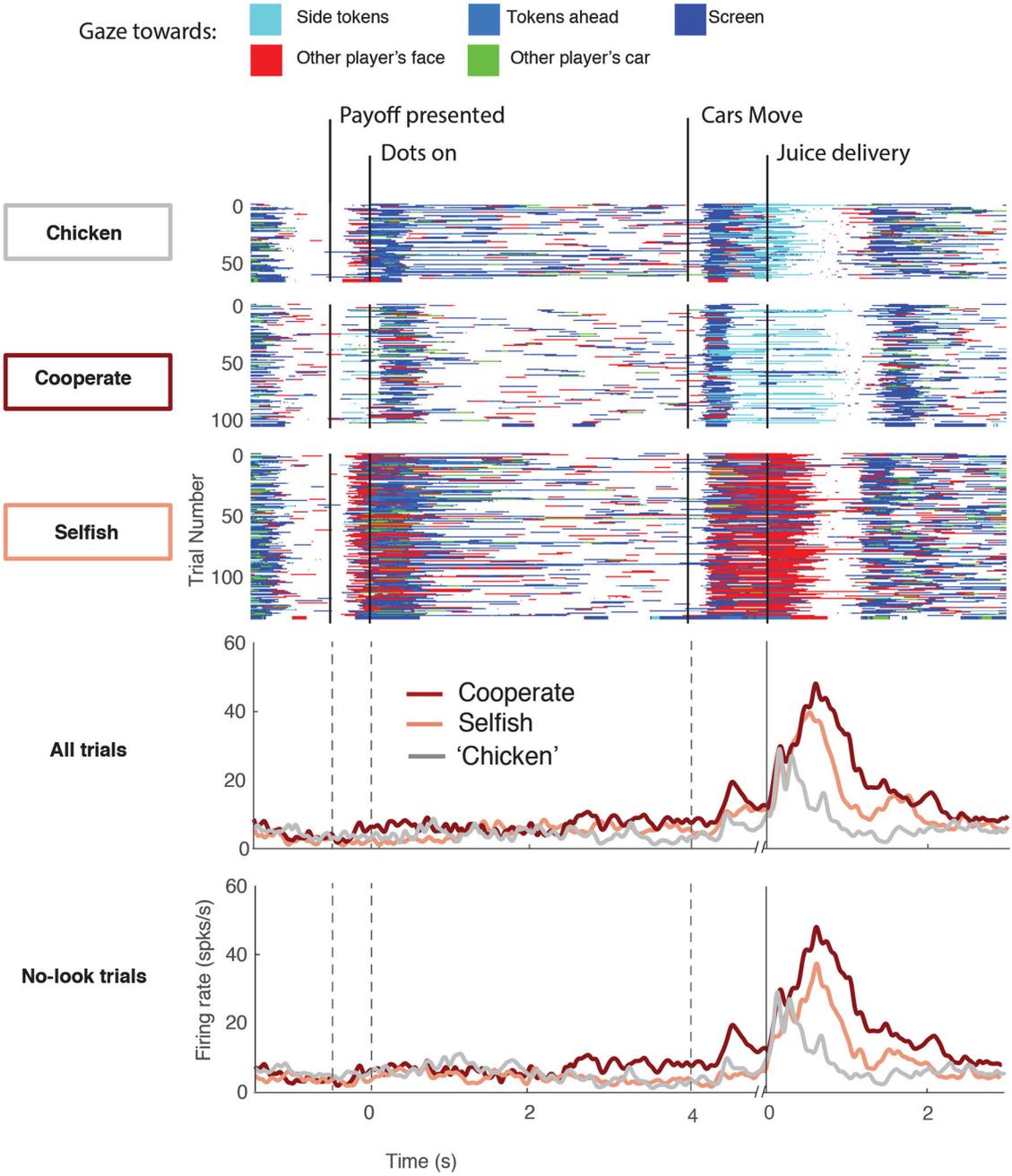


Figure 4: Example mSTS neuron signaling cooperative reward independent of gaze direction and reward size.

Top: Gaze on regions of interest (ROIs) for monkey B on trials with ‘chicken,’ cooperate, and selfish outcomes, aligned to target onset, dots onset, cars move, and juice delivery. Each row plots a rastergram of gaze on ROIs for a single trial (n=299 trials).

Rectangular boxes indicate other player’s face, other player’s car, and tokens ahead and to the side, while gaze elsewhere on screen labeled ‘Screen’ (See Extended figure 1A).

Bottom: Peri-stimulus time histogram (PSTH) for an example mSTS neuron on trials the monkey player chose cooperative, selfish, and ‘chicken’ rewards. Upper panel plots firing

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

rates on all trials; lower panel plots firing rates only on trials where the monkey did not look at the other player during juice delivery. Number of trials for each outcome (all trials / no look trials) : cooperate (108/ 104), selfish (130/ 30), 'chicken'(61/ 55)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

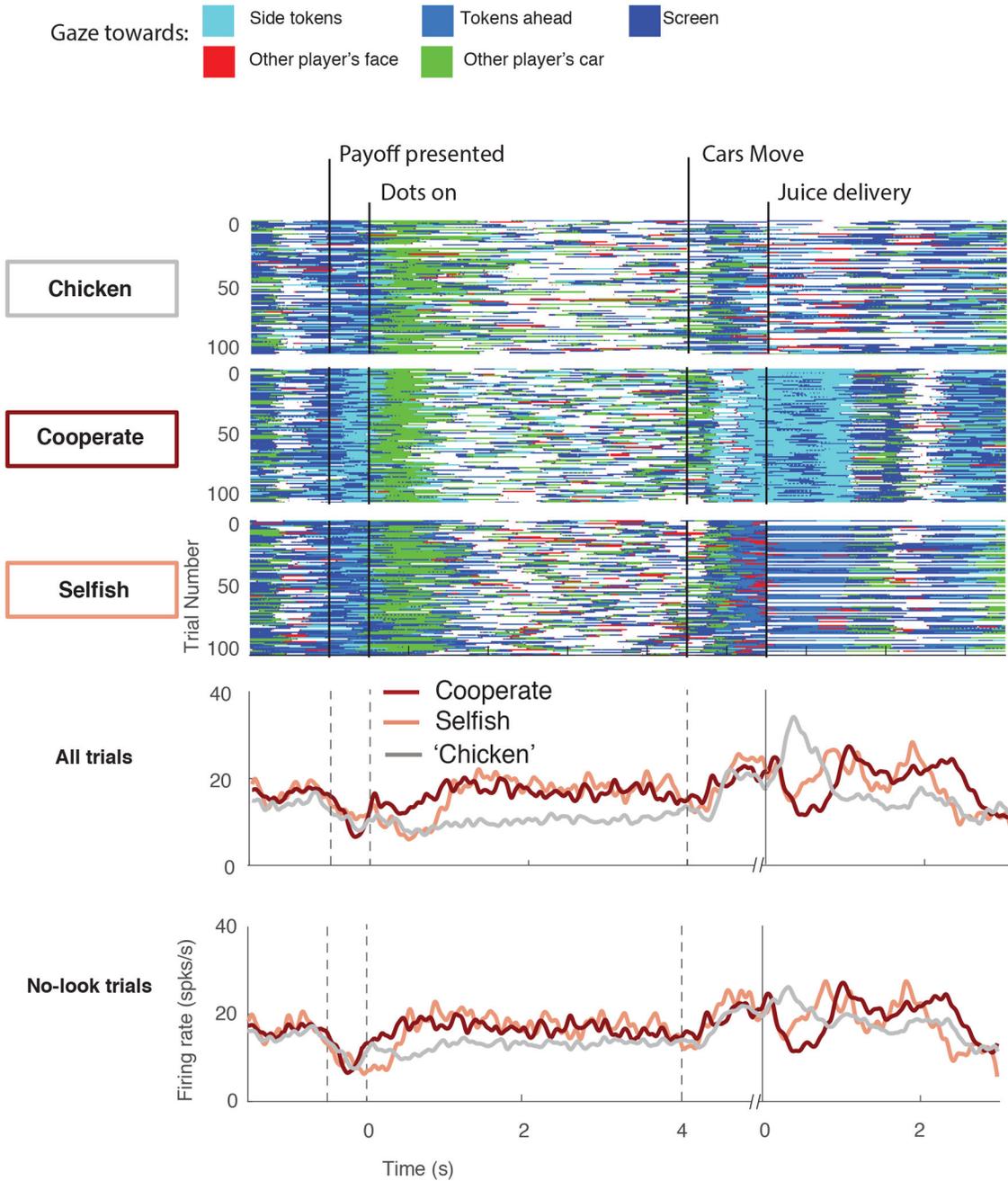


Figure 5: Example ACCg neuron sensitive to cooperative reward independent of gaze direction and reward size.

Top: Gaze on regions of interest (ROIs) for monkey B on trials with 'chicken,' cooperate, and selfish outcomes, aligned to target onset, dots onset, cars move, and juice delivery. Each row plots a rastergram of gaze on ROIs for a single trial (100 trials for each condition is shown, a subset of the 609 trials in the session).

Rectangular boxes indicate other player's face, other player's car, and tokens ahead and to the side, while gaze elsewhere on screen labeled 'Screen' (See Extended figure 2A).

Bottom: Peri-stimulus time histogram (PSTH) for an example mSTS neuron on trials the monkey player chose cooperative, selfish, and 'chicken' rewards (n=609). Upper panel plots firing rates on all trials; lower panel plots firing rates only on trials where the monkey did not look at the other player during juice delivery. Number of trials for each outcome (all trials / no look trials) : cooperate (202/ 195), selfish (118/ 92), 'chicken'(289/ 213)

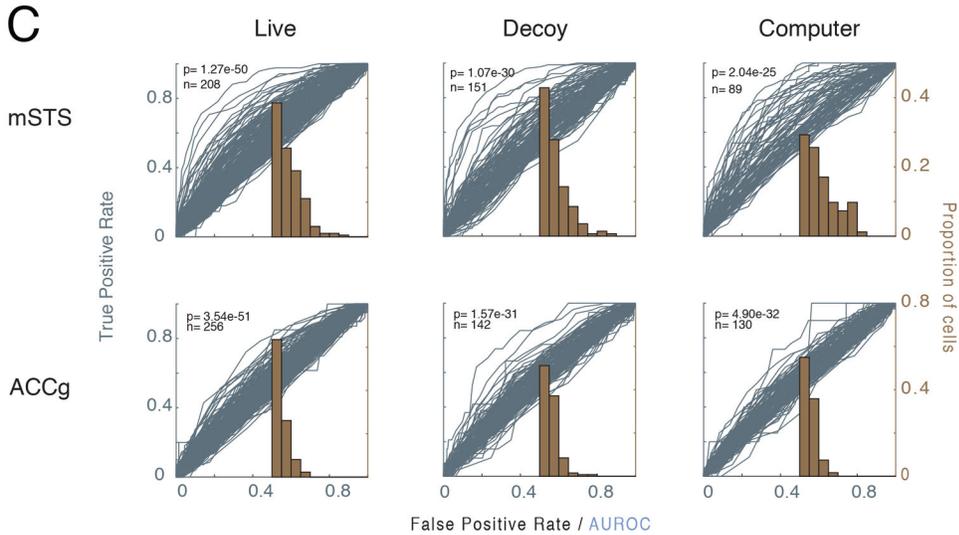
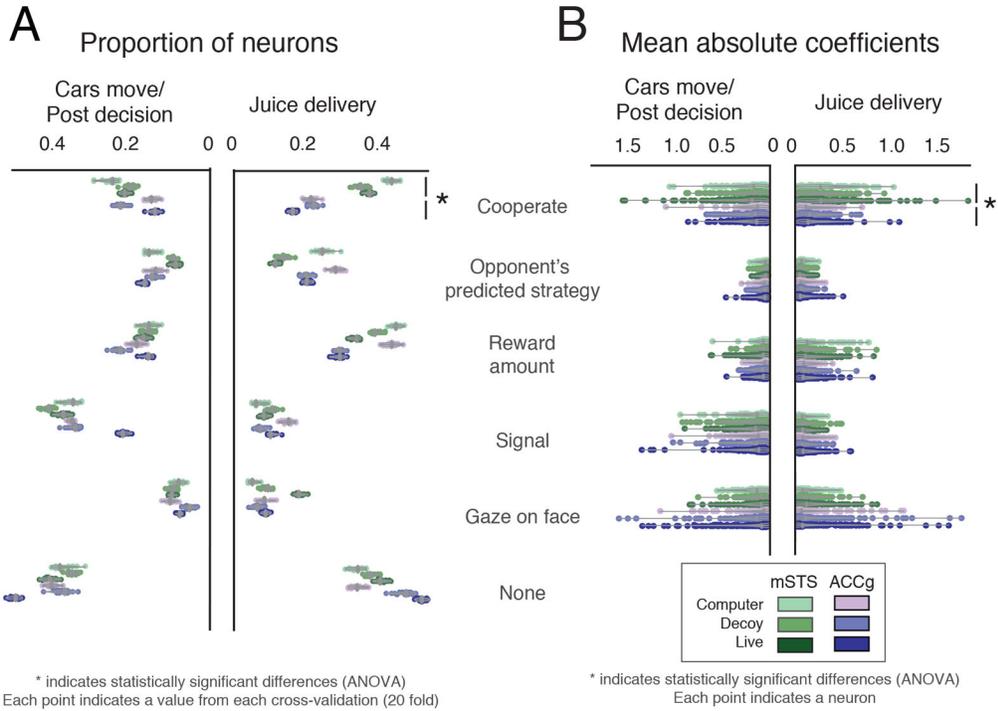


Figure 6: Neurons in ACCg and mSTS signal strategic information.

A) Proportion of neurons in ACCg and mSTS in live, decoy, and computer conditions showing significant firing rate modulations based on 20-fold cross validated LMs (leaving out 1/20 of the trials per neuron each time) as a function of joystick direction signal strength (dot motion coherence), reward size, gaze towards other player’s face, other player’s predicted strategy (p^*) and cooperation during the period 0-500ms after cars began translating and 250-1250ms after juice delivery (grey shading in figures 4A & B). * indicate statistically significant differences between mSTS and ACCg (2-way ANOVA) followed by

Tukey-HSD. For details see Supplementary Table 1. Data exclude single player control trials.

B) Mean absolute LM coefficients for all neurons.

C) Performance metrics for decoding whether monkeys had cooperated based on firing rates during the juice delivery epoch in the live, decoy, and computer conditions. Top row, mSTS; bottom row, ACCg. Grey lines indicate ROC (receiver operating characteristic) curves for each neuron, while histograms show the area under the ROC curve for each neuron (axis to the right, showing proportions of cells in the condition). In all cases, population decoding performed better than chance ($AUROC > 0.5$, $p \ll 10^{-20}$). n denotes number of neurons in each condition.

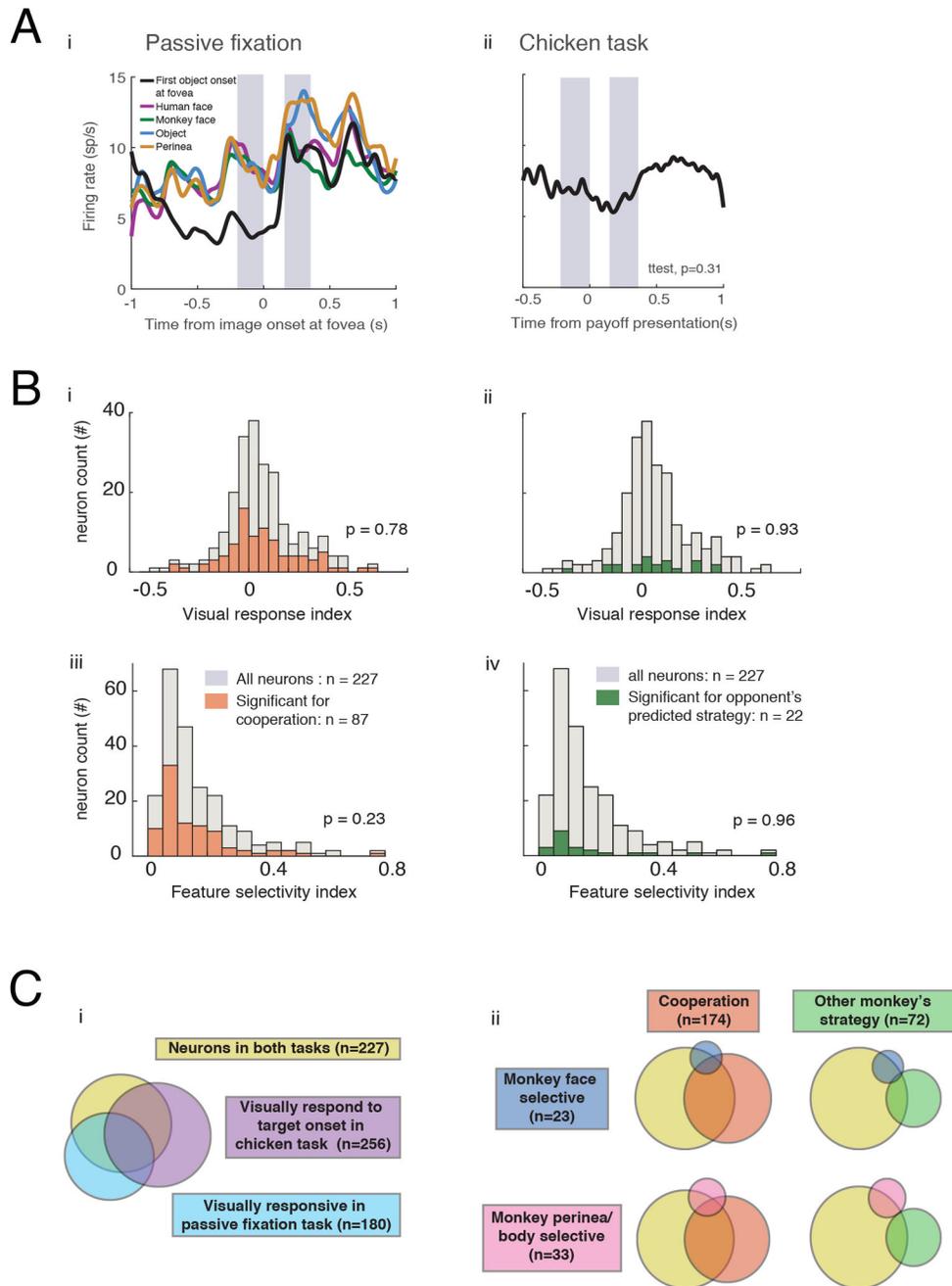


Figure 7: Strategic signaling by mSTS neurons is independent of visual responsiveness and feature selectivity.

A) (i) PSTH for an example mSTS neuron (same neuron as figure 4) recorded during fixation while luminance-balanced images were presented centrally for 400ms. Firing rates aligned to image onset (purple, human face; green, monkey face; blue, object; brown, monkey perineum). Black line plots firing rates aligned to onset of the first of five images presented on each passive fixation trial. (ii) PSTH for the same example mSTS neuron recorded during the chicken task aligned to token onset while the monkey maintained gaze centrally. Firing rates 200ms prior to payoff presentation did not differ from firing rates

150-350ms after payoff presentation ($p=0.31$, two-tailed unpaired t-test). For population, see Supplementary Table 3.

B) Population histograms plotting visual response index (i & ii) and feature selectivity index (iii & iv) for mSTS neurons recorded during both the passive fixation task and the chicken game. Colored bars indicate neurons showing significant effects of cooperation (i & iii) and other player's predicted strategy (ii & iv). Neurons signaling strategic information were not distinct from the overall population in terms of visual responsiveness or feature selectivity ($p = (i) 0.78, (ii) 0.93, (iii) 0.23, (iv) 0.96$; two-tailed unpaired t-test).

C) Venn diagrams showing overlap in visual and strategic signaling by mSTS neurons. Two-hundred twenty-seven mSTS neurons were studied in both the chicken game ($n=448$) and passive fixation task ($n=407$). (i) Visually responsive neurons in the chicken game were determined by comparing firing rates 200ms prior to payoff presentation to firing rates 150-350ms after payoff presentation (purple); visually responsive neurons in the passive fixation task were identified by comparing firing rates 200ms prior to the onset of the first image to firing rates 150-350ms after first image onset (purple). The intersection of the circles shows the fractions of neurons that were members of two or more feature sets. (ii) Upper row: subpopulation of mSTS neurons selective for monkey faces overlapped slightly with subpopulations that signaled cooperation (salmon) or the other player's predicted strategy (green). Bottom row: subpopulation of mSTS neurons selective for monkey perineal bodies overlapped slightly with subpopulation that signaled cooperation (salmon) and not at all with the subpopulation that signaled the other player's predicted strategy (green).