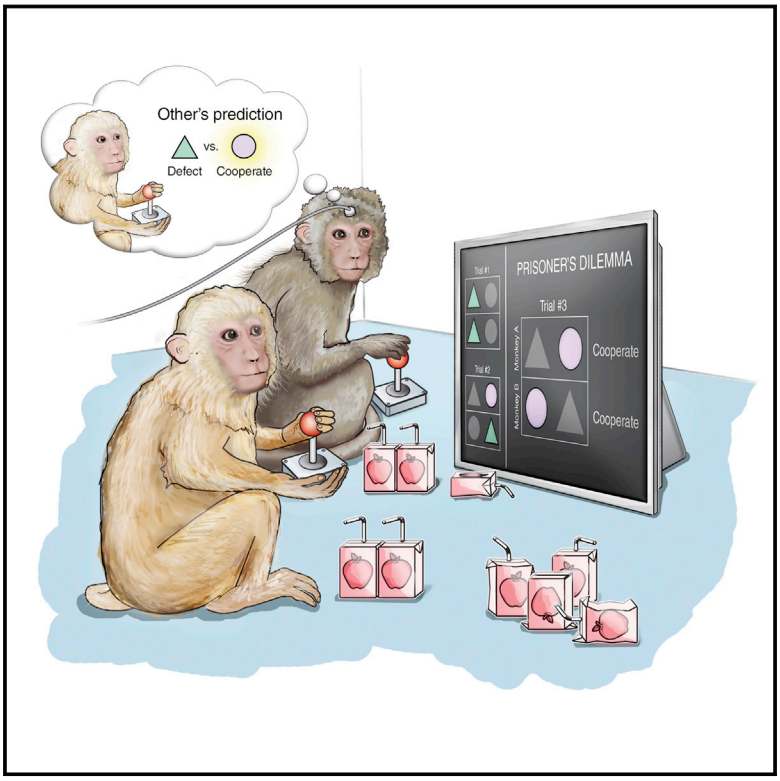
Article

Cell

Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates

Graphical Abstract



Authors

Keren Haroush, Ziv M. Williams

Correspondence

haroush.keren@mgh.harvard.edu (K.H.),
zwilliams@mgh.harvard.edu (Z.M.W.)

In Brief

Cells in the cingulate cortex of primates are able to predict the unknown intentions or state of mind of other individuals and are critically important in enacting cooperative social behavior, a framework that might be relevant for understanding interpersonal, economic, and political decision-making processes in humans.

Highlights

- Cingulate neurons predict another agent's unknown decisions during social interaction
- Other-predictive neurons are sensitive to social context, but not to expected reward
- Distinct cingulate neurons encode the individual's own decisions to cooperate or defect
- Disrupting cingulate activity selectively inhibits mutually beneficial interactions

Haroush & Williams, 2015, Cell *160*, 1–13
March 12, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2015.01.045

CellPress

Please cite this article in press as: Haroush and Williams, Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates, Cell (2015), http://dx.doi.org/10.1016/j.cell.2015.01.045

Cell

Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates

Keren Haroush[1,2,*] and Ziv M. Williams[1,2,*]
[1]Harvard-MIT Health Sciences and Technology, Harvard Medical School, Boston, MA 02114, USA
[2]Department of Neurosurgery, MGH-HMS Center for Nervous System Repair, Harvard Medical School, Boston, MA 02114, USA
*Correspondence: haroush.keren@mgh.harvard.edu (K.H.), zwilliams@mgh.harvard.edu (Z.M.W.)
http://dx.doi.org/10.1016/j.cell.2015.01.045

SUMMARY

A cornerstone of successful social interchange is the ability to anticipate each other's intentions or actions. While generating these internal predictions is essential for constructive social behavior, their single neuronal basis and causal underpinnings are unknown. Here, we discover specific neurons in the primate dorsal anterior cingulate that selectively predict an opponent's yet unknown decision to invest in their common good or defect and distinct neurons that encode the monkey's own current decision based on prior outcomes. Mixed population predictions of the other was remarkably near optimal compared to behavioral decoders. Moreover, disrupting cingulate activity selectively biased mutually beneficial interactions between the monkeys but, surprisingly, had no influence on their decisions when no net-positive outcome was possible. These findings identify a group of other-predictive neurons in the primate anterior cingulate essential for enacting cooperative interactions and may pave a way toward the targeted treatment of social behavioral disorders.

INTRODUCTION

Social interactions are unique from other behaviors in that they inherently require individuals to anticipate each other's unknown intentions and actions. Accordingly, individuals need to consider not only how their decisions affect their own personal outcomes but also how they may affect the outcomes of other individuals in a group and how these individuals may consequently respond. Such interactions, therefore, are not simply governed by the learned sensorimotor contingencies between action and outcome but are rather based on the ability to predict the unknown intentions or "state of mind" of others.

Whether and what neurons encode another's unknown actions and what role these signals play during joint decisions, made independently by two interacting individuals, remain unknown. Prior studies have demonstrated that frontal canonical cells, termed mirror neurons, encode another's known, observable actions, as well as actions performed by the individual himself (di Pellegrino et al., 1992; Rizzolatti and Sinigaglia, 2010). More recently, neurons have been similarly found to encode another's observed receipt of reward (Azzi et al., 2012; Chang et al., 2013; Hosokawa and Watanabe, 2012), as well as monitoring of other's errors (Yoshida et al., 2012, see Discussion). These findings have therefore provided a critical understanding of how another's known and observable actions may be represented at the neuronal level. However, they are distinct from those that may represent another's imminent decisions or intentions, which are fundamentally unobservable and unknown. While cells that predict another's unobservable intended actions have been widely hypothesized, and are a cornerstone of many theories on animal social behavior (Frith and Frith, 1999; Gallese and Goldman, 1998; Rilling et al., 2004; Sanfey et al., 2006; Vogeley et al., 2001), their existence has never been demonstrated.

A second unresolved question is how putative neural signals related to self and other's decisions may affect achieving mutual goals. Mutually beneficial interactions are ubiquitous among social animals (Bshary et al., 2008; Clutton-Brock, 2009; de Waal, 2000; Stephens et al., 2002; Warneken and Tomasello, 2006) and are cardinal to our understanding of socially-guided decisions. While competitive interactions, which allow an individual to profit at the expense of the other, have been previously investigated (Donahue et al., 2013; Hosokawa and Watanabe, 2012; Lee et al., 2005; Seo et al., 2014), the single-neuronal basis of mutually beneficial interactions, favorable to both individuals, have not been explored.

Finally, whereas certain areas may harbor signals that encode elements of social decision-making (Abe and Lee, 2011; Apps et al., 2012; Apps and Ramnani, 2014; Azzi et al., 2012; Behrens et al., 2008; Carter et al., 2012; Chang et al., 2013; Delgado et al., 2005; Donahue et al., 2013; Hampton et al., 2008; Lee et al., 2005; Rilling et al., 2002; Rudebeck et al., 2006; Sanfey et al., 2003; Tomlin et al., 2006; Yoshida et al., 2012), it has not yet been determined what causal contribution neurons in these areas may play in modulating mutual decisions.

A formal framework for studying mutually beneficial joint decisions is by the iterated prisoner's-dilemma (iPD) game (Clutton-Brock, 2009; Rilling et al., 2002; Stephens et al., 2002). This task incorporates two crucial properties: one is that the outcome is contingent upon the mutual concurrent decisions of both individuals, and therefore no one decision guarantees an individual's outcome, and the other is that both decisions can be either concordant or discordant (Camerer, 2003). Therefore, the key to succeeding in the game relies on one's ability to anticipate the other's concurrent, yet unknown intentions. Moreover, this dissociation of self and other decisions, concordant and discordant interactions, and the dissociation between one's decision

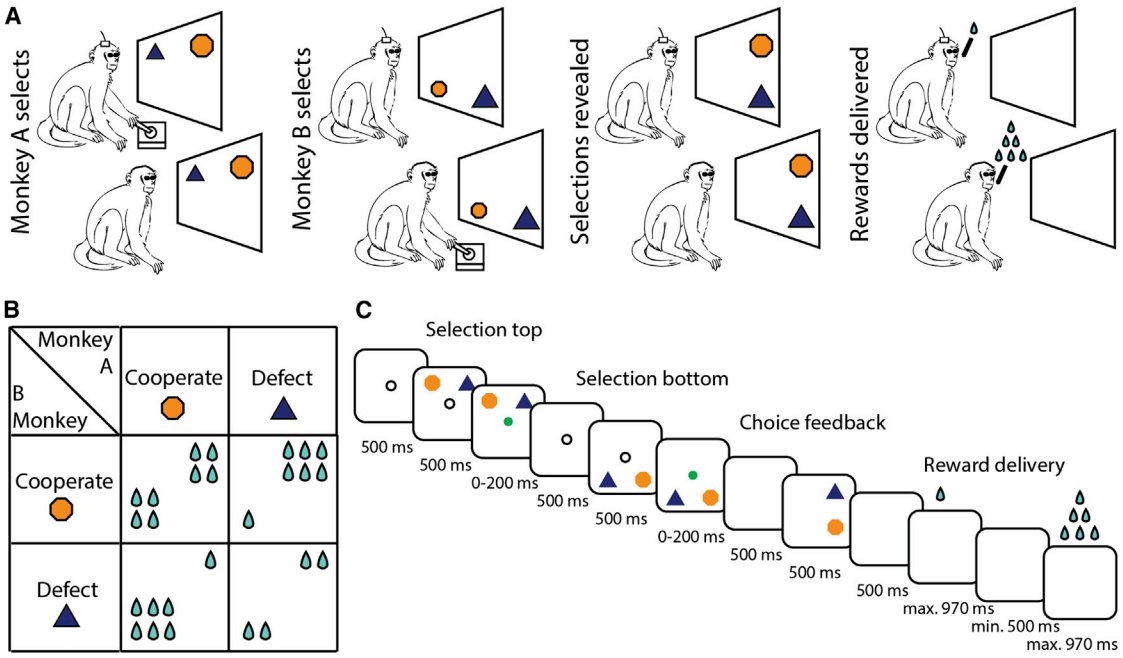


Figure 1. Task Design

(A) Experimental set-up. The monkeys sat side-by-side, facing a screen. On each trial, they covertly chose, in succession, to cooperate (orange hexagon) or defect (blue triangle). Following delay, both choices were revealed on screen and reward was delivered.

(B) Payoff matrix. Reward outcome for all possible choice combinations. Cooperation and defection were defined operationally by whether mutual benefit or loss is incurred.

(C) Trial timeline. The order in which the monkeys made their selections was randomized on each trial.

and reward, allows one to identify neuronal signals within the population that specifically encode another's yet unknown decisions and importantly dissociate them from those that reflect one's own planned decision and expected reward.

Here, we used a joint-decision paradigm to study mutual decisions in primates and provide evidence of neurons that predict another agent's intentions and modes of cooperation. We specifically focused on the dorsal region of the anterior cingulate cortex (dACC) because of its broad connectivity with frontal and temporal-parietal areas known to be involved in interactive behavior (Behrens et al., 2009; Paus, 2001) as well as its role in encoding social interest in other individuals based on functional imaging (Behrens et al., 2008) and ablative studies (Rudebeck et al., 2006). We find that many dACC neurons encoded the monkey's own decision to cooperate. Furthermore, a substantial and largely distinct group of neurons encoded the opponent monkey's decisions when they were yet unknown. These other-predictive neurons were uniquely sensitive to social context compared to other population cells and encoded no information about the monkey's own decisions or expected reward. At the population-level, dACC neurons reliably predicted the other's decisions with accuracy that remarkably approached those of behavioral decoders when based on prior selections. Finally, transient disruption of dACC activity directly and specifically inhibited mutually beneficial interactions based on prior decisions, but did not affect other decisions based on receipt of reward.

These findings together provide direct examination of how individual neurons represent another's unknown intentions or covert "state of mind," demonstrate the distinct encoding of other decisions from self-decisions and reward, ascertain the distinct roles that self- and other-encoding cells play in enacting joint decisions between simultaneously interacting animals, and demonstrate a causal link between cingulate activity and the specific enactment of mutually beneficial decisions.

RESULTS

Increased Cooperation following Mutual Cooperation

Four pairs of adult male Rhesus monkeys (*Macaca mulatta*) performed an iPD game whereby each animal chose on each trial between two response options over multiple successive trials (Figure 1A). The choice terms, cooperation and defection, were derived from iPD literature (Camerer, 2003). These were defined operationally by the payoff matrix illustrated in Figure 1B and are not referred to here in an anthropomorphic way. If both animals selected cooperation, both received the highest mutual reward whereas if one of the animals defected, that animal received the highest individual reward. The lynchpin of this game, however, was that if neither monkey cooperated, they would both receive a lower reward than if they both chose to cooperate. Accordingly, each individual decision could result in either high or low reward depending on the other's choice, and reward could not be predicted solely from any individual decision. Moreover, since the monkeys performed multiple trials, the decision of an individual to cooperate or defect on one trial may influence the other's subsequent decisions and, therefore, affect the future

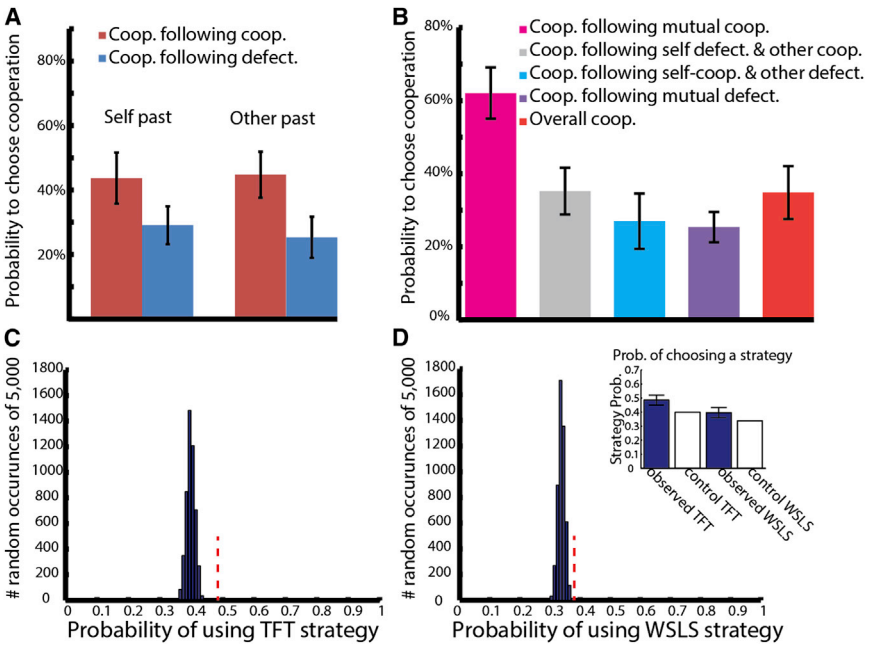


Figure 2. Mutually Beneficial Interactions Increase Cooperation

(A) Conditional probability of a monkey cooperating given that it cooperated or defected on the preceding trial (left) and conditional probability of a monkey cooperating given the opponent cooperated or defected on the preceding trial (right). Error bars represent SEM.

(B) Probability of selecting cooperation following both monkey's prior mutual selections. Red bar denotes overall cooperation probability. Mutually beneficial interactions led to an increase in subsequent cooperation (this was not evident when playing a computer opponent or in separate rooms, see text).

(C) Probability of following tit-for-tat (TFT) strategy. Histogram shows probability for 5,000 control Monte Carlo realizations of surrogate behavioral data. Red dashed line indicates experimental data value.

(D) Probability of following win-stay-lose-switch (WSLS) strategy. Red dashed line indicates experimental data value. Inset denotes observed data values of both strategies (blue bars), error bars represent SEM, white bars denote mean of surrogate control values.

See also Figure S1.

potential for mutual benefit. Here, we used this setup to differentiate between potential neuronal signals that encoded self-decisions, other-decisions, and expected reward as both monkeys jointly, simultaneously made their own choices.

The monkeys sat side-by-side, facing a screen that displayed different targets representing the choice to cooperate or defect (note, that facial expression observations or eye contact were not possible here by design). Neither monkey saw the other monkey's selection until after they made their own selection plus an additional blank screen delay. Then both selections were revealed on-screen followed by reward (Figure 1C). To further rule out implicit signals such as auditory cues that may contribute to predictions of the other's decisions, we randomly alternated the order in which monkeys made their selections (see below).

Behaviorally, we find that the monkeys were more likely to select defection over cooperation. The monkeys performed 1,346 trials over seven sessions; they chose defection in 65.3% of trials and cooperation in 34.7% of trials (chi-square = 123.7, df = 1, $p < 10^{-29}$). They selected cooperation simultaneously on 17.1% of trials, significantly higher than chance level (chi-square = 44.07, df = 1, $p < 10^{-11}$) and both defected on 37.6% of trials, significantly less than chance level (chi-square = 22.27, df = 1, $p < 10^{-6}$). Similar to prior observations in humans (Kuhlman and Marshello, 1975; Rapoport and Chammah, 1965), the monkeys were less likely to cooperate if the other previously defected (26% ± 6%; 2 × 2-chi-square = 56.89, df = 1, $p < 10^{-13}$) (Figure 2A), indicating their understanding of the task by taking into account the other's past action when selecting their own. Moreover, the monkeys were most likely to cooperate if both monkeys cooperated on the preceding trial (62.1% ± 7.0%; chi-square = 76.7, df = 1, $p < 10^{-18}$) (Figure 2B), despite the fact that individual reward is maximized if a monkey defects when his opponent continues cooperating (note these choices did not reflect a simple tit-for-tat response; see Supplemental Information and Figure S1). In other words, the monkeys reciprocated mutual cooperation for continued mutual benefit. Finally, we examined the behavioral strategy followed by the monkeys by analyzing specific choice sequences and found that they were significantly different than chance (Figures 2C and 2D; see Supplemental Information).

Behavioral Controls

To determine whether the monkeys' choices were affected by social context, i.e., their interaction with another monkey, we repeated the task in the exact same set-up, only now replacing a monkey with a computer opponent (Chang et al., 2013; Hosokawa and Watanabe, 2012). The computer's choices were determined by the statistics of monkeys' choices on the previous sessions, described above (see Supplemental Information). We find that the monkeys were less likely to cooperate overall (19.1% ± 3.9% versus 34.7%; chi-square = 161.73, df = 1, $p < 10^{-36}$). Moreover, they were less likely to reciprocate cooperation following mutual cooperation (14.5% ± 3.0% versus 62.1%; chi-square = 73.25, df = 1, $p < 10^{-17}$) when playing a computer opponent, therefore leading to less mutually beneficial interactions.

To eliminate the possibility that the reduced cooperation resulted from differences in choice selection between the computer model and the behaving monkey, we performed an additional set of social control experiments. Here, the monkeys were placed in two separate rooms so that they could not see the other player or hear each other's licking sounds. In addition, the monkeys' juicers were placed outside the experiment room to eliminate any cues from juicer clicks. Under these conditions, the monkeys performed the same task as before with each other. The monkeys performed a total of 2,344 trials in five experimental sessions. By and large, we find the behavior of the

monkeys in this control to be similar to the behavior found in the computer opponent control. Namely, the overall probability of the monkeys to cooperate under these conditions significantly dropped to 14.2%, compared with 34.7% when playing together (chi-square = 432.08, df = 1, $p < 10^{-95}$). Furthermore, we did not observe the increased cooperation following mutual cooperation that was a signature of the monkeys' behavior when playing each other in the same room. Namely, the probability of cooperating following a mutual cooperation trial dropped to 17.4% compared with 62% when playing in the same room (chi-square = 38.76, df = 1, $p < 10^{-9}$). This value closely matches the computer control value of 14.5% (not significant [n.s.] difference). Therefore, the effect of social context on the behavior of the monkeys is corroborated by these two independent control experiments (i.e., computer control and other room control).

As noted above, the monkeys demonstrated their understanding of the task by taking into account past joint decisions when selecting their own. However, to further confirm that the monkeys understood the relationship between their choices and payoff, the monkeys performed an additional control version of the task in which they were presented with the same choices as before, but could now see the other's selection before responding (see Supplemental Information). We find that, on trials in which the other monkey first defected, the monkey maximized reward by subsequently selecting defection on 90.7% ± 2.2% of trials (i.e., within the same trial when no mutual beneficial outcome was possible). This held true even if the other monkey cooperated on the preceding trial (95.0% ± 3.0%). In other words, the monkeys did not reciprocate a prior offer of cooperation if they knew their opponent defected on the present trial. This did not reflect a simple reward maximization behavior (see Supplemental Information).

Single Neuronal Encoding of Another Individual's Unknown Decisions

We recorded 363 neurons in the dACC in two of the four monkeys during task performance. Of these, 185 neurons significantly responded to the task (stepwise linear regression of neuronal firing rate with both monkeys' current and past decisions as predictor variables, corrected for comparisons across pre- and post-selection periods) (Figures 3A–3D and S2; Table S1; Experimental Procedures; Supplemental Information). In total, 24.3% of neurons encoded the monkey's own choices on the current trial; 15.7% responded differentially to choosing cooperation versus defection during the pre-selection period (immediately before the monkey's selection) while 11.4% responded differentially during the post-selection period (immediately after the monkey's selection; $p < 0.05$) (Figure 3A). There was a 2.33-fold ± 0.26-fold change in absolute activity between cooperation and defection when considered across all such neurons ($p < 0.05$). While the sign of the modulation of neural activity was similar in most neurons when the monkeys chose to defect, responses were more variable across neurons when the monkeys chose cooperation. Approximately half of these neurons (54.7%) had an increase of activity whereas the other half presented a decrease in activity (Figure 3C, left panel). In other words, many dACC neurons encoded the monkey's decision to cooperate or defect.

The key for succeeding in this game was the ability to anticipate the other monkey's concurrent decisions. Analyzing neural activity during the time when monkeys were still unaware of the other's concurrent selection, we found that the activity of many neurons was modulated by the other monkey's yet unknown upcoming choice. A total of 32.4% of neurons demonstrated significant differences in activity when the other monkey concurrently selected cooperation versus defection. Most of these (27.6%) encoded the opponent's unknown choice during the post-selection period (but prior to being informed of the other's response) and 7% during pre-selection period ($p < 0.05$) (Figure 3B). There was a 1.81-fold ± 0.07-fold change in absolute activity between other's cooperation and defection when considered across all such neurons ($p < 0.05$) (Figure 3C, right panel; note that the total number of neurons encoding current decisions was larger when considering past responses; see Supplemental Information and further below).

Neurons encoding the opponent monkey's choices and neurons encoding the monkey's own choices demonstrated little overlap with each other (Figure 3D). Only 4.3% of neurons responded to both the monkey's own decisions as well as the opponent's planned decisions. This was significantly lower than chance level, i.e., that expected by a product of the individual probabilities of encoding self and other (expected: 7.9%, chi-square = 4.97, df = 1, $p < 0.026$). This suggests that self and other related computations were carried out by largely distinct neuronal populations (Figures S3 and S4; Supplemental Information).

To further delineate and confirm the response characteristics of these neurons, we applied three additional approaches to re-analyze the data. First, we performed a choice probability (CP) index analysis examining the trial-by-trial encoding of single neuronal responses. CP index analysis results closely matched the stepwise regression results (35.7% of task responsive neurons had a significant CP index for encoding the other's choice post-selection, and 21.6% had a significant CP index for encoding self-decision pre-selection) (Figures 3E and S5A; Supplemental Information). Second, we performed an Akaike Information Criterion (AIC) analysis, which penalizes models containing multiple terms, to complement the term selection process in the stepwise linear regression (Figures S5B–S5E; Tables S2A and S2B). Finally, we performed an unsupervised population analysis in the form of a mixture of linear regression models to test in a more unbiased fashion the behavioral factors to which neurons responded at the population level (Figures S6A–S6F). These analyses confirmed the existence of self and other encoding neurons and the prominence of other-predictive neurons in the dACC and further demonstrate that our findings based on the neuronal data were reproducible across statistical methods (see Supplemental Information).

Neurons Predicting the Other's Unknown Decision Are Sensitive to Social Context

To test the direct effect of social context on neural encoding, we recorded a total of 164 additional neurons from the dACC during the social control experiment in which the monkeys played together but in separate rooms. Of these, 84 neurons were found task-responsive using the same stepwise regression analysis as

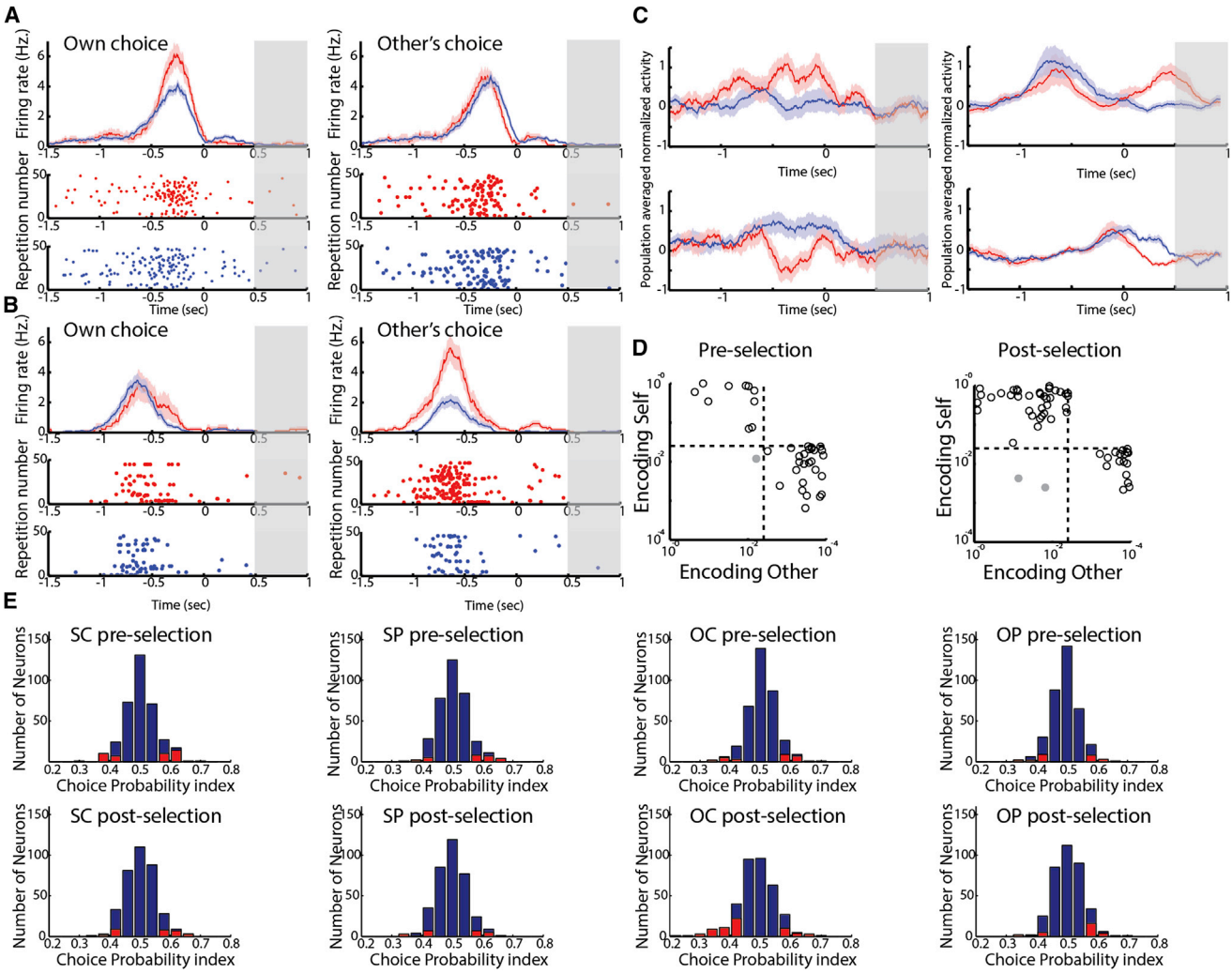


Figure 3. Distinct dACC Neurons Encode Self and Other's Decisions

Peristimulus histograms as mean firing activity ± SEM and raster plots for individual neurons. Cooperation trials are denoted in red and defection in blue. Time zero denotes monkey's own selection.

(A) Left: an example of a neuron that encoded the monkey's own current decision to cooperate or defect. Right: the same neuron did not encode the opponent's yet unknown decision. Gray bar indicates the time when both decisions were revealed to the monkeys (on half of trials; see text).

(B) Example of a neuron that encoded the opponent monkey's yet unknown decision to cooperate or defect (right), but did not encode the monkey's own current decision (left).

(C) Population responses based on the monkey's own current decisions for neurons that had a significantly higher activity during self-cooperation versus self-defection (top left) and significantly lower activity during self-cooperation versus self-defection (bottom left); and population responses for neurons that had significantly higher activity during other-cooperation versus other-defection (top right) and significantly lower activity during other-cooperation versus other-defection (bottom right).

(D) Functional partitioning within the population between neurons encoding the monkey's own current decisions and the opponent's yet unknown decisions. Log-log-scale scatter plots of individual neurons p values obtained from the regression analysis during pre- (left) and post-selection (right) periods (only significant neurons are shown). Dashed lines denote significance thresholds. Gray points denote neurons that significantly encoded both the monkey's own decisions and the opponent's decisions.

(E) Neurons with significant modulation based on choice probability (CP) analysis. Top row: pre-decision time period, bottom row: post-decision time period. Columns from left to right correspond to different behavioral variables (SC, self-current; SP, self-past; OC, other-current; OP, other-past). Red bars indicate significant neurons as obtained by bootstrap estimate.

See also Figures S2, S3, S4, S5, and S6 and Tables S1, S2A, S2B, and S3.

above ($p < 0.025$ for any main or interaction effect, either during the pre or post selection period; see Table S3). We found that only 14.3% of task responsive cells predicted the other's choice, significantly less than the 27.6% observed in the main task (chi-square = 7.42, df = 1, $p < 0.006$; post-decision). In contrast, a significantly larger fraction of task-responsive neurons encoded

the monkey's own decision in the separate room control (21.4% during the pre-selection period and 26.2% during the post-selection period, compared to 15.7% and 11.4% respectively in the main task; pre-selection: chi-square = 2.083, df = 1, p = 0.149; post-selection: chi-square = 18.193, df = 1, p < 0.00002). One possible explanation for the higher number of neurons encoding the monkey's own decisions is that there were more trials recorded per session during the separate room control. However, if this was the only factor, we would also expect to have a concurrent increase in the number of other-predictive neurons, which was not the case. Moreover, the increase in neurons encoding self-decisions indicates that the drop in other-predictive neurons was not simply due to a difference in the raw number of overall cooperation/defection trials. Therefore, this considerable reduction in the fraction of other-predictive neurons indicates that other-predictive neurons are significantly and selectively sensitive to social context.

Neurons Encoding the Other's Unknown Decisions Do Not Encode Expected Reward

While certain cingulate cells are known to encode received and expected reward (Seo and Lee, 2007; Sheth et al., 2012; Williams and Eskandar, 2006), cells encoding self or other decisions were largely distinct from those that encoded expected reward. An important feature of the iPD game is that it enables one to dissociate neuronal signals encoding self and other decisions from those related to expected reward. Specifically, the monkey's own choice alone cannot guarantee a high or low reward. Therefore, predicting one's own reward inherently requires an accurate prediction of the opponent's yet unknown selection. Nonetheless, to demonstrate more directly that the activity of cells predicting other-decisions is not explained by encoding of expected reward, we provide four lines of evidence based on examining the neuronal responses across multiple behavioral outcome contingencies.

First, we directly examined the encoding of expected reward during the decision period. We found that none of the other-predictive neurons was significantly modulated by self-reward across all four reward contingencies determined by the payoff matrix (see Supplemental Information for statistical tests). Second, we examined the differences in firing rate modulation between encoding of other decision and encoding of self-reward across the recorded population. We found that the firing rate modulation of other-predictive neurons was strong and significantly different from the general population when considering differences in the other's choice to cooperate or defect (Figure 4A), but not when aligning trials according to differences in the monkey's own expected reward, i.e., comparing trials in which the monkey cooperated or defected when the other choose to defect (Figure 4B) and when the other chose to cooperate (Figure 4C). Note that while we did find neurons in the dACC that showed strong modulation to self and other reward (as previously reported by Azzi et al., 2012; Chang et al., 2013; Hosokawa and Watanabe, 2012), these were distinct from the other-predictive neurons (Figure S7A; Supplemental Information). Third, we examined the reward feedback period itself, as it may have been possible that other-predictive neurons only encode reward weakly during the decision period when outcome is uncertain, but are more strongly modulated by reward when it is certain or known. However, we found that this was not the case (Figure 4D). In fact, compared to other cingulate cells, which overall demonstrated an enhanced modulation to expected reward during feedback, other-predictive neurons demonstrated a slight, non-significant reduction in modulation (Figure 4E). Finally, to test whether other-predictive neurons could be simply sensitive to raw difference in amount of reward irrespective of choice, we repeat the comparison between feedback time modulation and decision time modulation, but for the contingency that yielded the maximal difference in reward, and find no difference in modulation of the other-predictive neurons (Figure 4F).

In summary, we demonstrate that the response properties of other-predictive neurons were not explained by simple encoding of the monkey's own expected reward (see Supplemental Information). These results are further bolstered by the finding above that other-predictive neurons encoded no significant information about self-decisions and that they were highly sensitive to social context compared to other population cells.

dACC Populations Accurately Predict the Other's Decisions on a Trial-by-Trial Basis

Activity in the dACC was significantly predictive of self and other's choices on a trial-by-trial basis when considered across the entire population (Figures 5A and 5B). We constructed a linear decoder to predict the monkeys selections based on population activity (see Supplemental Information). Evaluating model performance on validation trials not used for model training, we find that cingulate populations predicted up to 66.1% ± 0.9% of the recorded monkey's own current choices (multivariance analysis of variance [MANOVA], $p < 10^{-4}$), with predictions being most pronounced in the pre-selection period (Figure 5C). Surprisingly, population activity correctly predicted the other monkey's yet unknown choices on up to 79.4 ± 1.1% of trials (MANOVA, $p < 10^{-5}$), with predictions being most pronounced in the post-selection period (Figure 5D). Prediction of other's unknown choices was significantly more accurate than prediction of monkey's own current choices (paired t test, $p < 10^{-5}$).

To more directly examine the role that the cells selected as other-predictive neurons by the regression analysis play in population decoding of the other's yet unknown decision, we next ran the decoder using only this subset of the neuronal population. We find that the accuracy of predicting the other monkey's decision was not affected and remained up to 78.1% ± 0.8% (MANOVA, $p < 10^{-9}$) correct, despite the fact that the decoder had access to far less cells. However, the accuracy of decoding the monkey's own decisions drastically dropped and was only up to 54.7% ± 0.9% (MANOVA, p = 0.37, n.s.). These specific effects found in restricting the analysis to this subset of neurons further support the above ascribed role of other-predictive neurons, as well as the functional distinction between these cells and those that encode the monkey's own selections.

Finally, we considered whether implicit cues between the two monkeys could explain these predictions. Note that an important aspect of the task design was that the monkeys made their selections in random temporal order before their responses were revealed. Accordingly, we tested the population predictions

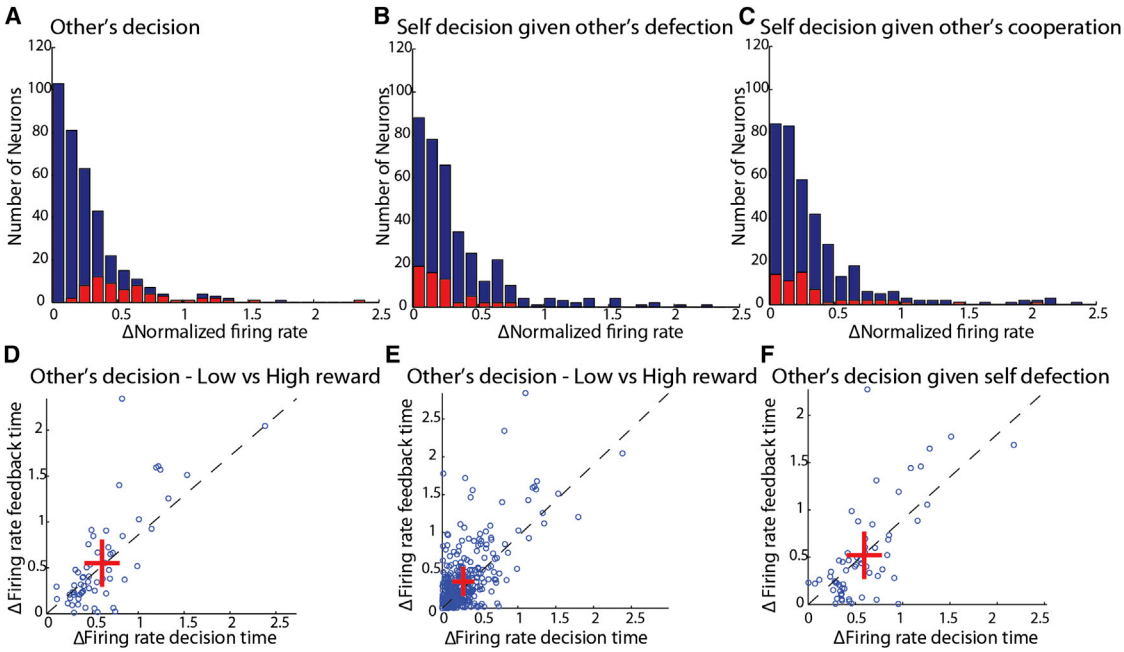


Figure 4. Other-Predictive Neurons Do Not Encode the Monkey's Own Expected Reward as Shown across Multiple Reward Contingencies

(A) Histogram of normalized difference in firing rate between trials in which the other monkey defected versus cooperated. Red bars indicate other-predictive neurons. Blue bars indicate the full population. The distributions were statistically different.

(B) Histogram of normalized difference in firing rate between trials in which the monkey chose defection versus cooperation, conditioned on the other choosing defection. Red bars indicate other-predictive neurons. Blue bars indicate the full population. No significant difference was found between distributions.

(C) Histogram of normalized difference in firing rate between trials in which the monkey chose defection versus cooperation, conditioned on the other choosing cooperation. Red bars indicate other-predictive neurons. Blue bars indicate full population. No significant difference was found between distributions.

(D) Scatter plot of firing rate difference between trials in which the other defected versus cooperated, for firing rate during decision time (x axis) and feedback time (y axis) in other-predictive neurons. There is no increase in differential activity when reward is known. Crosses represent mean ± SEM.

(E) Scatter plot of firing rate difference between trials in which other defected versus cooperated, for firing rate during decision time (x axis) and feedback time (y axis) in the full population. Here, there was a significant increase in differential activity when reward is known.

(F) Scatter plot of firing rate difference between trials in which the monkey chose defection versus cooperation, conditioned on other's defection, for firing rate during decision time (x axis) and feedback time (y axis) in other-predictive neurons. Here, there is no increase in differential activity when reward is known.

See also Figure S7.

when considering only trials in which the monkey played first, i.e., when the other monkey hadn't yet made his selection. We found that predictions of other's unknown choices maintained high accuracy (up to 70.7% ± 0.8%) and similar accuracies were found when considering only trials in which the monkey played second (68.5% ± 7.2%), ruling out the possibility that prediction is an artifact of an implicit signal disclosing the other monkey's choice. Note lower accuracy was expected due to using half the number of trials.

Behavioral Trial-by-Trial Decoders

To search for a possible basis for neural prediction of the other's concurrent selections, we examined predictions based on both monkeys' prior behavioral history. Using a locally-optimal classification model considering the monkeys' selections four trials back, we estimated on validation trial data the accuracy of predicting the opponent monkey's unknown concurrent choices. We find that model prediction accuracy was up to 79.8%, similar to neuronal decoding (similar accuracies were found for predicting self-selections, see Supplemental Information). To further explore the behavioral basis of the neuronal predictions of other's decisions, we tested trial-by-trial correlation between the behavioral and population-activity predictors, revealing significant correlations based on both monkeys' past selections ($r = 0.31$, $p < 0.0003$). These correlations of other's predictions were not evident when behavioral predictions were based on only a single monkey's past decisions or reward (see Supplemental Information). This suggests that population predictions were based on the prior choices of the two monkeys rather than any individual's past response or reward.

Neurons Keeping Track of Past Interactions

Consistent with the above findings, we find that many neurons within the population kept a dynamic record of the monkeys' prior selections. Figure 5E illustrates such a neuron; when the monkey chose to currently defect (left panel), responses did not differ when, on the preceding trial, the opponent chose to defect versus cooperate. In contrast, when the monkey himself cooperated (right panel), neuronal activity was significantly inhibited on trials in which the opponent previously defected (i.e., the monkey cooperated despite the opponent previously defecting) compared to those in which the opponent cooperated

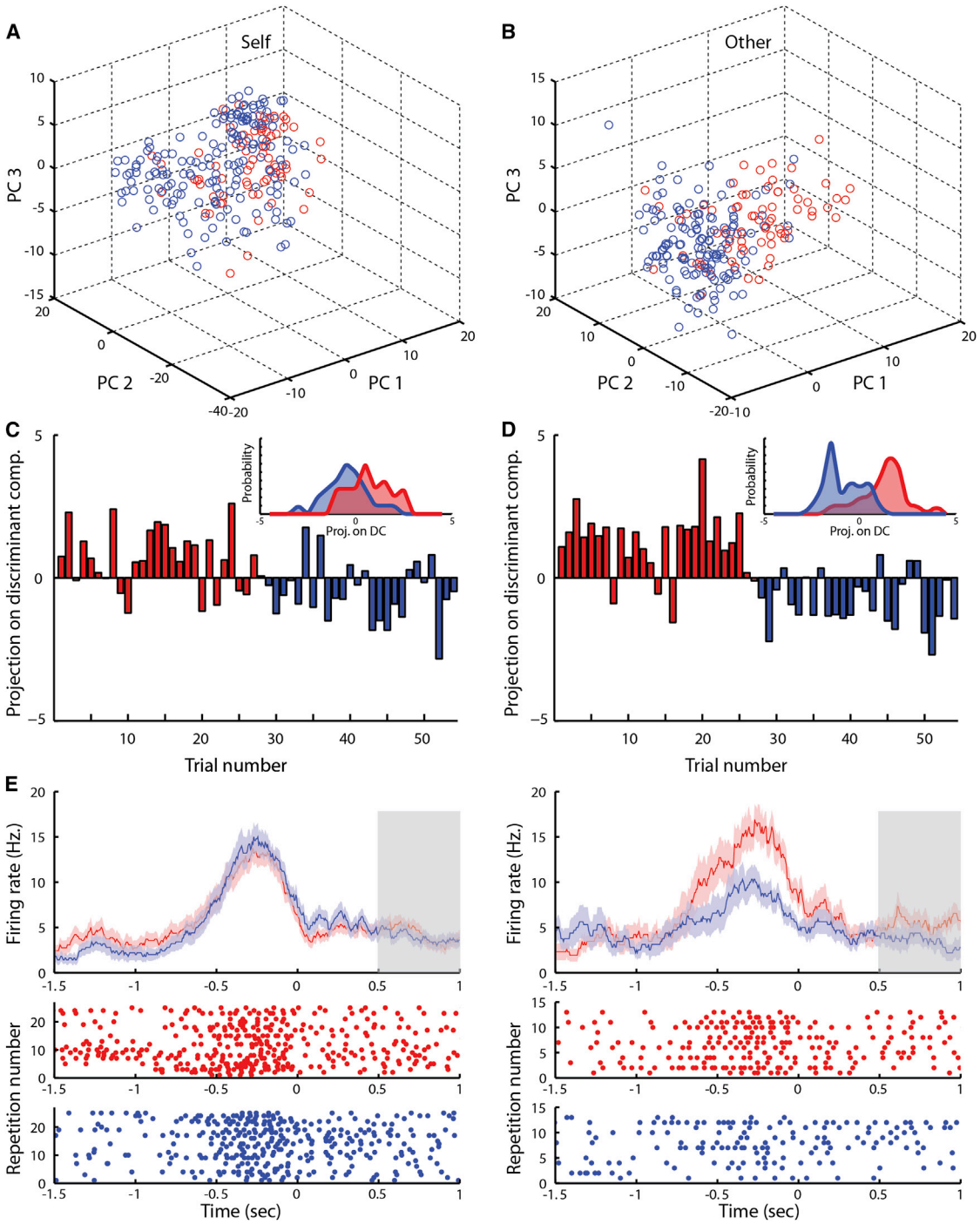


Figure 5. Trial-by-Trial Population Prediction of the Other's Yet Unknown Decision

(A and B) Principal component (PC) analysis over a sample session. Plotted in first three PC space, each circle represents the activity of all neurons recorded simultaneously on a single cooperation (red) or defection (blue) trial (see Supplemental Information).

(A) Self-current pre-decision activity.

(B) Other's-concurrent (yet unknown) post-decision activity.

(C and D) Linear decoding model. Each bar represents projection of the activity of all simultaneously recorded neurons during a single trial on first discriminant component (color code above). Positive values predict cooperation and negative defection. Insets (top right) plot distribution of projection values for cooperation (red) and defection (blue).

(C) Self-current pre-selection projection.

(D) Other's-concurrent projection during post-selection.

(legend continued on next page)

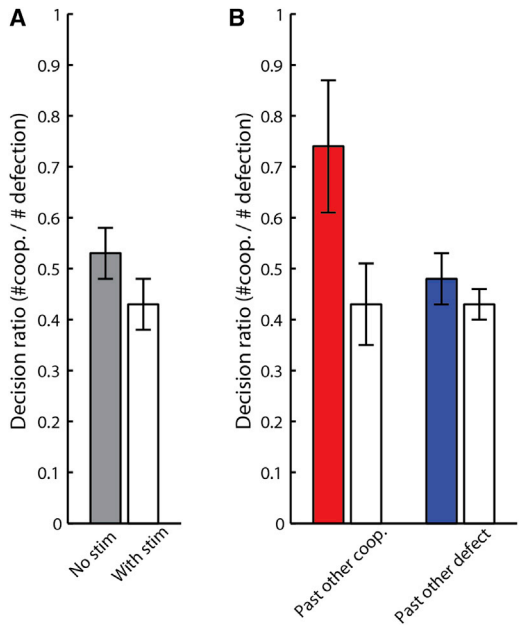


Figure 6. dACC Stimulation Selectively Inhibits Mutually Beneficial Interactions

White bars represent stimulation trials.

(A) Proportion in which the monkeys chose cooperation over defection ± SEM (decision-ratio of 1 indicates equal proportion of selecting either).

(B) Decision-ratio given the opponent's past decisions to cooperate (left), or defect (right).

(i.e., reciprocating opponent's preceding cooperation). In addition we found neurons that differentially encoded the joint outcomes on preceding trials (see Figure S7B and Supplemental Information for further details).

Cingulate Disruption Selectively Inhibits Mutually Beneficial Interactions

Given the above physiological findings, we next investigated whether disruption of the dACC may influence the monkeys' mutual choices. A series of electrical pulses was delivered to the dACC on half of 3,026 randomly selected trials in blocks (1,000 ms triggered at image presentation; 100 μA, 200-μs biphasic pulse durations with cathodal phase leading; see Supplemental Information).

Stimulation had a significant and selective effect on the monkeys' decisions. Here, we defined the "decision-ratio" as the number of trials in which the monkey selected cooperation over defection (i.e., a ratio of 1 indicates equal selection of cooperation versus defection). When no stimulation was given, the decision-ratio was 0.53 (corresponding to 34.7% cooperation, as also found in the main task). When stimulation was administered, the decision-ratio dropped to 0.43, i.e., monkeys were less likely to cooperate when stimulated ($t(6) = 3.18$, $p < 0.01$) (Figure 6A). This effect was highly dependent on the opponent monkey's preceding selection. When the opponent previously cooperated and no stimulation was given, the decision-ratio was 0.74, meaning that monkeys were more likely to choose cooperation if the opponent previously chose cooperation. However, during stimulation, following opponent's cooperation, the decision-ratio significantly dropped to 0.43 ($t(6) = -5.57$, $p < 0.0007$) (Figure 6B). In contrast, following opponent's defection when no stimulation was given, the decision ratio was 0.48 and, when stimulation was given, it was 0.43 ($t(6) = -1.12$, $p = 0.15$). In other words, stimulation had no effect on the monkey's current decision if the opponent previously defected, but when the opponent previously cooperated, stimulation reduced the decision-ratio to a level equal to the opponent previously choosing defection. Moreover, stimulation had no effect on risk behavior since the rate of cooperation when the other monkey defected on the preceding trial was not affected by stimulation (even though the risk of cooperation under such a condition is higher; i.e., the probability of the opponent to defect following defection is twice higher than following cooperation).

Finally, to further confirm that stimulation did not simply affect decisions based on past reward, we employed a zero-sum game task in which monkey's reward was contingent on the other's response, but individual profit was always at the expense of the other and no mutual positive outcome was possible (i.e., playing under Pareto optimality conditions) (Nash, 1950). We found no effect of stimulation on monkeys' choices during the zero-sum game, based either on the monkeys' preceding selection or preceding receipt of reward (Figure 7; Supplemental Information). Taken together, we conclude that stimulation in the dACC abolished specifically the incorporation of recent positive interactions, rather than any past interaction, into the monkey's own current decision, resulting in less mutually beneficial interactions.

DISCUSSION

Identifying neurons that reflect another individual's covert intentions or "state of mind" has been a long sought goal in neuroscience and a central proposed tenet of social decision making (Frith and Frith, 1999; Rilling et al., 2004; Sanfey et al., 2006; Vogeley et al., 2001). Here, we discover neurons that selectively encode another individual's yet unknown decisions during joint interactions. We confirmed that no explicit cues were relayed between the two monkeys during the task by using alternating trials in half of which the monkey from which we obtained recordings played first. We also demonstrated reliable population predictions of the other's decisions even on trials in which the other monkey had not yet made his selection. Remarkably, other-predictive cells during joint interactions constituted over a third of the cingulate task-responsive population and were more prevalent than cells encoding the monkey's own present selections. Notably, other predictive neurons were highly sensitive to social

(E) Peristimulus histograms as mean firing activity ± SEM (top) and raster plots of a neuron encoding the monkey's own current decision during the pre-selection period and modulated by the other's past decision. Trials separated according to monkey's own current decision to defect (left) or cooperate (right) and opponent's decision on a preceding trial to cooperate (red) or defect (blue; see text). Time zero denotes monkey's own selection. Gray bar indicates feedback period.

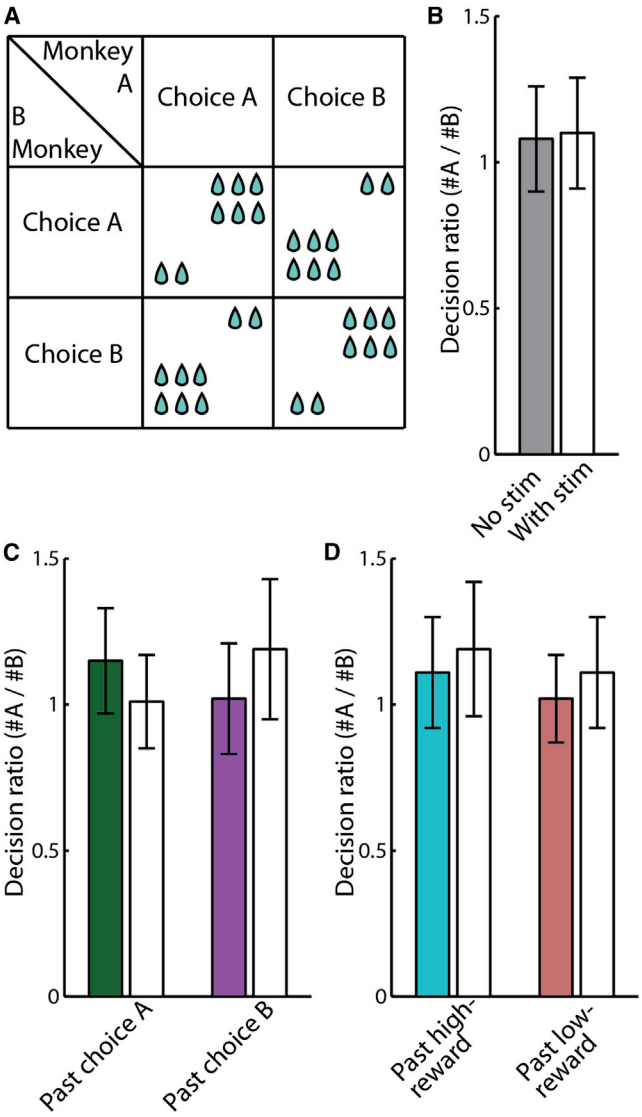


Figure 7. Stimulation Has No Effect when No Mutually Beneficial Interactions Were Possible

(A) Zero-sum game payoff matrix.

(B–D) Bars represent the decision-ratio on stimulated (white) and non-stimulated (colored) trials during the zero-sum game (see Supplemental Information). Error bars represent SEM. (B) Overall decision-ratio. (C) Decision-ratio was not affected by opponent's selection of choice A (left) or choice B (right) on the preceding trial. (D) Decision-ratio was not affected by the monkey's own past reward. Left bars: the monkey previously received a high reward. Right bars: the monkey previously received a low reward.

context and were not modulated by self-decisions or expected reward. Consistently, population predictions of the opponent's selections were more accurate than those reflecting the monkey's own selections and, in fact, predicted the other monkey's decisions with accuracies that were near optimal compared to behavioral decoders that considered both monkey's past behaviors. Taken together, these findings provide understanding of the population partitioning by which individual neurons in the primate cingulate cortex encode information about other social agents.

Game theory provides a framework for dissecting specific aspects of joint decision making, namely the contributions of self and other choices to shared outcome. Signals related to another's yet unobservable actions, in particular, are a distinct feature of mutual interactions in that one participant's concurrent decision affects the other's outcome and therefore inherently requires each participant to anticipate the other's intentions or state of mind.

These predictive signals are fundamentally distinct from previously reported neurons which reflect another animal's known and observable actions. These include canonical mirror neurons that reflect one's observed behavior and do not distinguish between self and other (di Pellegrino et al., 1992; Rizzolatti and Sinigaglia, 2010), neurons that encode another's observed receipt of reward (Azzi et al., 2012; Chang et al., 2013), and neurons that monitor other's observable errors (Yoshida et al., 2012). Importantly, the prediction neurons reported here are distinct from the findings of the latter study, in which neurons monitored the other's errors while the monkeys explicitly observed each other's selections on the same shared task (with each monkey alternating between actor and observer every other trial) (Yoshida et al., 2012). Moreover, encoding of the other's error occurred within the monkeys' movement time window (<200 ms before other's response) and in a setup which allowed them to directly observe each other's movement-preparatory cues. Here, decisions were made jointly, the other's decisions were inherently unobservable and unknown, and their neural encoding could be found many seconds before the other monkey made a selection.

A central feature of non-competitive games such as iPD is that no particular decision guarantees a high or low reward and different outcomes can be experienced either mutually or individually. This dissociation enabled us to examine the computations that contributed to self and other predictions and differentiate them from those that contribute to the encoding of reward outcome. More importantly, it allowed us to examine what particular computations were associated with interactions that were mutually beneficial compared with those that were not. For instance, the monkeys were almost twice as likely to cooperate if they both cooperated on the preceding trial, indicating an intention to reciprocate mutual cooperation. Here, we find that neurons that encoded a monkey's decisions largely did not encode his past or future receipt of reward even though, in combination, these neural signals could be used to predict the monkey's shared outcome. Many neurons, however, were also highly modulated by the two monkey's prior selections. For example, certain neurons differentially encoded the monkey's present decision to cooperate, based on the other monkey's preceding selection of cooperation or defection. Similarly, at the population level, neuronal predictions strongly correlated with predictions made by the behavioral decoder when considering both monkeys' past selections, indicating that neural predictions were based on the past interaction of both individuals.

Consistent with these physiological findings, we observed that disruption of the dACC by stimulation reduced the monkey's

likelihood of cooperation, an effect which was most evident when the opponent cooperated on the preceding trial. Stimulation therefore affected reciprocation of the other's cooperation, but did not affect the animal's ability to incorporate any past decision or outcome since no effect was observed when the opponent defected on the previous trial, or when testing the monkey's decisions in a zero-sum game. This is consistent with previous studies employing a computer opponent in zero-sum games that showed that the dACC does not differentially encode the monkey's decisions during such interactions (Donahue et al., 2013; Seo and Lee, 2007). Therefore, during joint interactions, the dACC specifically mediated mutually beneficial decisions based on the recent history of the interaction.

The monkeys were clearly affected by the social context of their interaction, as they significantly changed their behavior when playing either against a computer opponent or in separate rooms, consistent with prior reports (Carter et al., 2012; Chang et al., 2013; Hosokawa and Watanabe, 2012). Moreover, other-predictive neurons were selectively influenced by social context compared other population cells, suggesting that these cells encoded information that was specific to other social agents rather than any information about the environment which affected outcome. The monkeys also selected the appropriate responses when their opponent's decisions were known, suggesting that they understood the consequent payoff. While the joint nature of the task precludes the possibility of identifying "involuntary errors" by the individual animals, we find that the monkeys made incorrect selections on <10% of sequential control trials making such rare occurrences highly unlikely to qualitatively affect the study's results. This conclusion is also supported by the finding that the population prediction of the opponent's decisions was robust to substantial deletion of trials. However, as with any animal or human study that investigates interactive behavior, what internal thought process truly motivates these different behaviors can only be speculated upon. On this point, we note that cooperation is based on the observable action of two interacting individuals, rather than its hidden motivation, and is defined explicitly as the selection of actions capable of leading to joint benefit but which can also lead to loss if the action is not mutual.

Taken together, the present findings support the proposed role of the dACC in encoding a dynamic model of the environment (Adolphs, 2009; Karlsson et al., 2012; Sheth et al., 2012) but considerably expand it into the inclusion of mutual interactions which require an explicit representation of another's yet unknown behavior. The two distinct groups of neurons found in the dACC, encoding the self versus predicting the other's decisions, may therefore be uniquely suitable to allow the soon-available actual decision of the opponent and known decision of the acting monkey to update the internal model of their joint decisions in a way analogous to delta-learning (Pouget and Snyder, 2000) or an actor-critic (Parush et al., 2011; Williams and Eskandar, 2006; Witten, 1977) framework. Given the broad anatomical connectivity of the dACC to areas that encode aspects of socially-guided interactions, including the temporal-parietal junction, superior temporal sulcus, amygdala and orbitofrontal cortex, the dACC is likely to be part of a wider network of areas, sometimes referred to as the "social brain." The observed role of the dACC in predicting another's intentions contributes to our understanding of this proposed network. For instance, disruption of its activity markedly degraded cooperative behavior, suggesting that dACC activity may be necessary for constructive interaction between individuals and social learning. Such deficits are particularly prominent in individuals with autism-spectrum disorders or antisocial behavior in which anticipating another's intentions or state of mind and incorporating them into one's actions are severely affected (Frith and Frith, 1999; Lombardo and Baron-Cohen, 2011). Our neuronal findings in combination with the behavioral effects observed with stimulation may therefore pave the way toward targeted treatment in the dACC for these or similar disorders in which dysfunctional social behavior is a predominant feature.

EXPERIMENTAL PROCEDURES

Task Design

Four adult male Rhesus monkeys (*Macaca Mulatta*) across four paired combinations were trained to play an iterated prisoner's dilemma (iPD) game. On successive trials, two images (an orange hexagon and a blue triangle) were randomly displayed on the left and right of the screen (Figure 1A). Each monkey selected one of the two images using a joystick and was not shown the other monkey's concurrent selection. The outcome of each monkey's selection depended on both of their concurrent choices, according to the payoff matrix shown in Figure 1B. Based on these payoffs, the orange hexagon was operationally defined as "cooperation" since mutual cooperation led to the highest mutual reward (Camerer, 2003). The blue triangle was operationally defined as "defection" since unilateral defection led to the highest individual reward. However, if both monkeys defected, they each received less reward than if they both cooperated. Note, importantly, that the terms cooperation and defection are used here solely to indicate the potential for mutual benefit or loss dependent on the opponent's selection. Mutual cooperation and mutual defection indicates that both monkeys made the same choice. See Supplemental Information for trial structure details.

Neuronal Recording and Stimulation

Single-Unit Isolation and Recordings

All procedures were performed under approval by the Massachusetts General Hospital institutional review board and were conducted in accordance with Institutional Animal Care and Use Committee (IACUC) guidelines. Prior to recordings, floating micro-electrode arrays (MicroProbes for Life Sciences) were surgically implanted in each monkey. The electrodes were implanted in the dACC through a wide craniotomy under stereotactic guidance (David Kopf Instruments). The location of the arrays was confirmed by direct visual inspection of the sulcal and gyral anatomy with the electrode tips located 8 mm from the cortical surface. Each array had 36 microelectrodes spaced horizontally 400 μm apart. Electrode leads were secured to the skull and attached to connectors with the aid of titanium miniscrews and dental acrylic.

Recordings began 2 weeks following surgical recovery. A Plexon multichannel acquisition processor was used to amplify and band-pass filter the neuronal signals (150 Hz–8 kHz; 1 pole low-cut and 3 pole high-cut with 1,000× gain; Plexon). Shielded cabling carried the signals from the electrode array to a set of six 16-channel amplifiers. Neural signals were then digitized at 40 kHz and processed to extract action potentials by the Plexon workstation. Classification of the waveforms was performed using template matching and principal component analysis based on waveform parameters. Only single-, well-isolated units with identifiable waveform shapes and adequate refractory periods were used. When an individual electrode recorded more than one neuron, a high degree of isolation was required in order to include each as a single-unit ($p < 0.01$, multivariate ANOVA across the 1+ two principal components). We did not include multi-unit activity.

Electrical Stimulation Protocol

During stimulation trials, the monkeys performed the iPD and zero-sum games in separate sessions. Each session was composed of randomly selected

30–40 stimulated trials followed by another 30–40 trials in which no stimulation was delivered. Stimulation was administered as a brief series of alternating rectangular positive to negative voltage pulses. Stimulation parameters were 100 μA and 200 Hz biphasic pulses, with cathodal phase leading. Average impedance at the time of the stimulation experiments was 100–500 kΩ. Here, all 32 electrode contacts were simultaneously stimulated per array. Stimulation was given for 1,000 ms and included the baseline and image presentation periods. Stimulation ended prior to presentation of the go cue and prior to the monkey's selection.

Statistical Analysis

A stepwise linear regression was conducted in order to determine how the different task parameters modulated the neuronal activity. In this analysis, parameters are incrementally added to the model, starting with the parameter that explains the most variance and continuing on to the parameters that most explain the remaining variance, terminating when parameters no longer significantly explain the residual variance. The model included the four main effect parameters, as described below (self-current, other-current, self-past and other-past) as well as their pairwise interactions (see Equation 1),

$$r(t) = a + \sum_{i=1}^{4} \beta_i^{Main} M_i^{Main} + \sum_{i=1}^{6} \beta_i^{Inter} M_i^{Inter} \quad \text{(Equation 1)}$$

where r(t) is current trial firing rate, $M^{Main} = \{s(t), s(t-1), o(t), o(t-1)\}$ are the four main effects and $M^{Inter} = \{s(t)s(t-1), s(t)o(t), s(t)o(t-1), s(t-1)o(t), s(t-1)o(t-1), o(t)o(t-1)\}$ are the six second order interaction terms; s(t) is current self selection, o(t) is current other selection, and (t − 1) indicates preceding trial.

For brevity, "self" refers here to the selections of the monkey in which neural recordings were performed and "other" refers to the selections of the opponent (i.e., selecting to cooperate or defect). In addition, "current" refers to the two monkeys' current selection (i.e., the trial from which neuronal activity was being evaluated) and "past" refers to the two monkeys' selections on the previous trial. The depended variable is the averaged neuronal firing in the 500 ms period before response selection (i.e., choosing cooperation versus defection) and during the 500 ms period after selection, referred to as "pre-selection" and "post-selection," respectively. Note that we chose to use a stepwise linear regression this analysis since the task parameters and samples were neither balanced nor independent (see further details in Supplemental Information). Multiple complimentary analyses, including a four-way analysis of variance, AIC analysis, and mixture of regressions analysis, yielded qualitatively similar results.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.01.045.

ACKNOWLEDGMENTS

This project was funded by NIH 5R01-HD059852, the Presidential Early Career Award for Scientists and Engineers, and the Whitehall Foundation. Data are available as Supplemental Information. We thank John Assad, Wael Asaad, Shaul Druckmann, Shaul Hochstein, Daeyeol Lee, Israel Nelken, and Sameer Sheth for insightful discussions, Caitlin Commins, Christine Emmanuel, Rebecca Gwaltney, Morgan Jamiel, and Kaitlin Sodon for technical assistance, and Katie Ris-Vicari for the graphical abstract.

Received: April 13, 2014
Revised: October 25, 2014
Accepted: January 5, 2015
Published: February 26, 2015

REFERENCES

Abe, H., and Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. Neuron *70*, 731–741.

Adolphs, R. (2009). The social brain: neural basis of social knowledge. Annu.Rev. Psychol. *60*, 693–716.

Apps, M.A.J., and Ramnani, N. (2014). The anterior cingulate gyrus signals the net value of others' rewards. J. Neurosci. *34*, 6190–6200.

Apps, M.A.J., Balsters, J.H., and Ramnani, N. (2012). The anterior cingulate cortex: monitoring the outcomes of others' decisions. Soc. Neurosci. *7*, 424–435.

Azzi, J.C.B., Sirigu, A., and Duhamel, J.R. (2012). Modulation of value representation by social context in the primate orbitofrontal cortex. Proc. Natl. Acad. Sci. USA *109*, 2126–2131.

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. Nat. Neurosci. *7*, 404–410.

Behrens, T.E.J., Hunt, L.T., Woolrich, M.W., and Rushworth, M.F.S. (2008). Associative learning of social value. Nature *456*, 245–249.

Behrens, T.E.J., Hunt, L.T., and Rushworth, M.F.S. (2009). The computation of social behavior. Science *324*, 1160–1164.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis. Neurosci. *13*, 87–100.

Bshary, R., Grutter, A.S., Willener, A.S.T., and Leimar, O. (2008). Pairs of cooperating cleaner fish provide better service quality than singletons. Nature *455*, 964–966.

Camerer, C. (2003). Behavioral Game Theory: Experiments in Strategic Interaction (Princeton University Press).

Carter, R.M., Bowling, D.L., Reeck, C., and Huettel, S.A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. Science *337*, 109–111.

Chang, S.W.C., Gariépy, J.F., and Platt, M.L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. Nat. Neurosci. *16*, 243–250.

Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. Nature *462*, 51–57.

de Waal, F.B.M. (2000). Primates—a natural heritage of conflict resolution. Science *289*, 586–590.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. Nat. Neurosci. *8*, 1611–1618.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. Exp. Brain Res. *91*, 176–180.

Donahue, C.H., Seo, H., and Lee, D. (2013). Cortical signals for rewarded actions and strategic exploration. Neuron *80*, 223–234.

Frith, C.D., and Frith, U. (1999). Interacting minds—a biological basis. Science *286*, 1692–1695.

Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. Trends Cogn. Sci. *2*, 493–501.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. Proc. Natl. Acad. Sci. USA *105*, 6741–6746.

Hosokawa, T., and Watanabe, M. (2012). Prefrontal neurons represent winning and losing during competitive video shooting games between monkeys. J. Neurosci. *32*, 7662–7671.

Karlsson, M.P., Tervo, D.G.R., and Karpova, A.Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. Science *338*, 135–139.

Kuhlman, D.M., and Marshello, A.F.J. (1975). Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. J. Pers. Soc. Psychol. *32*, 922–931.

Lee, D. (2008). Game theory and neural basis of social decision making. Nat. Neurosci. *11*, 404–409.

Lee, D., McGreevy, B.P., and Barraclough, D.J. (2005). Learning and decision making in monkeys during a rock-paper-scissors game. Brain Res. Cogn. Brain Res. *25*, 416–430.

Lombardo, M.V., and Baron-Cohen, S. (2011). The role of the self in mindblindness in autism. Conscious. Cogn. *20*, 130–140.

Nash, J.F. (1950). Equilibrium points in N-person games. Proc. Natl. Acad. Sci. USA *36*, 48–49.

Parush, N., Tishby, N., and Bergman, H. (2011). Dopaminergic balance between reward maximization and policy complexity. Front. Syst. Neurosci. *5*, 22.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat. Rev. Neurosci. *2*, 417–424.

Pouget, A., and Snyder, L.H. (2000). Computational approaches to sensorimotor transformations. Nat. Neurosci. *3*(*Suppl*), 1192–1198.

Rapoport, A., and Chammah, A.M. (1965). Prisoner's Dilemma; A Study in Conflict and Cooperation (Ann Arbor: University of Michigan Press).

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. Neuron *35*, 395–405.

Rilling, J.K., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2004). The neural correlates of theory of mind within interpersonal interactions. Neuroimage *22*, 1694–1703.

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. Nat. Rev. Neurosci. *11*, 264–274.

Rudebeck, P.H., Buckley, M.J., Walton, M.E., and Rushworth, M.F.S. (2006). A role for the macaque anterior cingulate gyrus in social valuation. Science *313*, 1310–1312.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2003). The neural basis of economic decision-making in the Ultimatum Game. Science *300*, 1755–1758.

Sanfey, A.G., Loewenstein, G., McClure, S.M., and Cohen, J.D. (2006). Neuroeconomics: cross-currents in research on decision-making. Trends Cogn. Sci. *10*, 108–116.

Seo, H., and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. J. Neurosci. *27*, 8366–8377.

Seo, H., Cai, X., Donahue, C.H., and Lee, D. (2014). Neural correlates of strategic reasoning during competitive games. Science *346*, 340–343.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. *16*, 1486–1510.

Sheth, S.A., Mian, M.K., Patel, S.R., Asaad, W.F., Williams, Z.M., Dougherty, D.D., Bush, G., and Eskandar, E.N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. Nature *488*, 218–221.

Stephens, D.W., McLinn, C.M., and Stevens, J.R. (2002). Discounting and reciprocity in an iterated prisoner's dilemma. Science *298*, 2216–2218.

Tomlin, D., Kayali, M.A., King-Casas, B., Anen, C., Camerer, C.F., Quartz, S.R., and Montague, P.R. (2006). Agent-specific responses in the cingulate cortex during economic exchanges. Science *312*, 1047–1050.

Vickery, T.J., Chun, M.M., and Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. Neuron *72*, 166–177.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N.J., Fink, G.R., and Zilles, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. Neuroimage *14*, 170–181.

Warneken, F., and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. Science *311*, 1301–1303.

Williams, Z.M., and Eskandar, E.N. (2006). Selective enhancement of associative learning by microstimulation of the anterior caudate. Nat. Neurosci. *9*, 562–568.

Witten, I.H. (1977). Adaptive optimal controller for discrete-time Markov environments. Inf. Control *34*, 286–295.

Yoshida, K., Saito, N., Iriki, A., and Isoda, M. (2012). Social error monitoring in macaque frontal cortex. Nat. Neurosci. *15*, 1307–1312.

EXTENDED EXPERIMENTAL PROCEDURES

Task Design

iPD Game Trial Structure

At the onset of each trial, the monkeys were given a central fixation point for 500ms (baseline period). Following, the two images were displayed for 500ms (image presentation). After an additional 0-200ms variable delay, a go cue (central green circle) appeared, indicating that the monkey could move his joystick to either image. The monkeys were trained to respond only to the images appearing at the top or bottom of the screen, and did not respond when stimuli appeared on their counter-part of the screen. After the first monkey's selection, the screen would turn blank for an additional 500ms delay (delay period) and the same presentation cycle repeated for the second monkey on the other half of the screen (top versus bottom).

One example of such a sequence would proceed in the following manner: after fixation, a hexagon and triangle are displayed at the top of the screen, on the left and right, for 500 ms. Monkey A makes his selection. After another 500 ms blank screen delay, a hexagon and triangle are displayed on the bottom of the screen for 500 ms. Monkey B makes his selection, and the screen turns blank for 500 ms. Importantly, the succession in which the monkeys were given the top versus bottom images and the location of the images on the left and right of the screen were varied randomly on each new trial.

Only after both monkeys made their successive selections, and a 500ms blank screen delay period, the selections of the two monkeys were revealed (feedback period). Their selections were each displayed on the corresponding part of the screen for 500ms (bottom/top, left/right). Following an additional 500ms blank screen delay, one of the monkeys, in random order, received reward (which depended on both selections, as in the payoff matrix; reward period). Following 1470ms from reward initiation, and at least 500ms following completion of reward delivery (500 ms for highest reward, longer delay for less reward), the other monkey was rewarded. Finally, the screen would turn blank for 1500 ms and the next trial would commence. The task was run using a customized software package written in MATLAB (MathWorks, Inc., Natick, MA) that provides millisecond temporal precision of all task events (Asaad and Eskandar, 2008).

Zero-Sum Game Trial Structure

In a set of controls used for the stimulation experiments, the monkeys additionally performed a zero-sum game (Camerer, 2003; Seo and Lee, 2012). The trial structure and timing was the same as the iPD game. However the images used and payoff matrix were different (Figure 7A). In this task, if both monkeys chose the same image, monkey A got the higher reward, and monkey B got the lower reward. If both monkeys chose different images, monkey B got the higher reward, and monkey A got the lower reward. Therefore, the enhancement of an individual's reward is always at the expense of the other, and no choice could lead to mutually enhanced reward for both monkeys. We confirmed that the monkeys understood the payoff by using an additional sequential task whereby the monkeys could see each other's choices before making their own choice, and verified that the monkeys maximized their reward following their opponent selection (data not shown).

Behavioral Results

Monkeys' Choices during iPD Were Affected by Past Outcomes

Monkeys' choices depended on the joint decision of both monkeys. While the overall cooperation probability was 34.7%, the monkeys cooperated significantly more following mutual cooperation (62.1%, see main text). However, cooperation reciprocity did not reflect a simple tit-for-tat response, i.e., an automatic replay of the opponent's previous response, since the monkeys were less likely to cooperate when the other previously cooperated but the monkey himself defected (35.2% $\pm$ 6.4%; Chi-square = 0.05, df = 1,p = 0.83, Figure 2B). Note that in this case the monkey's cooperation did not differ from their overall cooperation probability. In contrast, when the situation was reversed, i.e., the other defected on the preceding trial while the monkey himself cooperated and got the least possible reward, the monkey was more likely to switch response on the next trial and defect, thereby displaying a lose-switch response (27.0% $\pm$ 7.6%; Chi-square = 6.19, df = 1,$p < 0.013$, Figure 2B).

Next, we considered the conditional probabilities of each response based on their comprehensive behavioral histories. Toward this end, we calculated the probability of the monkey choosing cooperation when both monkeys defected on preceding consecutive trials (from 1 to 4), the conditional probability of choosing cooperation when the monkey himself defected on preceding consecutive trials, regardless of the other monkey's choice on the previous trials (1-4) and the conditional probability of choosing cooperation when the other monkey defected on preceding consecutive trials (1-4).

For each condition, we first assembled all the trials that met that criterion (e.g., trials preceded by 4 mutual defections in a row). In order to obtain the specific conditional probability of choosing cooperation, we then checked in how many of these trials the monkey chose cooperation and divided that by the number of trials that met the criterion. Note that error bars here indicate between session variability and that there are fewer instances of a particular conditioned response as the history of consecutive actions gets longer (i.e., since there are progressively fewer unique instances for progressively longer runs). Therefore, experimental sessions that did not have more than 5 trials that matched a specific category (e.g., 4 mutual defections in a row) were removed from the analysis for that specific condition. Note that, for cooperation following mutual defection, only one session had 4 consecutive mutual defection trials that preceded cooperation, and therefore there was no error bar. As can be seen in Figure S1A, the main observable effect is that, following repeated choices of self-defection (3 and above, *left* panel), the monkey is more likely to continue choosing defection (i.e., the probability of choosing cooperation is significantly reduced. Means: 1-consecutive choices 35% $\pm$ 6%, 2-consecutive 34% $\pm$ 6%, 3-consecutive 26% $\pm$ 1%, 4-consecutive 24% $\pm$ 1%; Paired t test: 1 versus 3: $t(5) = 2.57$, $p < 0.025$, 1 versus 4: $t(5) = 2.76$, $p < 0.019$).

We then tested the likelihood of the monkeys to repeat a given choice, cooperation or defection, across different repetition lengths. For each of the choices, we calculated the probability of finding a repeated choice of a given order and compared it to that expected by chance. The probability expected by chance equals the overall probability of that choice to the power of the repetition length. For example, the by-chance probability of a sequence length two of defections is p(defection) x p(defection) = 0.643 × 0.643 = 0.413. The results of this analysis are presented for repetitions of up to length 4 (since there were very few occurrences of repetitions of length greater than 4). The bar plots in Figures S1B and S1C indicate the average experimental probability of the sequence across sessions (and its standard error), and the red circles connected by lines indicate the probability expected by chance.

As can be seen, the probability of long sequences of defection is lower than expected by chance, and the probability of long sequences of cooperation is higher than expected by chance. In other words, the monkeys had a tendency to avoid long repetitions of defection, but had a slightly higher tendency for long sequences of cooperation. Intuitively, this seems like a sensible approach in a game like iPD where the highest reward possible would be if both players would consistently cooperate. Alternatively, it is possible that the monkeys used long sequences of cooperation to lure the other monkey into cooperation in order to then defect to increase their reward.

Evaluating Behavioral Strategies during iPD

Examining the behavioral strategy followed by the monkeys, note that from a theoretical perspective, all iPD player strategies considered for a memory length of 1, are defined by the conditional probability of selecting cooperation given the outcomes of the previous trial (see e.g., Nowak and Sigmund, 1993). Specifically, strategies are defined by the four conditional probabilities of selecting cooperation following the four trial result options: mutual cooperation (both monkeys cooperated), self-defection and other-cooperation, self-cooperation and other-defection, and mutual defection. Figure 2B depicts the probability that the monkey cooperated following these four outcomes which essentially define the possible strategies (note the right most bar is the unconditional probability of cooperation). For instance, under the above definitions, a pure tit-for-tat strategy (TFT) would result in the four conditional probabilities being [1 1 0 0], respectively, and a win-stay-lose-switch (WSLS) results in the four conditional probabilities being [1 0 0 1].

Accordingly, for each session, we quantified the fraction of trials consistent with the monkey employing each of these principal strategies. To determine the statistical significance of this probability, we generate 5000 control self and other full session surrogate-choices, matched to the number of trials per session, in a Monte Carlo fashion considering the 0^{th} order statistics. This probability was fixed to the experimentally found average of 35.7% for cooperation and 64.3% for defection. We then compare the actual data value of the probability of following each strategy to the control distribution.

The average probability of following the TFT strategy was 0.48 ± 0.035, compared to 0.39 in the control distributions, and the average probability of following the WSLS strategy was 0.40 ± 0.036, compared to 0.34 in the control distributions (Figures 2C and 2D). In other words, the monkeys followed these strategies more closely than would be expected by chance (Monte Carlo estimated p-value: TFT, $p < 0.0002$, WSLS, $p < 0.0028$). In the individual sessions, we find that for the TFT strategy, the experimental probability of following the strategy was beyond the 95^{th} percentile of the control, Monte Carlo distribution for 4 out of the 7 sessions. For the WSLS strategy, in 2 out of the 7 sessions the probability of following that strategy was beyond the 95^{th} percentile of the control, Monte Carlo distribution. Note that, since this study focused on the electrophysiological responses during the iPD task, a more in-depth analysis of these unique behavioral strategies would require acquiring a more specialized behavioral dataset.

Evaluating the Understanding of Social Context with a Computer Opponent Control

To assess whether the monkeys were influenced by the social context of the task, i.e., playing against another monkey, we used a control in which we replaced the opponent monkey with a computer opponent mimicking the monkey's behavior. In two different experiments, we tested the monkey's behavior while the computer opponent was programmed to use different strategies based on the monkeys' behavior when playing against a monkey opponent. In both experiments we found significant differences in monkey behavior between monkey and computer opponents. Specifically, we find an overall decrease in the probability of cooperation, and did not observe any increase in cooperation following mutual cooperation, as was evident when playing against the various monkey opponents (see main text).

In one experiment, the probability of the computer's choices was determined by the average choice probability of the monkeys, i.e., a fixed 34.7% probability of choosing cooperation on a given trial. The two monkeys from which we obtained neural recordings in the main task played a total of 1,722 trials during this control. Here, they selected cooperation on 21.9% ± 1.9% of trials when playing against a computer opponent. This was significantly lower than when they played each other (Chi-square = 124.65, df = 1, $p < 10^{-28}$).

In the other experiment, the probability of the computer's choices depended on the previous trial outcome and additionally accounted for all possible past mutual selections (i.e., strategies). Here, probabilities were determined according to the history-dependent choice probabilities of the two monkeys (i.e., 62.1% following mutual cooperation, 35% following opponent cooperation and self-defection, 27% following opponent defection and self-cooperation and 25% following mutual defection on a trial-by-trial basis). The two monkeys played 1,511 trials during this control. Overall, they significantly cooperated less against a computer opponent (19.1% ± 3.9%) compared to when they were playing each other (Chi-square = 161.73, df = 1, $p < 10^{-36}$). In addition, no significant increase in cooperation was found following mutual cooperation despite the computer fostering such interactions by elevating its own cooperation probability (14.5% ± 3.0%, Chi-square = 1.05, df = 1, $p = 0.31$; unlike when they played each other). This was similarly true in the fixed probability experiment described above ($p = 0.38$; note similar p-values were obtained for other history dependent probabilities, not shown), but was less surprising in that case since the computer did not modulate its own choice probability according to the preceding trial outcome.

In summary, the monkeys' behavior depended on whether they were playing with another monkey opponent, with a clear preference for cooperation in a social context. This was further corroborated by the behavioral findings when the monkeys played each

other in the separate rooms control (see main text), and taken together strongly indicate that the monkeys understood the social nature of their interaction.

Evaluating Understanding of the Payoff Matrix Using a Sequential Control Version of iPD

In a final set of behavioral control experiments, the performance of the monkeys was evaluated on a sequential version of the task, whereby they were able to observe each-other's concurrent choices upon selection, using the same payoff matrix used in the main task (Figure 1B). Here, the monkeys were given the same sequence of image presentations and delays (as described in the method section), but the choice of each monkey was underlined for 500ms with a small bar to indicate what they had chosen. Therefore, within a given trial, the second monkey in the sequence would be able to see what the first monkey in the sequence had chosen before his turn to make a selection. Similarly to the main task, on each trial the order in which the monkeys played was determined randomly. To clarify, in such a setup, if the first monkey chooses to defect, the possibility of mutual cooperation outcome in the given trial is eliminated and the appropriate response of the second monkey is to defect as well. Otherwise, he would get less reward cooperating (1 drop of juice) than defecting (2 drops). For the purpose of this test, the monkey that chose to defect as a first player always got a fixed amount of reward regardless of the other monkey's following selection (2 drops). The remainder of the payoff matrix was the same as the main task. To this end, the four monkeys were tested on a total of 970 trials. Of these, the monkeys chose defection as first actor in 48.5% ± 7.7% of the trials. When taking into account only these trials, the monkeys successfully maximized their reward by selecting to defect as second actors on 90.7% ± 2.2% of the trials. Note, that choosing to cooperate after the other was known to defect would lead directly to a personal loss and would not serve to foster future cooperation since the other chose to defect, not to cooperate. We therefore concluded that the monkeys understood the reward contingencies of the game and moved on to main task recordings.

Regression Analysis Results

Linear Stepwise Regression

We used a step-wise linear regression analysis to examine neural responses to the different features of the iPD task. For each trial, we considered the following parameters: monkeys' current own selection (self-current; SC), the opponent's current trial selection (other-current; OC), monkey's own selection in the previous trial (self-past; SP) and the opponent's selection on the previous trial (other-past; OP). Here, we focus on neuronal activity in the 500 ms period before response selection (i.e., choosing cooperation versus defection) and during the 500 ms period *after* selection, referred to as 'pre-selection' and 'post-selection' respectively.

Since the task involved the interplay between the decisions of both monkeys' current and prior decisions, we first determined the level to which these four parameters were interdependent by calculating the correlation coefficients between them. As expected, we find non-zero, but weak correlations between the above parameters, which allows us to move on with the analysis without the need to eliminate any of them ($|r| < 0.185$, See correlation matrix in Figure S2A).

A total of 185 neurons had a significant difference in activity between cooperation and defection for at least one of the model parameters ($p < 0.025$). Table S1 summarizes the model results as percentage of total neurons. As seen in the table, combining the fraction of neurons which responded to the difference between cooperation and defection during pre-selection period and those which responded during the post selection period results in a sum only slightly higher than the fraction of neurons responding to either pre or post decision, i.e., the fractions in the third row. In other words, overall there was little overlap between these two groups of neurons.

Neuronal Modulation by Past Decisions

In addition to current decision main effects reported in the main text (i.e., those relating to the two monkeys concurrent decisions to cooperate or defect), we found neurons that respond to their past decisions. In total, 14.6% of task responsive neurons were modulated by the monkey's own previous choices (SP), demonstrating a significant difference in firing activity during current selection when the monkey selected cooperation versus defect on the previous trial. This constituted a 1.91 ± 0.13 fold change in absolute activity when considered across all such neurons (Figure S2B).

Similarly, 13.5% of task responsive neurons were selective to the other's previous choices (OP), demonstrating a significant difference in activity during current selection when the other monkey selected cooperation versus defection on the previous trial. This constituted a 1.84 ± 0.15 fold change in absolute activity when considered across all such neurons (Figure S2C). Note that a substantial amount of time passes between the current and past trials (10.75sec ± 1.35, mean ± std) and, therefore, these effects are not merely a trivial consequence of neural summation.

Finally, only 2.2% of neurons that encoded the opponent's current decisions also encoded the opponent's past decisions (4.4% expected by chance level), indicating that largely distinct populations encoded other's current decisions and other past decisions. Similar partitioning between neurons encoding monkey's own current and past decisions is reported in the main text and in Figure 3D. Figures S3B–S3G depicts neural partitioning in different exemplary conditions (see "Population partitioning and neuronal selectivity" section below, and Figure S4 for single neuron examples encoding each of the main behavioral terms.).

Neuronal Modulation by Current Decisions Based on Past Decisions

The observed modulation by past decisions was not a simple read-out of past selections, but rather reflected the fact that some neurons' activity encoded current decisions differentially depending on the recent history of choices. For example, choosing cooperation despite the other choosing to defect on the previous trial lead to different responses compared to when the monkey chose cooperation when cooperation was previously established. Accordingly, as reported in the main text, we found neurons with significant interaction terms corresponding to such first order history effects. 13.0% of neurons responded differentially to the monkey's choice on the current trial depending on the opponent choice on the preceding trial (SCxOP, $p < 0.025$, linear stepwise regression, Figure 5E).

Similarly, 10.3% of neurons responded differentially to the monkey's own choice on the current trial depending on their own choice on the previous trial. Another 11.4% of neurons responded to the opponent monkey's current choice given the recorded monkey's own past selection, and another 11.4% responded to the current opponent monkey's choice given his own past choices. Note that the least common interaction terms in the regression model were those that contained interactions between the two monkeys at the same point in time, e.g., SCxOC.

Population Partitioning and Neuronal Selectivity

Most of the task-responsive neurons' variance was explained by only one significant main effect (or interaction) term. Figure S2D below displays the distribution of number of significant terms per neuron, pre-selection and post-selection. To dissect the contribution of individual terms, we calculated the distribution of model terms separately and grouped the neurons according to the specific term that explained the most variance. For example, the distribution of number of terms for all neurons for which most significant term was SC, term # 1. As can be seen, the distributions are qualitatively similar for all terms and no one term dominates the distribution above (Figure S2E).Furthermore, we examined the two-way partition of neurons between different parameters. For each pair of interactions, we present a scatter plot of the regression p-values for all neurons that had a significant term in at least one of the two parameters (Figure S3). Neurons that responded to both terms are located in the bottom left quadrant of the plot. Additional partitioning graphs are presented in the main text, Figure 3D.

Neuronal Response Properties Analyses Complementing Stepwise Regression

Akaike Information Criterion

Our stepwise regression clearly showed encoding of the other's concurrent decisions in a large fraction of the neuronal population. To shed additional light on the encoding properties of the neurons we recorded, we next performed additional Akaike Information Criterion (AIC) analysis and determine the normalized Akaike weights of each of the possible models. Here, we considered all the possible combinations of model terms and compared all these to the constant model (i.e., no behavioral terms, just a constant in the regression). Based on this new analysis, we obtained similar results to those presented with the stepwise regression above.

Specifically, we count a neuron as being responsive for a given term (e.g., self-present action) if the normalized Akaike weight of a model including this term was greater than the weight of the constant model. We then collect the data to yield the table of model percentages (see below). Finally, to help visualize this data, we plot as a heatmap the normalized Akaike weights for all models and all neurons. To assist in observing the salient components, we subtracted the weight of the constant model and set all the values below that to zero, thereby avoiding clutter in the visualization. We then sorted the neuron matrix according to the maximal (across all models) weight and present this sorted matrix in Figures S5B and S5C. The x axis is sorted according to the number of non-zero terms in the model and then their position. Table S2A summarizes the percentage of neurons that were better explained by each of the model terms than by the constant model. As can be seen in Figures S5B and S5C and Table S2A, the term predicting other concurrent decisions is the most frequent term and it is higher in the post-selection than in the pre-selection period. In general, these relative percentages are qualitatively similar to those previously reported in the main text for the stepwise regression analysis.

We next performed a similar analysis, but now including all interaction terms. Note that this entails a very large number of models, since all possible combinations of the 4 main effects and 6 interaction terms result in 1024 possible combinations (including the constant model). As before, we calculate the AIC values for each of these combinations for each neuron, and then normalize these values into AIC weights. For each neuron, we then calculate the particular model that had the highest weight and, using that, calculate the probability of the neuron's chosen model to contain each of the terms. We note that there is some concern that due to considering such a large number of possible models, this distribution across the population would be noisy. For that reason, we chose only the model that had the maximal AIC weight, and not all the models that had an AIC weight above the constant model as we did when considering only main effect terms. Nonetheless, note that this does not completely eliminate the concern that these results are contaminated with noise from examining such a large number of models for each neuron.

Figures S5D and S5E plots the normalized Akaike weights for all models and all neurons as a heatmap, after subtracting the weight of the constant model and setting all values below that to zero, thereby avoiding clutter in the visualization. We then sorted the neuron matrix according to the maximal (across all models) weight and present this sorted matrix. The x axis is sorted according to the number of non-zero elements in the model and then their position. We term this order model combination index. For instance, model combination index 1 is the vector [1, 0, 0, …., 0], model combination index 2 is [0,1, 0, …., 0], model combination index 11 is the vector [1,1,0,…,0], model combination index 12 is the vector [1,0,1,0,…,0], and the final model combination index, 1023, is [1,1,1,1,…,1]. We then summarize in Table S2B the percentage of neurons that were better explained by each of the model terms than by the constant model.

Here, again the distribution across the population *closely matched* what we found in the regression analysis reported in the main text. Namely, we found a large number of other-predicting neurons during the post-decision period, and their increased prevalence in this period compared to the pre-decision time (Table S2B; Figures S5D and S5E).

Clustering in Mixture of Regressions Models

Re-examining the entire dataset from the perspective of unsupervised learning, we next fit a mixture of regressions model to our data. Under this model, the conditional distribution of the firing rate of the i^{th} neuron in the population, y_i, given the trial variables x, is a mixture of normal distributions:

$$y_i|x \sim p_1\mathcal{N}\left(x'\beta_1, \sigma_1^2\right) + \ldots + p_k\mathcal{N}\left(x'\beta_k, \sigma_k^2\right)$$

Where the regression coefficients β_k and the variance σ_k^2 vary between each of the components 1,…,k, but are fixed across the population (i.e., not generated individually for each neuron).

We ran the mixture model for different number of clusters (linear regression components), k, from 2 to 10 (since 1 cluster would be identical to regular linear regression) while allowing all main effect variables, *x*. For each such model, we repeat the estimation multiple times and find the mixture component probabilities, *p*, as well as the coefficients, β.

By and large, the model was in good qualitative agreement with the main results of the paper and, in fact, demonstrated an even stronger representation of other-predictive neurons than in the step-wise regression analysis. Specifically, for all component clusters we find a sizeable cluster of neurons in which the most prominent variable of the linear regressions is the other monkey current response variable. In other words, there was a predominant (and, in fact, even more prominent) representation of other-predictive cells when compared to the step-wise regression analysis. A second common large cluster is one that represents the dependency of the firing rate on the monkey's own current response, although these responses were often found with weak conjunction with additional task parameter responses (such as other-past, or self-past decisions).

We considered models of 2-10 components (the "k" in the equation above). Intuitively speaking, each component potentially captures a specific linear relationship between the firing rate of a group of neurons and the behavioral variables. We start from two components since a single component would just be regular linear regression. The maximal number of 10 components was chosen for three reasons. First, the gain in log-likelihood appeared to be saturating when considering > 10 components as seen in Figure S6A. Second, starting from nine components and above, components that had a very small probability of association, namely < 1%, i.e., one or two neurons, began to become more commonplace, as seen in Figure S6B. Finally, for all of these models, a large component that was selective mainly for the other's decision was found, further supporting our main step-wise regression finding of neurons encoding the other's yet unknown concurrent decision (Figure S6C). Encoding of self-decision was always an additional component in the model, though at times more mixed with other decision variables than the more pure mode encoding other-decision.

We next performed the same form of mixture regressions analysis, only now considering both the 4 main behavioral variables and their 6 interaction terms. As before, we consider models with 2-10 components (note that since it wasn't obvious that the model likelihood was saturating we double checked the likelihood with models that include 11 and 12 terms, Figure S6D). Results were qualitatively similar to the analysis performed on the main variables only, as well as to the analyses reported in the original manuscript (Figures S6D–S6F). Namely, a component that had strong weight for the other's unknown choice was present in all models and was the most common component.

In summary, the mixture of regressions model approach recapitulates the main step-wise regression results reported in the main text, and concurs with the AIC analysis reported above. Specifically, these new analyses demonstrate; (1) that a predominant fraction of the population consists of 'other-predictive' neurons and (2) that these neurons mostly encoded little concurrent information about other task variables such as self-decision.

Choice Probability Index Analysis

To further elucidate the nature of single neuron response properties when encoding self and other decisions to cooperate or defect, we performed a Choice Probability (CP) Index analysis which correlates the trial-by-trial activity of each neuron to the monkey's choice to cooperate or defect. To this end, we calculated the Receiver Operator Characteristic (ROC) curve (Britten et al., 1996; Shadlen et al., 1996), separately for each neuron and for each time segment, i.e., pre- versus post- decision. Namely, we treated defection as the signal to be detected (this is somewhat arbitrary but was chosen since there are more defection trials). For each value of the firing rate, we assigned the predicted response of defection to each trial for which the firing rate was higher than this value and then compared it to the monkey's choice on that trial, yielding the fraction of true positives (Hits) and false positives (False Alarms). Testing all firing rate values, from that which yielded 0 fraction of both true positive and false positive to the maximal value that yielded a value of 1 for both false positive and false negatives, we obtain the ROC curve. We then calculated the area under the ROC curve by the sum of the trapezoids defined by the different firing rate values. This area is defined as the CP index.

Note that a CP value of 0.5 indicates no information yielded by the firing rate of that particular neuron regarding the monkey's choice. Values above 0.5 indicate preference for defection (i.e., higher firing rate for defection) and values below 0.5 indicate preference for cooperation. Values of 1 or 0 would indicate perfect correlation between the firing rate and the trial-by-trial choice of the monkey. We performed this analysis for all 4 task parameters (self-current choice, self-past choice, other-current choice, other-past choice) to be predicted by the firing rate.

In order to determine the significance of the CP values, we perform a permutation test (Britten et al., 1996). Namely, we generate 5000 random permutations of the trial identities for each neuron under each task parameter, calculate the ROC value for all permutations and determine the number of permutations that had values as far from 0.5 as those found for the actual experimental data. This number is divided by the number of random permutations to yield the approximate probability of finding such a value by chance. We determine a neuron to have a significant choice probability index if the permutation test yielded a probability lower than 0.025.

When testing all 363 recorded neurons, we found that 20.9% of the cells had a CP index significantly above chance for predicting the other's-choice (during post-selection period), while 14.0% of cells had a significant CP index for encoding self-decision (during pre-selection period). To facilitate a more direct comparison to the stepwise regression results, we also tested what fraction of the reported task-responsive neurons had a significant choice probability index. The results we found using CP index analysis closely match the stepwise regression results reported in the main text (35.7% of neurons had a significant CP index for encoding the other's choice post-selection, and 21.6% had a significant CP index for encoding self-decision pre-selection). Finally, to allow visual

impression of this data, we display the histograms of the choice probability index for all recorded neurons in Figure 3E. Red bars indicate significate CP index value neurons. Examples of ROC curves of single neurons for all task parameters are displayed in Figure S5A.

These analyses, together with the AIC and mixture of regressions analysis results, suggest that our findings based on the neuronal data stepwise regression were reproducible across multiple statistical methods.

Dissociating Encoding of Decisions from Encoding of Expected Reward

An important feature of the iPD framework is that it allows to explicitly dissociate one's decisions from expected reward. As discussed above, the monkey's own choice cannot guarantee a high or low reward, and the prediction of one's own reward requires one to accurately predict the opponent's yet unknown selection. Nonetheless, to more directly demonstrate that other-predicting neurons did not encode the monkey's own expected reward, we perform four new separate analyses. We specifically analyze neuronal responses based on differences in reward contingencies. For example, in a case where the other monkey chooses to cooperate, the monkey's own reward differs according to his own selection. If the monkey himself chooses to cooperate, his expected reward will be of 4 drops, but if he defects, the expected reward increases to 6 drops. In other words, there is a difference of 2 drops in the expected reward contingent on self (but not other) action. Importantly, this difference in the amount of reward is comparable to the difference in the monkey's own reward if the other chooses to cooperate versus defect, which is the condition used to define the other predicting cells. Additional comparisons can also be made between maximal and minimal net expected reward as discussed further below.

Analyzing neuronal responses during the decision period, we find that the firing rate modulation of other-predicting neurons did not encode the monkey's own expected reward. Furthermore, comparing neuronal responses during the decision period to the feedback period, when the monkeys already knew which reward they will receive, modulation of firing rate activity in other predicting cells was significantly less sensitive to the expected reward than the recorded population. These differences held for numerous high versus low reward combinations (see further below) indicating that lack of reward selectivity in other-predictive neurons was not simply due to differences in net amounts of reward. Therefore, while combining all heterogeneous cell types in the cingulate may, in principal, allow one to predict the monkey's own reward, ''other-predictive'' neurons are a fundamentally distinct class of cells selective for encoding the other monkey's yet unknown concurrent decision. As we further demonstrate in the main text, these neurons also encoded no information about the monkey's own decision and were also uniquely sensitive to social context compared to other population cells.

First, we compared the different self-reward per choice. We find that none of the other-predicting neurons displayed a significant modulation for all possible differences in self-reward (two tailed Welch's t test, corrected for unequal variance and sample size, $p > 0.05$). This held true also when using a non-parametric test that compared the distribution of neuronal responses across trials (Kolmogorov-Smirnov, $p > 0.05$). These analyses were conducted during the monkey's own post selection period, as in the corresponding analysis in the main text. Specifically, when evaluating trials in which the other chose to defect, only 4% of cells displayed a significant difference in the firing rate when the monkey's own expected reward differed, i.e., when the monkey's own selection was cooperation versus defection (two-sample *t*test, $p < 0.018$). Similarly, when the other monkey cooperated, only 4% of cells, non-overlapping with the 4% of neurons reported in the preceding comparison, had significant modulation in firing rate when comparing self-cooperation to self-defection (two-sample *t*test, $p < 0.024$). Therefore, neurons encoding the other's decisions did not demonstrate a firing rate modulation consistent with encoding the monkey's own expected reward. As we show further below in the fourth analysis, this lack of concurrent reward modulation was not simply due to differences in raw amount of reward (e.g., 1 versus 2 over 2 versus 6).

Second, to further describe the distinction between the activity of other-predicting neurons and encoding of self-reward, we compared firing rate modulation in the same conditions as above (i.e., comparing self-defection to self-cooperation, once conditioned on the other cooperating, and once on the other defecting) but now considering the distribution across the full population, and not just the other predicting neurons. To allow for intuitive visualization, we calculate for each neuron the average difference in firing rate between these two conditions. This difference in raw firing rate (delta) was normalized by dividing by the neuron's base line firing rate to account for differing firing rates between neurons and allow comparison across the population. This gives the size of the difference (delta) in terms of the ratio of the increase/decrease to that of the baseline firing rate. That is, a value of one corresponds to an increase in the firing rate of 100% compared to the baseline firing rate, i.e., a difference of 2 Hz between the conditions for a neuron firing at 2 Hz would be marked as 1. The resulting distribution of the absolute values of the normalized firing rate difference across the neurons are then plotted as the histograms and scatter plots of Figures 4A–4F. Similar results (described below) were also found when data was normalized not according to the baseline, but according to the higher firing rate value of the two conditions considered.

As can be seen in Figure 4A, considering the other-predicting neurons, the distribution of the average difference in firing rates between trials in which the other chose to cooperate versus defect, is centered far away from zero, i.e., they show significant modulation according to the other's yet unknown decision. In addition, when considering the full distribution of all recorded neurons, the other-predicting neurons highly significantly differed in their distribution from that of the full population (Wilcoxon rank sum test, $p < 10^{-15}$). Indeed, as can be seen, the other predicting neurons comprise the tail and not the center of the firing rate differences across the full population. Now, if these other-predicting neurons encoded the differences in expected reward, one would expect to see a significant modulation across other cases of differences in expected reward as well. However, this is not what we find.

For the same other-predicting neurons, when considering only trials in which the other monkey chose to defect, the normalized difference in firing rate between trials in which the monkey's own decision was to cooperate versus defect, was closer to zero-mean

(median 0.17, Wilcoxon rank-sum test, $p < 10^{-5}$, Figure 4B). When further comparing these other-predicting neurons to the population, other predicting neurons were actually less modulated in this case of trial contingencies than the general population (population median 0.21, Wilcoxon rank-sum test, $p < 0.05$, Figure 4B). That is, these cells showed little modulation in firing rate despite the fact that the monkeys expected to receive double the reward in one condition versus the other. In other words, while many neurons in the population did encode the monkey's own expected reward, these were not the other-predicting neurons.

A similar picture holds when considering only trials in which the other monkey cooperated (Figure 4C): the distribution of other predicting neurons was centered close to zero and did not significantly differ from the population distribution (median of other predicting neurons = 0.23, population median = 0.24, Wilcoxon rank-sum test, $p = 0.62$). Therefore, these neurons did not strongly encode the difference in the monkey's own expected reward. Once again, as can be seen from the population distribution, there were many neurons that did encode this difference in reward, but these were not the other-predicting neurons. All in all, this analysis shows that the response of the other-encoding neurons is not consistent with a simple parametric encoding of expected reward.

Third, in another independent line of analysis, we examined neural activity during the feedback period (Figures 4D–4F). Here, when the choices of both monkeys are revealed, but the reward has not yet been delivered, the expected reward is known. Therefore, if other-predicting neurons merely encoded expected reward, the difference in firing rate between trials which result in low or high reward, should be greater during the feedback period, than during the selection period (which was analyzed above and in the main text) in which the reward may be expected to be either high or low, but with less certainty. For each of the other-predictive cells, we calculate the difference in firing rate during both the selection and the feedback period. Then we normalize these differences in firing rate to the baseline firing rate of each cell in order to allow for direct comparison. Finally, we compare the values of the normalized difference in firing rate between the two time periods (selection versus feedback). Using a paired t test on the value of the difference between low and high reward, we find no significant difference in firing rate between the two time periods (paired t test, $p = 0.22$, Figure 4D). Therefore, the firing rate of these cells did not change according to the differences in expected reward.

Next, we considered all neurons, not just those labeled as other-predicting neurons. Here, we did find a highly significant increase in the difference in firing rate during the feedback period compared to the selection period (paired t test, $t(362) = -4.7$, $p < 0.00004$, Figure 4E). As seen in the paragraph above, the other-predicting neurons actually showed an opposite trend, of reduced difference in the firing rate in the feedback period. Overall, for the population, the mean normalized difference in firing rate increased from 0.28 ± 0.02, 0.35 ± 0.02 between selection to feedback. In contrast, for other-predictive neurons, the mean difference in firing rate slightly decreased from 0.60 ± 0.05 to 0.55 ± 0.07. This indicates that direct reward feedback enhanced the differential response of many cingulate cells to high versus low self-reward, but not of other-predictive neurons.

Fourth, we tested whether this lack of reward selectivity by other-predictive neurons could be affected by differences in raw amount of reward. Toward this end, we perform the comparison specifically for the two conditions that resulted in the highest reward differential, i.e., when the monkey chooses to defect, and then comparing trials in which the other monkey either cooperated or defected (that is, 6 versus 2 drops of juice, respectively). However, once again, we found no significant difference between the selection time period and the feedback time period for the other-predicting neurons (paired *t*test, $p = 0.13$, Figure 4F).

In summary, looking at multiple possible angles trying to find reward-related modulation, we observed that other-predictive neurons encoded little or no information about the monkey's own expected reward; either before they know which reward they would receive or after. However, this was not true of many other cells in the population which were highly sensitive to differences in reward when the monkeys already knew their selections, consistent with prior reports (Chang et al., 2013; Yoshida et al., 2012).

Even when using an over-permissive definition of reward encoding, whereby a neuron is counted as encoding reward if it responds to any difference of reward contingency, even if it lacks significant responses to other contingencies, we find little overlap between encoding of decision and encoding of reward during selection period. Namely, we find that, of cells that encoded SC decision, only 8.0% also encoded self-future reward (i.e., same-trial reward not yet received), and of cells that encoded OC decision, 6.3% encoded other-future reward. Similarly, there was little overlap between neurons encoding the monkeys' current decisions and signals related to known past reward. Of cells that encoded SC decisions, 6.0% encoded self-past reward, and of cells that encoded OC, 4.7% encoded other-past reward (step-wise linear regression on reward). Finally, we find little overlap between cells encoding past decisions and past reward. Of cells that encoded SP decisions, 6.7% encoded self-past reward, and of cells that encoded OP, 8.0% encoded other-past reward. Overall, the overlap between cells encoding decision and reward across all these intersections was at chance level (Chi-square = 0.0, df = 1, $p > 0.9$).

Reward Neurons Modulate Their Activity during the Other's Receipt of Reward

Demonstrating above that other-predictive neurons did not encode expected reward, we next tested the response of the general population of neurons to the other's receipt of reward. Toward this end, we performed the same step-wise regression analysis used in the main text for the choice related responses, only now testing the neurons firing rate around the time of self and other's receipt of reward (i.e., pre and post reward initiation). Note, that the monkeys received their reward sequentially in an order that changed randomly from trial to trial to allow separation, and therefore in half of the trials the recorded monkey received his reward first. For the purposes of this analysis, we combined the trials in which the monkey was first or last to receive his reward. We find a total of 102 out of 363 recorded neurons that had a significant response to reward ($p < 0.025$) as following: during the pre-reward period (500ms, following feedback presentation when monkeys already know how much reward will be given to them), 58 neurons encoded the monkeys own reward and 21 neurons significantly encoded the other's reward. During the 500ms following reward

initiation, 38 neurons significantly encoded the monkey's own reward and 19 neurons significantly encoded the other's reward. All in all, 34 neurons encoded both monkeys reward. Figure S7A displays peristimulus histograms of an example neuron encoding other reward, but none of the other task parameters.

Behavioral and Neural Prediction Models

Decoding the Opponent's Concurrent yet Unknown Decisions Based on Neural Population Activity

Many neurons encoded the opponent's current decisions at a time in which their selections were still unknown to the recorded monkey. To quantify how predictive these combined responses were of the opponent's choices, we turned to a population modeling approach since such predictive signals may be weakly encoded by individual neurons on a trial-by-trial basis but robustly represented by the combined activities of such neurons across the population. Here, we find that neural population predictions were highly accurate. More remarkably, predictions of the other's yet unknown decisions were significantly more accurate than predictions of one's own known selections.

To quantitatively test the ability to predict the opponent's choices on each trial, we built a linear predictor that used the recorded activity of all neurons. Namely, we found the direction in activity space that maximized the between-choice variability, normalized by the within-choice variability, known as Fisher's linear discriminant. Specifically, the prediction vector, v, is the eigenvector corresponding to the largest eigenvalue of the matrix $S_W^{-1}S_B$, that is:

$$S_W^{-1}S_B v = \lambda v$$

Where S_W is the within group scatter matrix and S_B the between group scatter matrix. This vector defines a projection from the activity of all neurons into a scalar value. This scalar value is then compared to a threshold, θ, and a trial was predicted to be a cooperation trial if the projection was above θ and a defection trial if the projection was below θ.

We divided the data into 75% training data from which the prediction vector v was determined along with a threshold, θ. The accuracy of this prediction was then tested on the 25% of the data that was not used from training. This procedure was repeated with a different random choice of the 75% of the data 50 times and averaged to mitigate variability that might arise from an unusual choice of the training/test data. The bar plots in Figures 5C and 5D show the projection value of the population activity on this direction in the activity space in each trial, for test, not train data, cooperation trials in red, defection in blue.

We find that cingulate populations predicted up to 66.1% ± 0.9% of the recorded monkey's *own* current choices before selection and 65.6% ± 1.0% after selection (mANOVA, $p < 0.0002$ and $p < 0.004$ respectively). When considering the opponent monkey's choices, however, cingulate neurons in the recorded monkey could predict up to 57.4% ± 1.2% of trials before their own selection and 79.4% ± 1.1% after (mANOVA, $p < 0.035$ and $p < 0.00001$ respectively). Remarkably, this post-selection activity was significantly more predictive of the opponent's unknown choices than pre-selection activity was of the acting monkey's own choices (paired t test, $p < 0.00001$, Figures 5C and 5D). Note that, at all time periods, the monkey did not know what the opponent had chosen until the feedback period, at least 0.5 s later (if first actor, or an average of 2.3 s on the other half of the trials), and that predictions kept their accuracy when taking into account only half the trials when the monkey played either first or second (see main text).

Finally, to help visualize the separation between the different parameters afforded by the neural data, we reduce the dimensionality of the data to three dimensions using Principal Component Analysis (PCA). We generate an activity data matrix by concatenating the number of spikes of each neuron for each individual trial. Based on this, we calculated the weighted sums of the population activity which captured the most variance of neuronal activity. We then plotted each trial as a point corresponding to the value of the population activity on that trial in the first three principal components. Consistent with the population decoding model results, we observed that the separation between the principal component scores when considering the opponent's current selections was more pronounced than that when considering the monkey's own current selections. This separation was also more pronounced in the post-selection period compared to the pre-selection period (Figures 5A and 5B).

Decoding the Opponent's Current yet Unknown Decisions Based on Prior Behavioral Selections

To search for the possible source of predictive neural activity in the dACC, we tested the possibility that the two monkeys' prior choices were behaviorally predictive of the opponent monkey's current choice. A linear regression model was used to model the relation between past and present behavioral responses of the two monkeys. As with the neuronal decoding, we use a Fisher's linear discriminant approach to predict behavioral choices from the history of behavior. Using the past 4 trial selections of both monkeys as the basis for prediction, we find the weighted linear sum across these terms through the eigenvalue of the matrix composed of the inverse of the within group scatter matrix times the between group scatter matrix (see neuronal predictions for equation). As with the neuronal data, we used a random subset of 75% of the data for training and 25% of trials for validation to decode the opponent's current decision on each trial. These models predicted the opponent's decisions up to 66.0% of trials. Increasing the length of the history beyond 4 trials did not improve model performance.

We also considered other possible modeling approaches to determine if better predictions could be made. Toward this end, we used a locally optimal non-linear classification tree model which considered the past history of the past 4 trials of both monkeys (Breiman et al., 1984). In brief, this model is composed of a series of decision points, in which at each decision point the data is subject to a binary split based on the value of one or more of the predictor parameters. We constructed each tree using a random subset of 75%

of the data for training. The splitting criterion was the Gini diversity index and the tree was constructed until each decision node was pure, i.e., had only cooperation or only defection trials. Then the tree was pruned, i.e., decision branches were removed from the bottom nodes toward the upper nodes, to the level that maximized the cross-validated (within the training data) prediction accuracy. Finally, the remaining 25% of the data was used to determine final prediction accuracy. This model predicted the opponent's unknown decisions in up to 79.8% ± 1.3% of trials. This behavioral prediction approximately matches the neuronal data based prediction presented in the main text and suggests that the information for computing the neuronal prediction may be available in the past history of both monkeys' prior selections. Similarly, the model predicted the monkey's own selections up to 79.9% ± 1.2%. Prediction accuracy was comparable when prediction was based on the four-back history of reward received by both monkeys rather than the history of monkeys' actual selections (76.7% ± 1.6%, 81.0% ± 0.9% for other and self respectively).

Correlating Neural Population Decoding and Behavioral Decoding

To further explore the underlying behavioral basis of population activity predictive of the other's unknown concurrent decisions, we tested trial-by-trial correlations between the neural population predictor and different versions of the behavioral predictor, based on different inputs. Since the behavior predictor outputs the identity of the trial (cooperation versus defection) whereas the population decoder outputs a scalar value (the projections of the firing rate on the discriminant component), we first transformed the population decoder output into a decision (cooperation versus defection) by comparing it to its threshold.

We find significant correlations on a trial-by-trial basis between the population decoder and the behavior decoders based on both monkeys' behavioral history ($r = 0.31$, $p < 0.0003$) and somewhat lower correlations for decoding given both monkeys' reward history ($r = 0.21$, $p < 0.01$). Omission of either monkey's information related to either monkey's past decisions or past reward resulted in largely reduced, non-significant correlations (behavior prediction based only on other's past behavior $r = 0.12$, $p = 0.17$ versus only on self-past behavior $r = -0.02$, $p = 0.60$; behavior prediction based only on other's past reward $r = 0.15$, $p = 0.14$ versus only on self-past reward $r = 0.08$, $p = 0.33$). Note that the same number of data points were used when calculating both versus individual behavioral histories. This indicates that the population neural activity which is predictive of the other's unknown concurrent actions is most related to the past selections of both monkeys. Though non-significant, we note that the correlation between the population decoder and behavioral predictor based on only the other's past selections was higher than a behavioral decoder based only on self-past selections.

Influence of Mutual Past Decisions on Neural Activity

Using linear regression analysis, we have systematically described sub-populations of dACC neurons encoding multiple different facets of the task: First, neurons that encode the current decision of the monkey himself and the expected selections of the other. Second, neurons that encode past decisions of the self and other. Third, neurons that encode current selections modulated by single monkey's past-response, that is neurons that encode self-current selection given self-past selection, neurons that encode self-current given other-past selection, other-current given self-past or other-current given other-past.

Arguably, however, crucial context for a current decision is found in the past decision of both monkeys. In Figure S7B, we considered trials in which the monkey chose to defect after he cooperated in the previous trial. We then separated the trials to those in which the other monkey cooperated on the previous trial (shown in red); that is, the monkey switched response from cooperation to defection following mutual cooperation, and those in which the other defected on the previous trial (in blue); that is the monkey switched response after being defected upon by the other and receiving the least possible reward (displaying lose-switch response). As can be seen, there is a significant difference in firing rate as the monkey defects following these two different contexts. Note that there is a significant difference in firing rate despite the fact that in both cases the monkey performed the same actions two trials in a row (self-defect following self-cooperation), and only the other monkey's previous action differed. Note that when breaking up data this way, fewer trials are available for analysis (in this example, there weren't many trials in which the actor chose to defect following mutual cooperation).

We found 61 neurons that kept track of such current self-decisions modulated-by-mutual-past decision outcomes ($p < 0.025$, with the following iterations: self-defection following self-cooperation, self-defection following self-defection, self-cooperation following self-cooperation and self-cooperation following self-defection). Of these, 56% neurons had either a main effect or interaction effect in the regression model (detailed above).

Stimulation Results

Zero-Sum Game Control

As reported in the main text, we found that stimulation decreased the probability of the monkey to cooperate when the opponent cooperated on the preceding trial, but no effect was observed when the opponent defected on the preceding trial. To further confirm that stimulation did not affect decisions based on any past reward or opponent selection, we turned to a task in which monkey's reward was contingent on the others response but no mutual positive outcome was possible. Unlike iPD, where accumulating net positive mutual outcome is possible, zero-sum games are defined by their Pareto optimality, meaning that the enhancement of profit by one individual always diminishes the profit of others (Camerer, 2003). Here, we examined whether dACC stimulation affected the monkey's decisions based on the zero-sum payoff matrix shown in Figure 7A (see also Extended Experimental Procedures).

We find no effect of stimulation on monkeys' choices when considering either preceding reward or other's preceding selection. Consistent with prior animal and human studies, and as predicted by the optimal strategy (Barraclough et al., 2004; Lee, 2008;

Nash, 1950), the monkeys were equally likely to select either choice within a given trial (total of 2,516 trials, 1,221 without stimulation, decision-ratio 1.08 ± 0.18). Moreover, their decisions were not modulated by the preceding choices of the opponent (1.15 and 1.02, p = 0.19), thereby maximizing their receipt of reward. However, unlike the iPD game, stimulation did not affect the overall decision-ratio (1.10 ± 0.19, p = 0.42; Figure 7B), nor did it have an effect when data was separated according to the opponent's preceding decision (1.01 and 1.19, p = 0.21 and 0.14, respectively, Figure 7C). Importantly, there was no significant effect of stimulation on these responses based on whether the monkey received high or low reward on the preceding trial (decision-ratio of 1.11 and 1.02 without stimulation, 1.19 and 1.11 with stimulation; p = 0.28 and 0.25, respectively; Figure 7D). Therefore, while the dACC may play a role in encoding zero-sum game decisions (Seo and Lee, 2007; Vickery et al., 2011), these findings suggest that the effect of dACC stimulation during the iPD game did not simply effect decisions based on any past reward or opponent response.

SUPPLEMENTAL REFERENCES

Asaad, W.F., and Eskandar, E.N. (2008). A flexible software tool for temporally-precise behavioral control in MATLAB. J. Neurosci. Methods *174*, 245–258.

Breiman, L., Friedman, J.H., Olshen, R., and Stone, C.J. (1984). Classification and Regression Trees (CRC Press).

Nowak, M., and Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature *364*, 56–58.

Seo, H., and Lee, D. (2012). Neural basis of learning and preference during social decision-making. Curr. Opin. Neurobiol. *22*, 990–995.

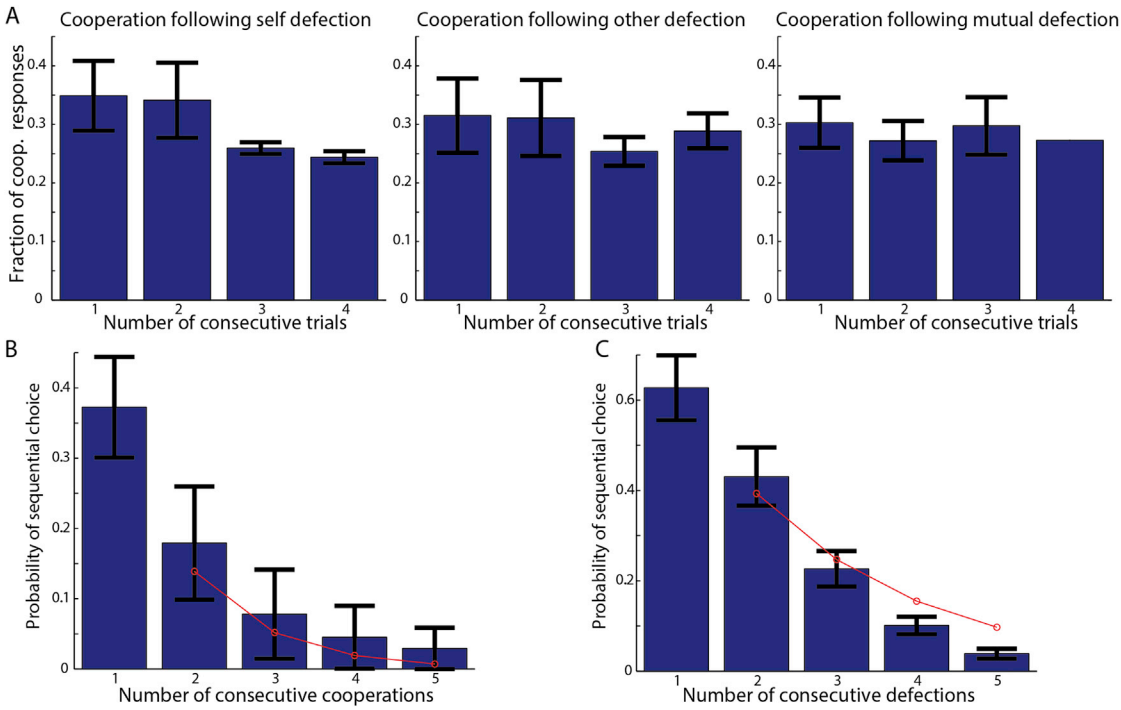


Figure S1. The Effect of Past Decisions on Cooperation, Related to Figure 2

(A) Probability of choosing cooperation following a sequence of consecutive choices for left: sequence of monkey consecutively choosing defection, middle: sequence of other monkey consecutively choosing defection, right: sequence of both monkeys choosing defection. x axis denotes length of past sequence (1-4 trials). Error bars represent SEM.

(B) Probability of sequentially choosing cooperation. x axis denotes sequence length. Bars indicate probability, red circles indicate value expected by chance from zero-order probability of choosing cooperation.

(C) Probability of sequentially choosing defection. x axis denotes sequence length. Bars indicate probability, red circles indicate value expected by chance from zero-order probability of choosing cooperation.

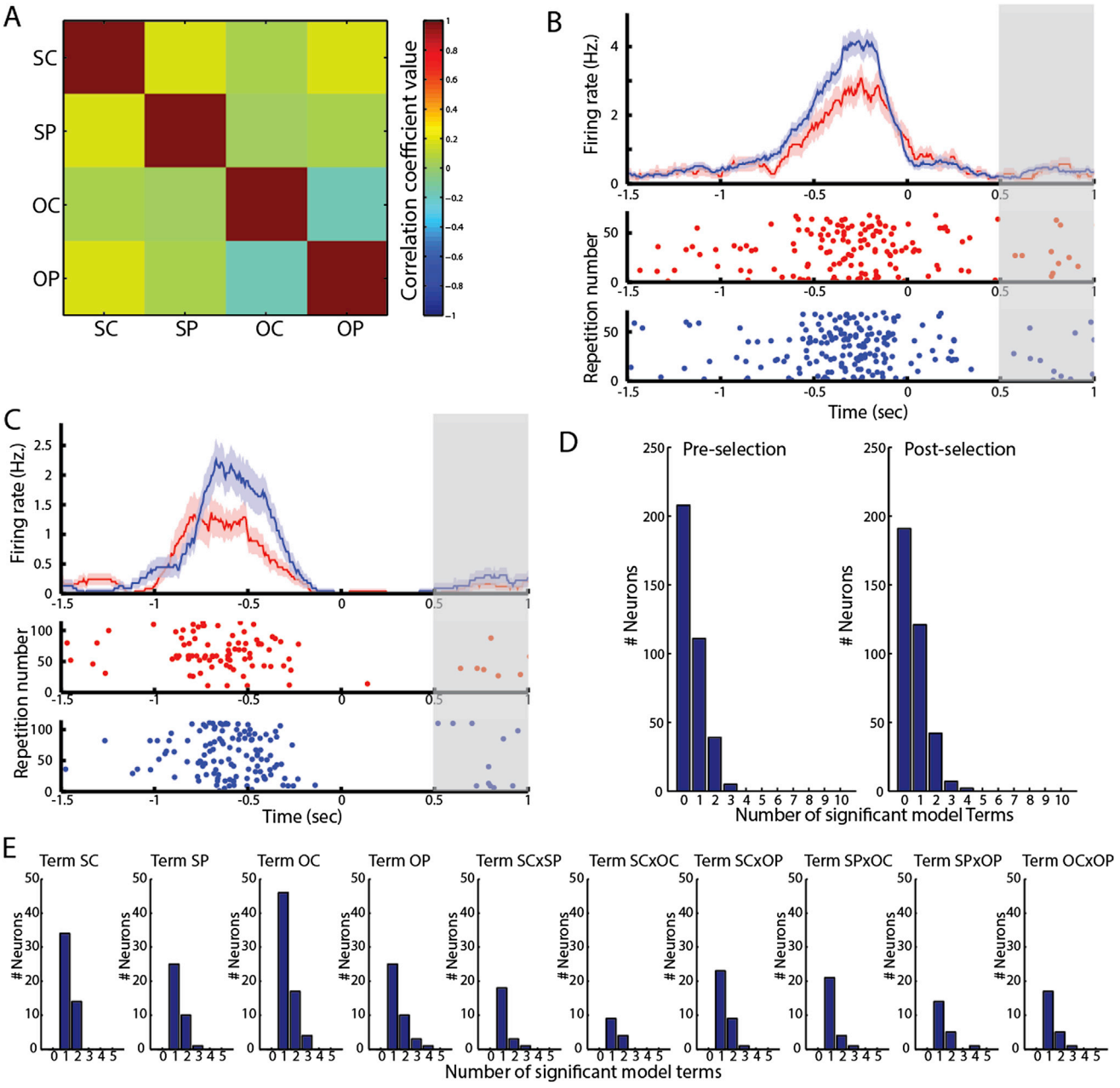


Figure S2. Regression Model Analysis, Related to Figure 3

(A) Correlation coefficient between main effect parameters. Low correlation coefficients were found across all trials between the four main effects denoted as: self-current: SC, self-past: SP, other-current: OC, other-past: OP.

(B and C) Example of single neurons encoding past decisions. Peristimulus histograms as mean firing activity ± SEM and raster plots for individual neurons. Red denotes cooperation trials, blue denotes defection trials. (B) Neuron encoding self-past decisions. (C) Neuron encoding other's-past decisions.

(D) Distribution of number of significant model terms across all neurons. Histogram displays number of significant model terms for pre-decision (left) and post-decision (right) time segments. Only very few neurons responded significantly to more than 2 terms.

(E) Distribution of number of significant model terms according to neuron's most significant term tested in the two 500ms time segments, before and after decision. Histogram displays number of significant model terms, breaking down the population according to the most significant term for that neuron for all terms (from left to right; terms 1-4 are main effects, terms 5-10 are interactions).

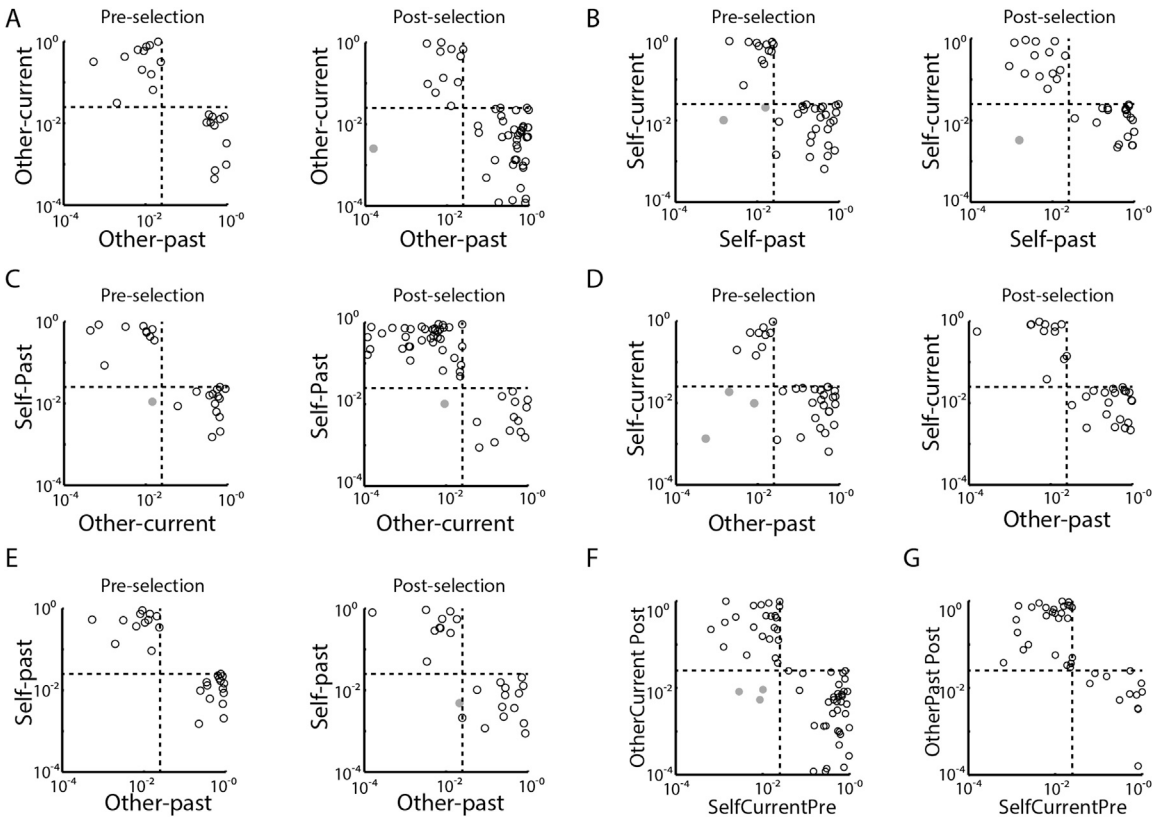


Figure S3. Pairwise Significance Value of Regression Terms Indicates Population Partitioning Based on Self- and Other-Decision Encoding, Related to Figure 3

Log-log scale scatter plots of regression model term p-values plotted against each other. Each dot corresponds to a neuron. Dashed lines indicate significance threshold. Only neurons that significantly encoded at least one of the plotted terms are shown. Gray circles denote neurons that were significantly modulated by both terms. Note that largely distinct populations encoded each two terms. Left plots indicate pre-selection data, right post-selection in (A)–(E).

(A) Opponent's past (known) and current (yet unknown) decisions.

(B) Monkey's own past and current decisions.

(C) Monkey's own past decisions and opponent's current (yet unknown) decisions.

(D) Monkey's own current decisions and opponent's past decisions.

(E) Monkey's own past decisions and opponent's past decisions (See also Figure 3D).

(F) Monkey's current decisions during pre-selection period and opponent's current (yet unknown) decisions during the post-selection period.

(G) Monkey's current decisions during pre-selection period and opponent's past decisions during the post-selection period. Note additional terms all qualitatively similar to (F) and (G) not shown.

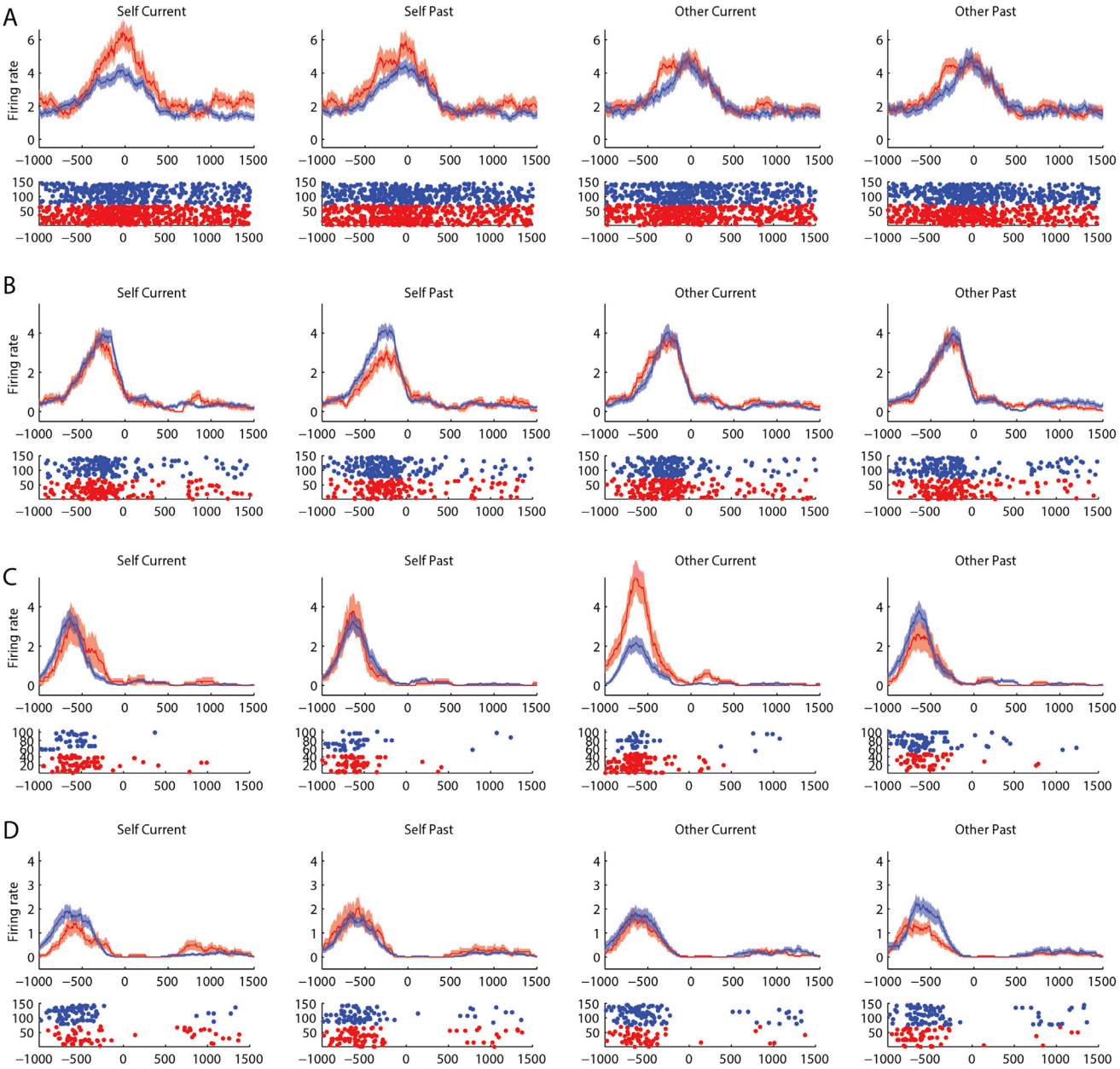


Figure S4. Examples of Neurons Encoding Single Behavioral Terms, Related to Figure 3

Peristimulus histograms as mean firing activity ± SEM and raster plots for individual neurons. Average activity for cooperation (red) and defection (blue). Trials sorted according to different behavioral variables.

(A) Neuron encoding the monkey's current decision.

(B) Neuron encoding the monkey's past decision.

(C) Neuron encoding the other monkey's current decision.

(D) Neuron encoding the other monkey's past decision.

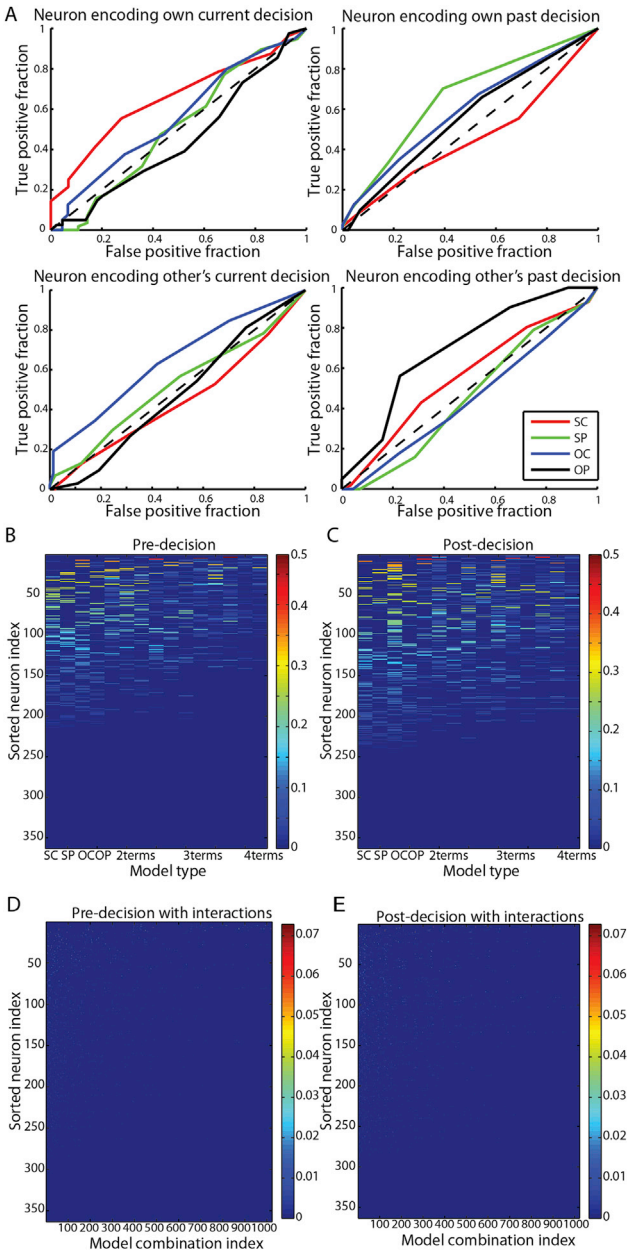


Figure S5. ROC and AIC Analyses Recapitulate Main Stepwise Regression Results, Related to Figure 3

(A) Single neurons reliably encoded behavioral variables on trial-by-trial basis. Receiver-operating-characteristic (ROC) curves from examples of neurons encoding self-current (top left), self-past (top-right), other past (bottom-right) and other-current (bottom left). Color indicates behavioral variable.

(B–E) Akaike Information Criterion analysis. (B) Values of normalized Akaike weights for all combinations of the 4 behavioral terms across all neurons. Constant model value was subtracted and values less than zero set to zero in order to aid visualization. Neurons sorted according to maximal weight in any term. Analysis was performed on pre-decision time period. (C) Values of normalized Akaike weights for all combinations of the 4 behavioral terms across all neurons. Constant model value was subtracted and values less than zero set to zero in order to aid visualization. Neurons were sorted according to maximal weight in any term. Analysis was performed on post-decision time period (note the prominent values for the other-current behavioral term). (D) Values of normalized Akaike weights for all combinations of the 4 behavioral terms and their 6 interactions across all neurons. Total of 1024 terms sorted according to number of terms in model (1 term left, all 10 terms right) and within same number of terms by term order (SC,SP,OC,OP respectively). Constant model value was subtracted and values less than zero set to zero in order to aid visualization. Neurons sorted according to maximal weight in any term. Analysis performed on pre-decision time period. (E) Values of normalized Akaike weights for all combinations of the 4 behavioral terms and their 6 interactions across all neurons. Total of 1024 terms sorted according to number of terms in model (1 term left, all 10 terms right) and within same number of terms by term order (SC,SP,OC,OP respectively). Constant model value was subtracted and values less than zero set to zero in order to aid visualization. Neurons sorted according to maximal weight in any term. Analysis performed on post-decision time period.

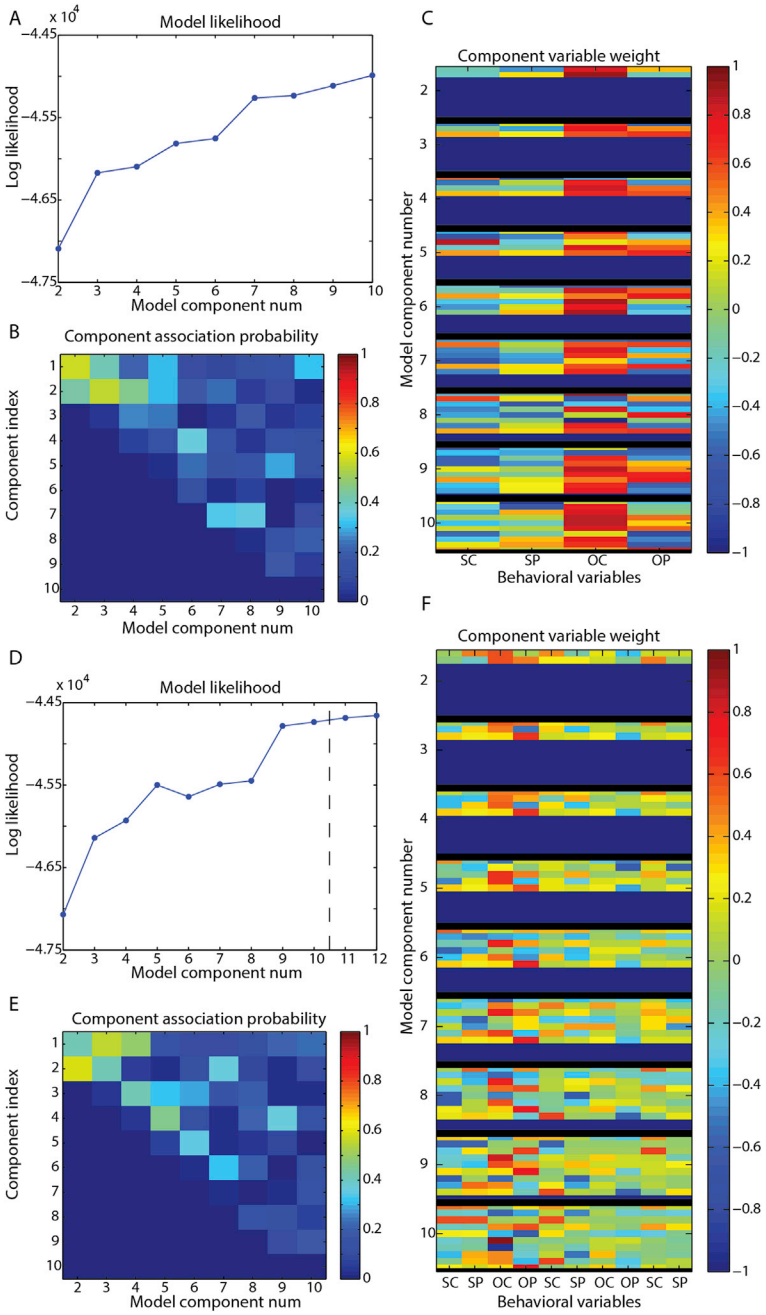


Figure S6. Mixture of Linear Regression Models Analysis Recapitulates the Main Stepwise Regression Results, Related to Figure 3

(A–C) Mixture of linear regressions using main behavioral terms only. (A) Log-likelihood of data for models with different component numbers. (B) Association probability for each component in each model. x axis indicates number of components in model, y axis indicates component index within model. (C) Normalized weight of each of the four behavioral terms in the regression. Models organized according to number of components in model from top (2 components) to bottom (10 components).

(D–F) Mixture of linear regression models with main and interaction behavioral terms. (D) Log-likelihood of data for models with different component numbers. Dashed line indicates separation between models with up until 10 components, presented in (D)–(F), and models with 11,12 components that were fit to ensure saturation of log-likelihood function and detailed model not shown. (E) Association probability for each component in each model. x axis indicates number of components in model, y axis indicates component index within model. (F) Normalized weight of each of the four behavioral terms in the regression. Models organized according to number of components in model from top (2 components) to bottom (10 components). 10 rows shown per model, regardless of number of components in model, solid blue rows indicate lower than 10 component number. Note that in each model there is at least one component with a strong weight for the other-current variable.

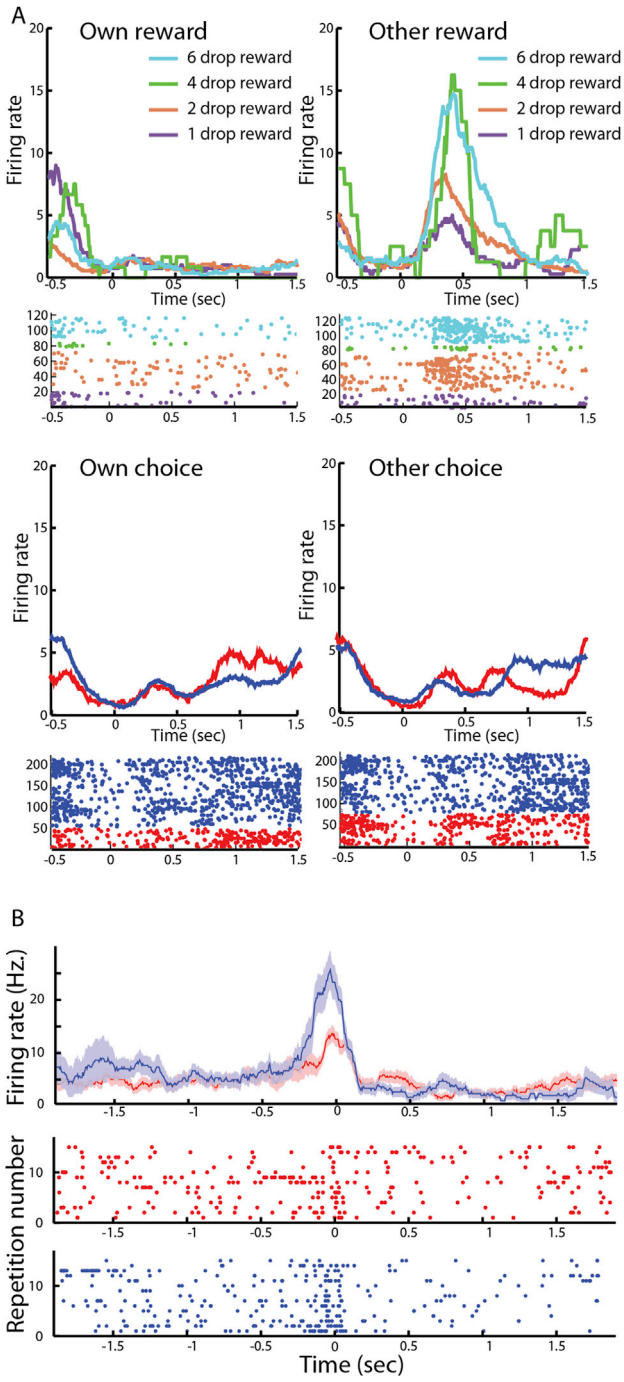


Figure S7. Examples of Single Neurons, Related to Figure 4

(A) Example of a neuron encoding other's reward and not the monkey's own reward or self/other decisions. Top left: peristimulus histogram as mean firing rate for contingencies resulting in different own reward values, and corresponding raster plots. Top right: peristimulus histogram as mean firing rate for contingencies resulting in different other-monkey reward values, and corresponding raster plots. Bottom left: peristimulus histogram as mean firing rate for the monkey's own choice: cooperation (red) and defection (blue), and corresponding raster plots. Bottom right: peristimulus histogram as mean firing rate for other monkey's choice: cooperation (red) and defection (blue), and corresponding raster plots.

(B) Example of a neuron encoding current decision given both monkey's past decisions. Peristimulus histogram as mean firing activity ± SEM and raster plot. Here, the monkey chose defection in both cases after choosing cooperation on the preceding trial. Blue corresponds to trials in which the other monkey defected on the preceding trial and red corresponds to trials in which the other monkey cooperated on preceding trial (see Supplemental Information).