Human and macaque pairs coordinate differently in a dyadic decision game with face-to-face action visibility

Sebastian Möller[1,2,3], Anton M. Unakafov[1,2,3,4,5], Julia Fischer[3,6,7], Alexander Gail[1,2,3,8,a], Stefan Treue[1,2,3,8,a], Igor Kagan[1,3,a,*]

[1]Cognitive Neurosciences Laboratory, German Primate Center – Leibniz Institute for Primate Research, Kellnerweg 4, 37077, Goettingen, Germany

[2]Georg-Elias-Müller-Institute of Psychology, University of Goettingen, Gosslerstrasse 14, 37073, Germany

[3]Leibniz ScienceCampus Primate Cognition, Kellnerweg 4, 37077, Goettingen, Germany

[4]Max Planck Institute for Dynamics and Self-Organization, Am Fassberg 17, 37077, Goettingen, Germany

[5]Campus Institute for Dynamics of Biological Networks, Hermann-Rein-Straße 3 37075, Goettingen, Germany

[6]Cognitive Ethology Laboratory, German Primate Center – Leibniz Institute for Primate Research, Kellnerweg 4, 37077, Goettingen, Germany

[7]Department of Primate Cognition, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, University of Goettingen, Germany

[8]Bernstein Center for Computational Neuroscience, Am Fassberg 17, 37077, Goettingen, Germany

[a] Shared senior authorship

[*]Corresponding author: ikagan@dpz.eu

Short title: Decision coordination in humans and monkeys

Abstract

Interactions of group-living primates with conspecifics range from cooperation to competition. Game theory allows testing the strategies that underlie such interactions, but in classical theory, agents act simultaneously or sequentially. Many real-world decisions, however, are made while directly observing partner's actions. To investigate social decision-making under conditions of face-to-face action visibility, we developed a setup where two agents observe each other and reach to targets on a shared transparent display, enabling naturalistic interactions we call "transparent games". Here we compared human and macaque pairs in the transparent version of the coordination game "Bach or Stravinsky", which rewards coordination but entails the conflict about which of the two individually-preferred coordinated options to choose. Most human pairs developed coordinated behavior, and 53% adopted dynamic coordination via turn-taking to equalize the payoffs. All macaque pairs also converged on coordination, but in a simpler, static way: persistently selecting one of the two coordinated options or one of the two display sides. Two animals that underwent training with a turn-taking human confederate learned to coordinate dynamically. When tested as a pair, they mostly converged on the faster monkey's preferred option, and a dynamic coordination emerged as animals spontaneously took turns in leading to their respective preferred option and following to the other's. The observed choices were captured by modeling a probability to see the other's action before own movement. Importantly, such competitive turn-taking was unlike the benevolent turn-taking in humans, who equally often initiated switches to and from their preferred option. Our findings demonstrate that dynamic coordination is not restricted to humans – although it serves a selfish motivation in macaques – and emphasize the importance of action visibility in the emergence and maintenance of coordination.

Keywords: cooperation, competition, coordination, game theory, inter-species comparison, reward, nonhuman primates, rhesus monkey

Introduction

The majority of primate species lives in complex social groups, in which interactions range from intense competition to cooperation [1–5]. To adjust their behavior optimally, individuals need to assess not only their own actions and goals, but also other group members' current actions, while taking into account the history of interactions with and between these individuals [4,6,7]. Coordination is essential for maintaining cohesion between group members, avoiding conflicts,

and achieving individual and joint action goals. Here we focus on understanding how such coordination can be achieved and maintained in a dyadic setting involving two individual agents, in "transparent" conditions in which both agents can observe each other's evolving actions and social cues [8,9].

To shed light on the evolution of coordination, researchers have turned to nonhuman primates to investigate how their behavior compared to that of humans [10–12]. Macaques play a particularly important role in this context, as the most common nonhuman primate model for studying the neuronal basis of higher socio-cognitive functions [4,13,14]. The realization that primate brain functions are best understood under social conditions in which they evolved sparked a surge of interest in behavioral repertoire and neural correlates of economic and social factors underlying social cognition in humans and nonhuman primates [15,16].

Game theory, developed to study strategic interactions in rational decision-makers [17], offers a powerful framework to investigate dyadic social interactions [18]. Game-theoretical approaches derived from mathematical and experimental economics emerged as an important tool in behavioral ecology [19] and more recently in the burgeoning fields of neuroeconomics and the neuroscience of social decision-making [20–22]. One fundamental class of games are the so called 2×2 games [23–25] in which each agent chooses one of two actions and the outcome depends on the combination of their choices, as defined by the payoff matrices describing the gain or loss the two agents earn for the four possible combinations. Such choices can be presented as one-shot games or as iterated games. In the latter, agents interact repeatedly, as is often the case in real life. Repeated interactions encourage tracking of interaction history, as well as the formulation of predictions regarding the other agent's decisions.

A particularly interesting class of 2×2 games for understanding many realistic scenarios are non-zero-sum games, in which the agents need to coordinate their choices to maximize individual and/or joint reward [26]. As opposed to the strictly competitive zero-sum games, non-zero-sum games have both cooperative and competitive elements, with aligned and opposing interests. A number of such games have been used in studies of coordination behavior in human and nonhuman primates. In the Stag Hunt (or Assurance) game, agents choose between maximizing the individual as well as joint reward and minimizing the individual risk [27]. Humans *(Homo sapiens)*, chimpanzees *(Pan troglodytes)*, and capuchin monkeys *(Cebus apella)* converge at above chance level on mutually beneficial high-reward / high-risk choices in the iterated Stag Hunt game; humans did so at a higher rate than apes, and apes at a higher

rate than capuchins [28–30]. The performance of rhesus monkeys *(Macaca mulatta)* was more similar to humans than capuchins, especially when no information about the current choice of the partner was available [31]. Interestingly, in a computerized Stag Hunt game, humans adapted more closely than macaques to a simulated partner's stag choice probability, showing that different species might employ distinct but overall fairly successful strategies to maximize reward in a dyadic context [32].

Similar to Stag Hunt, in Prisoner's Dilemma (PD), the *joint* reward of two agents is maximized if they both choose to cooperate. Yet, an individual agent obtains the highest reward when defecting while the partner cooperates, and in this case the cooperating partner receives the lowest reward. Thus, there is a dilemma, or conflict between the self-interested temptation to defect and a riskier cooperation. In a single round of PD rational agents should both defect. For mutual cooperation to emerge in PD, agents need to "trust" the partner, i.e. to expect the partner to cooperate for the mutual benefit. Humans often show a bias towards cooperative behavior [33–35]. Remarkably, it was demonstrated that in an iterated PD game rhesus macaques, while mostly selecting defection, chose mutual cooperation significantly more and mutual defection significantly less often than expected by chance, especially after preceding cooperation [36]. This suggests that macaques can at least partially overcome the selfish motives and reciprocate cooperation.

Another type of conflict is implemented in the Conflict game (also known as Hawk-Dove, or game of Chicken) [37]. These anti-coordination games model a competition over a shared resource that can be monopolized by one agent, while an actual clash is highly detrimental to both. Here, compared to PD, if both agents exhibit cooperative behavior (e.g. yield), their joint reward is less or equal to the joint reward of the two anti-coordinated choices (e.g. fight-yield and yield-fight) that maximize the individual reward of one of the agents. Hence, to achieve optimal and balanced distribution of rewards, agents need to alternate between the two anti-coordinated options. Humans, capuchins and rhesus macaques often converged on anti-coordination in this game, but importantly, only humans alternated between the two individually-optimal choices [37]. Such alternating behavior resulted in the maximal and roughly equal payoff for both agents and is called "cooperative turn-taking" [38]. It indicates that in the Conflict game, humans but not nonhuman primates strive for fairness.

In the games considered so far, reward-optimizing rational agents should converge on a behavior in which they either coordinate on one option (Stag Hunt and Prisoner's Dilemma) or

anti-coordinate on opposite options (Conflict game). A third type of games that emphasizes the coordination on *either of two options* is known as Bach or Stravinsky, or Battle of the Sexes (BoS) [39]. Similarly to Stag Hunt and Prisoner's Dilemma, but unlike the Conflict game, BoS models a cooperative interaction in which agents should optimally converge on the same option. Each agent has an individually preferred option, but coordinating on either one of these two options adds the same bonus to both agents. This renders any coordinated choice better than no coordination for both agents, but one coordinated choice best for the first agent and another for the second agent. The coordinated choices represent two pure-strategy Nash equilibria. For both agents, the rational choice is to coordinate, but unlike the conceptually simpler optimal convergence on one option in Stag Hunt and PD, BoS includes an inherent conflict about who profits the most. Consequently, when playing iteratively, turn-taking strategies are needed to avoid unequal distribution of rewards.

This combination of cooperation and conflict offers an interesting opportunity for studying social interactions that has not been sufficiently explored in nonhuman primates. Games of the BoS type have revealed that humans indeed often take turns – i.e. they switch from coordinating on one option (preferred by one player) to coordinating on the other option (preferred by another), again indicating human propensity for fairness [40–43]. Interestingly, stable cooperative turn-taking frequently took place in 5-year-old children, but neither in 3-year-old children nor in chimpanzees [42]. This begs the question whether turn-taking requires special social abilities that are unique to humans (and potentially only emerge later in life). We therefore compared the behavior of (adult) humans and rhesus macaques in the BoS economic game.

Traditionally, economic games are either played *simultaneously* (neither agent knows the choice of the other before making its own decision) or *sequentially* in a certain order; yet real dyadic interactions often play out in a real-time with the partner's actions in direct sight [44,45]. Thus, the timing of one's own and other's actions becomes part of the strategy space [46]. Such a "transparent" continuous time setting can change choice strategies of the agents as compared to the classic simultaneous and sequential settings [38]. Especially for nonhuman species, coordinating based on mutual choice history might be more demanding than coordinating based on the immediately observable behavior of others. For example, visual feedback about partner's choices improves the convergence on a high reward / high risk equilibrium in the iterated Stag Hunt game in humans, capuchins and rhesus macaques [31], and such coordination in chimpanzees is facilitated if one of the agents consistently acts faster than the partner, suggesting a leader-follower strategy [29]. Similarly, there were substantial differences in

capuchins' and rhesus' behavior in the Conflict game when they had access to the current choice of a partner [37]. In humans, a real-time anti-coordination game revealed that action visibility (and the possibility to change an already initiated action) increased efficiency and fairness [47]. Besides those few pioneering studies, the experimental and theoretical effects of action visibility on (anti)coordination games have not been systematically addressed.

To account for possible changes in strategies during action visibility compared to simultaneous ("opaque") discrete choices, we recently developed the concept of "transparent games" that extends classical evolutionary game-theoretic analysis to real-time interactions in which the visibility of partners' actions depends on their relative reaction times [8,9]. Here, we use the theoretical insights from that work and analyze short- and long-term dynamics of choices, mutual information and reaction times in humans and rhesus monkeys playing a transparent version of the iterated movement-based BoS game. Since many direct interactions between primates take place face-to-face [48–52], we designed a novel dyadic interaction platform that allows two human or monkey agents to observe each other and to act on the same visual objects on a vertical touch-sensitive transparent display between them. The advantage of this configuration is that agents can monitor and react to each other's actions in real-time, emulating naturalistic interactions while still maintaining well-controlled laboratory conditions. Extending beyond static choice proportions, our approach captures dynamics of such interactions, affording a mechanistic model of trial-to-trial choice behavior driven by within-trial temporal patterns.

To facilitate inter-species comparability, we paired human or macaque agents without explicit task instructions. We expected that the transparency and instantaneous coaction [45] would facilitate efficient coordination, as well as reciprocal cooperative turn-taking in humans. We predicted that macaques would also utilize the information about partner's actions, and anticipated one of the two outcomes. Macaques might exhibit a form of turn-taking, or converge on one of the two pure equilibria, i.e. coordinate to increase individual (and joint) rewards, but without turn-taking to balance rewards between players.

Results

In our experiments, two human or macaque agents were sitting face-to-face and shared a transparent vertical workspace between them (**Figure 1A,B**). Agents could simultaneously see the stimuli and each other's actions. In the variant of the BoS paradigm we used, each agent

chooses one of two simultaneously presented color targets (red and blue). The left/right position of each color was randomized across trials (50% left, 50% right). The resulting rewards follow the payoff matrix shown in **Figure 1C** and depend on the choices of both agents. Prior to the first dyadic session, each agent in a pair was individually trained to associate one of the two color targets to a larger reward, so that they were biased to value different colors. In the dyadic trials, selecting the same target resulted in an additional reward on top of the individually trained values, hence choosing the same target was always better than choosing different targets. But in such coordinated trials, the agent whose individually preferred ("own") color was selected received a larger reward. This paradigm probes the ability to realize that coordinated target selection results in higher rewards, and the ability to perceive and deal with the inherent unfairness introduced in each trial by the unequal rewards for selecting the same target.
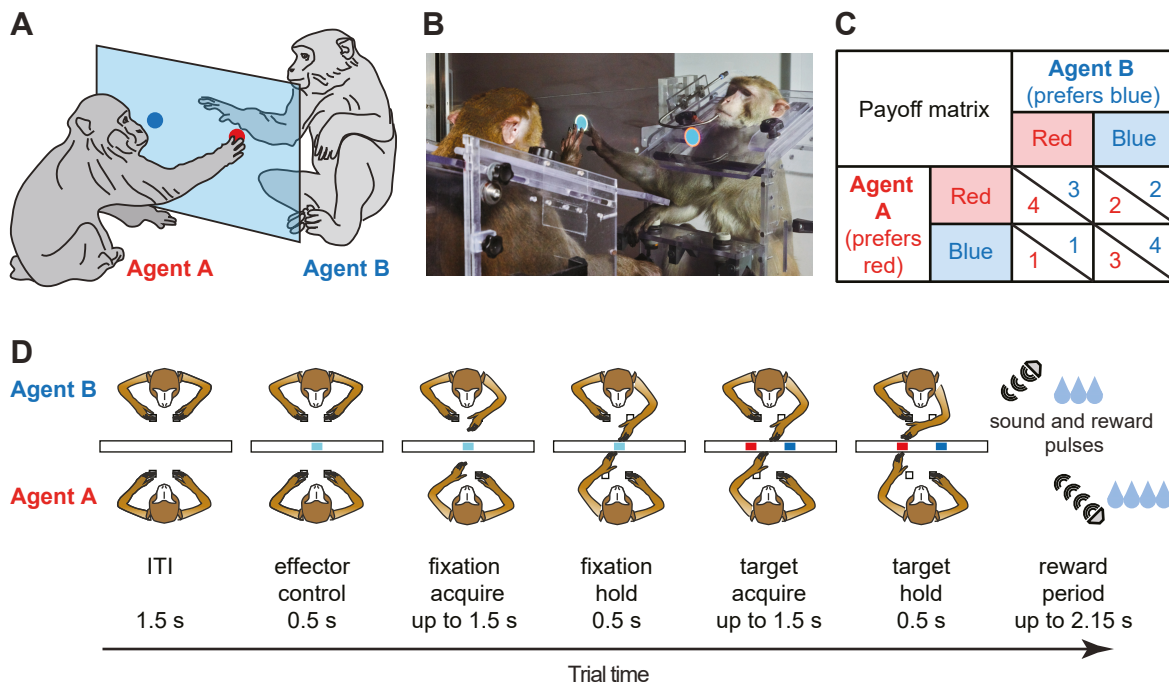

Figure 1. Experimental setup. (A) A sketch of the dyadic interaction platform showing two macaques reaching to targets in a shared workspace. **(B)** A photo of two macaques monkeys working in the dyadic interaction platform. Here both agents reach to the same target with the left hand. **(C)** The payoff matrix that describes the reward outcome for each player for all four possible combinations of individual choices (numbers show the monetary/liquid reward units for agent A in red and agent B in blue). The coordinated selection of the same target resulted in a "bonus" of additional 2 reward units. **(D)** Schematic top-view on two macaque agents in different stages of the task, and the timing of one trial.

Human pairs mainly converge on fair coordination

We recruited naïve human subjects to assess performance in the BoS paradigm. Each subject was trained individually for 50-100 solo trials to operate the touch panel and to associate a higher auditory pulse count/monetary reward with one of the two color targets (**Table S1**). We

then paired these pre-trained agents and let them explore the full payoff matrix. Subjects were instructed not to talk during the experiments (see Methods and Supporting Information for the full set of instructions given to the subjects). We mainly analyzed the last 200 dyadic trials of each session, to allow the subjects to explore all choice combinations and to converge on a strategy (Methods).

Human strategies ranged from strict adherence to the pre-trained higher value colors to joint selection of the same target with alternating between the two colors trial-by-trial or in blocks. **Figure 2A,B** illustrates an exemplary human pair that developed such alternating turn-taking behavior, using *share of own choices* (SOC) - the likelihood for each agent to select the own preferred color, and the *share of left choices* (SLC) - the likelihood for each agent to select the target on the left side. These measures provide direct insight into each agent's value and side biases. Here the SOC curves show that they mostly jointly selected A's or B's color. After an initial period of long blocks of trials, this pair switched to shorter blocks. The SLC curves show that the agents did not display a side bias.

The SOC of all 19 pairs (**Figure 2C**) shows 53% (10 of 19) with balanced values close to 0.5 (indicating alternating between colors), and the remaining 47% (9 of 19) with at least one agent close to one or close to zero (indicating fixed color selection). **Figure 2D** reveals that most agents had no side bias, since the SLC was around 0.5. **Figure 2E** shows the average reward (AR) for each agent and the average joint reward, which in most pairs was above chance level (2.5 units), and in about half the pairs close to the optimum for the dyad (3.5).

From the sequence of choices and achieved rewards we derived three measures of coordination (Methods). The significant *dynamic coordination reward* (DCR) describes the average amount of reward a pair earned above (or below) the reward expected for independent choices with the observed choice frequencies per agent; zero DCR indicates coordination by chance alone, values significantly different from zero indicate above chance levels of coordination. The DCR values (**Figure 2H**) and the fairly balanced average rewards (**Figure 2E**) show that 10 of 19 human pairs converged on a form of "turn-taking", or reciprocal coordination. These are the same 10 pairs that showed roughly balanced SOC and SLC values around 0.5 (**Figure 2C,D**). To better understand what drives high dynamic coordination seen in **Figure 2H**, we calculated the *mutual information* (MI) between the sequences of target color choice (MI target, **Figure 2F**) and the side choice (MI side, **Figure 2G**) of both agents; these measures indicate how well one agent's choice of color or side can be predicted from the other's choice

(Methods). Comparing the DCR and the two MI reveals that high DCR values coincide with non-zero values of both MI side and MI target, indicating that dynamic coordination corresponds to both correlated side and color choices between the agents.
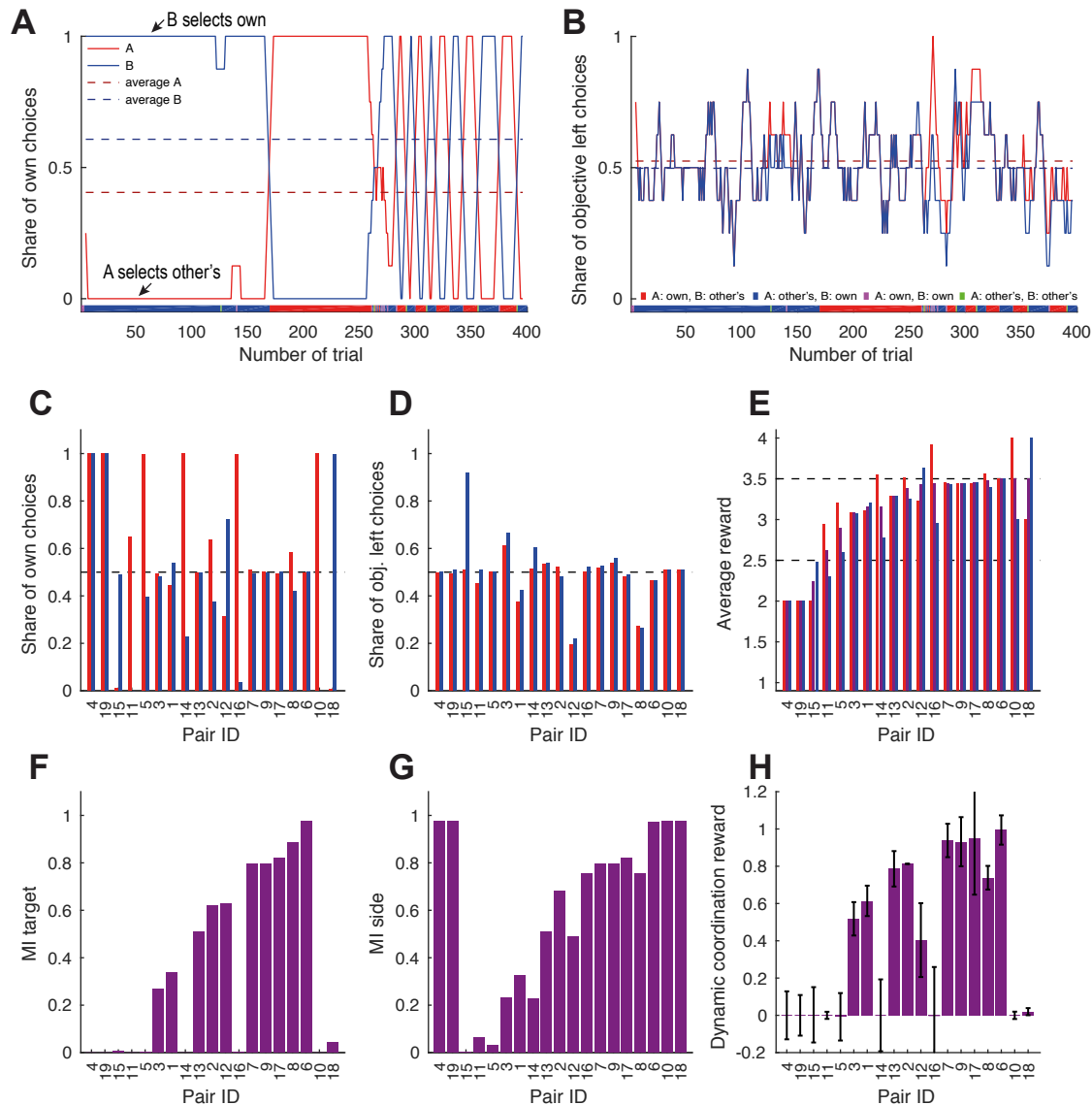

Figure 2. Performance of naïve human pairs. (A) Example session showing for both agents (A: red; B: blue) the share of own choices (SOC) as running average over 8 trials. In coordinated choices one agent always selects the own target (SOC = 1), while the other selects the non-preferred target (SOC = 0). The dashed lines show the session average SOC for each agent. The colorbar below shows for each trial the selected combination (red: A and B selected A's preferred target; blue: A and B selected B's preferred target; magenta: both agents selected their individually preferred target; green: both agent selected a non-preferred target). Here both human agents coordinated by selecting the same target in blocks, switching to short blocks in the latter part of the session. **(B)** The share of objective left choices (SLC, from A's viewpoint) as running average over 8 trials. The dashed lines show the session average SLC for each agent. Here side selection fluctuated around random level 0.5 for both players. **(C-H)** Aggregate measures of choice behavior and coordination for 19 naïve human pairs sorted by the average joint reward (the example in (A) and (B) is from pair 2). Individual measures for A are shown in red, for B in blue, measures for the pair are in purple. **(C)** The session average SOC for every pair. The dashed line at 0.5 indicates the level of equal selection of both color targets. **(D)** The session average SLC for each pair. The dashed line at 0.5 indicates equal selection of both color targets. Only agent B of pair 15 and pairs 12 and 8 showed a strong side bias.

(E) The average reward for each agent and average joint reward. The dotted line at 2.5 shows the reward value for random choices, the line at 3.5 shows the maximum value for the average joint reward. The difference between red (A) and blue (B) bars indicates the unfairness of the reward distribution between the agents (zero indicates perfect fairness). **(F)** The mutual information (MI) between the target color choices of both agents. **(G)** The MI between the side choices of both agents. **(H)** The dynamic coordination reward (DCR). The error bars show 95% confidence intervals, 53% of pairs (3, 1, 13, 2, 12, 7, 9, 17, 8, and 6) showed significant DCR.

The 9 human pairs that did not show significant dynamic coordination used strategies other than turn-taking. These alternative strategies included color-based strategies such as: (i) both agents largely converging on a fixed color (5, 14, 16, 10, 18; indicating that at least one of the two agents understood the value of coordination), (ii) exclusively selecting the respective own preferred color (4, 19); and (iii) mixed color-side strategies where one agent consistently selected the non-preferred color while the other seemed to switch randomly between the two colors (11) or selected the right side (15). To summarize these results, more than half of the human pairs exhibited dynamic coordination and aimed for fairness.

Macaque pairs converge to simpler coordination strategies

Macaques were individually trained for multiple weeks to internalize the individually preferred reward values of the two color targets. All six macaques developed a strong preference for the large reward (range 78-100%, mean 95% large reward color selection the last solo session), before being paired with a conspecific (**Table S2**). Generally, macaques changed their choice behavior more slowly than human participants who often exhibited clear changes within a single session. We therefore collected the data from macaque pairs for multiple sessions to see how their behavior evolved over time and what strategy they converged on, instead of only sampling each macaque pair in a single session (note that we use the term "strategy" to conform to the standard game-theoretical notion, indicating a decision but making no assumptions about the underlying understanding which led to that decision, cf. [28]). **Figure 3A,B** shows the color and side choice behavior for the first dyadic session of a typical example pair of monkeys. The SOC curves indicate that both agents had a strong preference for their own preferred colors; the SLC curves indicate no strong side selection bias. Later, in the sixth dyadic session (**Figure 3C,D**), the behavior had changed, with both agents displaying less own color selection and a stronger convergence on the left side of the screen (average SOC values closer to 1 as compared to 0.5 in the early session). This shift in strategy was gradual (**Figure S1A-F)** and resulted in an overall better outcome for both agents with nearly equalized reward (**Figure S1C**). The DCR did not reach significance for any session (**Figure S1F**) indicating that this pair did not employ a dynamic coordination strategy.
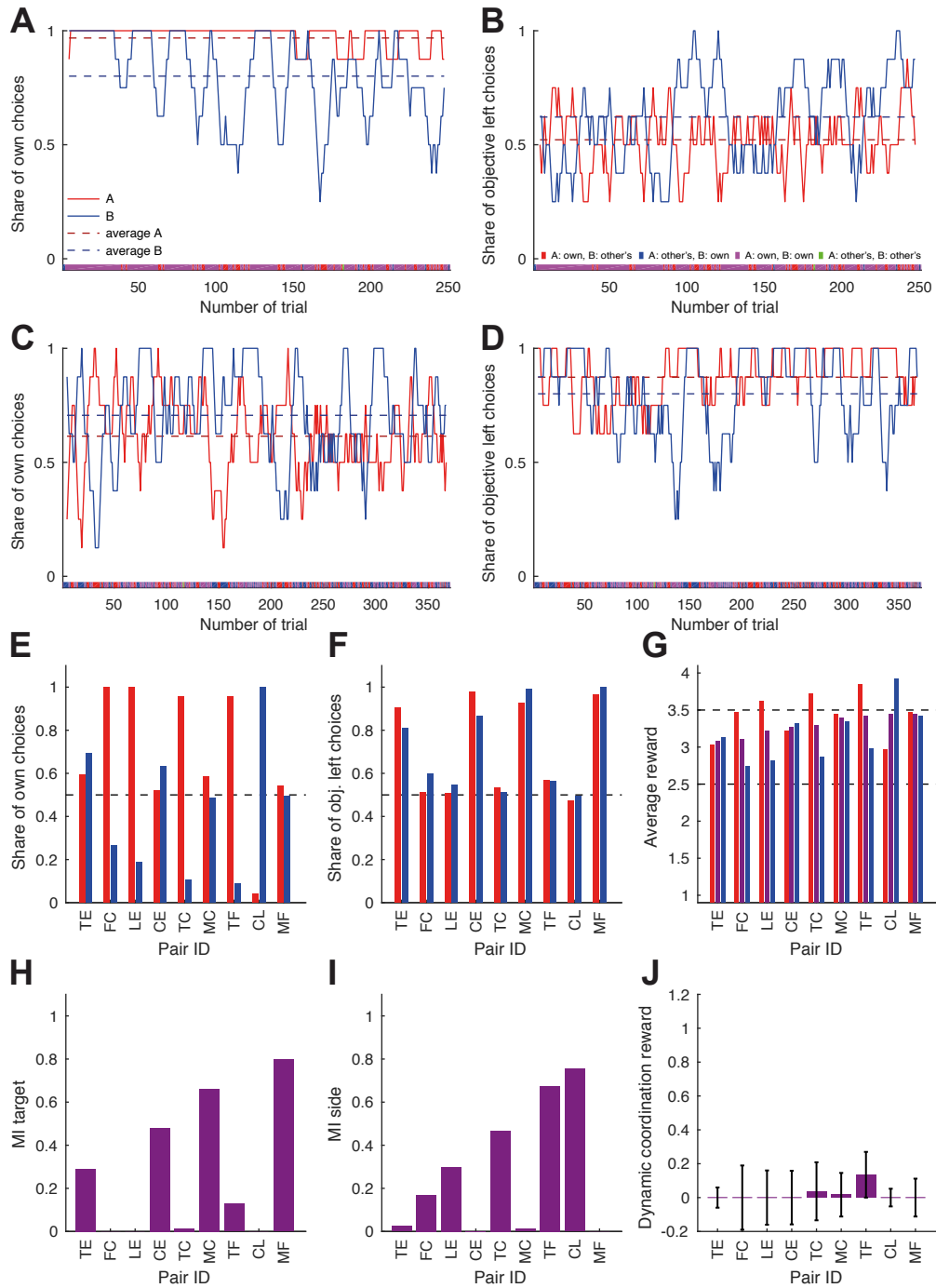
Figure 3. Performance of naïve monkey pairs. (A, C) The running average of the share of own choices from an early (A) and a late session (C) of pair TE. **(B, D)** The running average of the share of objective left choices from an early (B) and a late session (D) for the same pair. Conventions are the same as in Figure 2A,B. The pair shifted from each selecting their individually preferred target to mostly jointly selecting the objective left side. Panels **(E-J)** show aggregate measures of choice behavior and coordination for a late session for each pair sorted by the average joint reward; conventions as in Figure 2C-H). Overall macaque pairs tended to either converge on jointly selecting either the same side or on jointly selecting one of the two color targets.

A similar picture emerges in the set of 9 pairs tested (with 6 unique macaques) (**Figures 3E-J** and **S2**). Comparing the average rewards in early and late sessions (**Figure 3G, S2C**) shows that most pairs over time reached better coordination, as the average reward across all pairs increased (**Figure S3**; early median 2.75 versus late sessions median 3.29, N: 9; p < 0.0039, two-sided exact Wilcoxon signed rank test). In 8 of 9 pairs, the proportion of coordinated trials increased significantly (p < 0.05, Fisher's exact test of the proportion of coordinated versus uncoordinated trials in the early and late session).

Comparing early and late SOC and SLC (**Figure 3E,F**, and **S2A,B**) shows that this increase in coordination was mostly achieved either by converging on the same color target (pairs FC, LE, TC, TF and CL) or by converging on the same side (TE, CE, MC, and MF), also evident in non-zero values of target or side MI (**Figure 3H,I**). The low and with one exception non-significant DCR values in **Figure 3J** indicate that the resulting coordination was not achieved by dynamic turn-taking.

Converging on either the same color or the same side are both simpler strategies than dynamic coordination, by virtue of allowing each agent to make decisions without paying attention to the trial-specific actions of the partner. To summarize, macaque pairs learned to coordinate using simple strategies that did not require trial-by-trial integration of partner's action.

Comparison of behavior between species

To compare the strategies of naïve human and macaque pairs, we plotted the mutual information values for side and color choices of each pair, for the last 200 trials (**Figure 4**). The distance of each pair's location to the origin denotes the strength of (anti)coordination. Locations close to the x-axis denote static side strategies, locations close to the y-axis denoting static target color strategies. Ten of 19 human pairs were close to the main diagonal, corresponding to similar values of side and target MI. These locations denote trial-by-trial or block-wise turn-taking – a signature of dynamic coordination. In contrast to humans, nine macaque pairs developed simple static side-based (4 pairs) or single color-based (5 pairs) coordination, with different levels of coordination strength, but did not employ dynamic coordination.
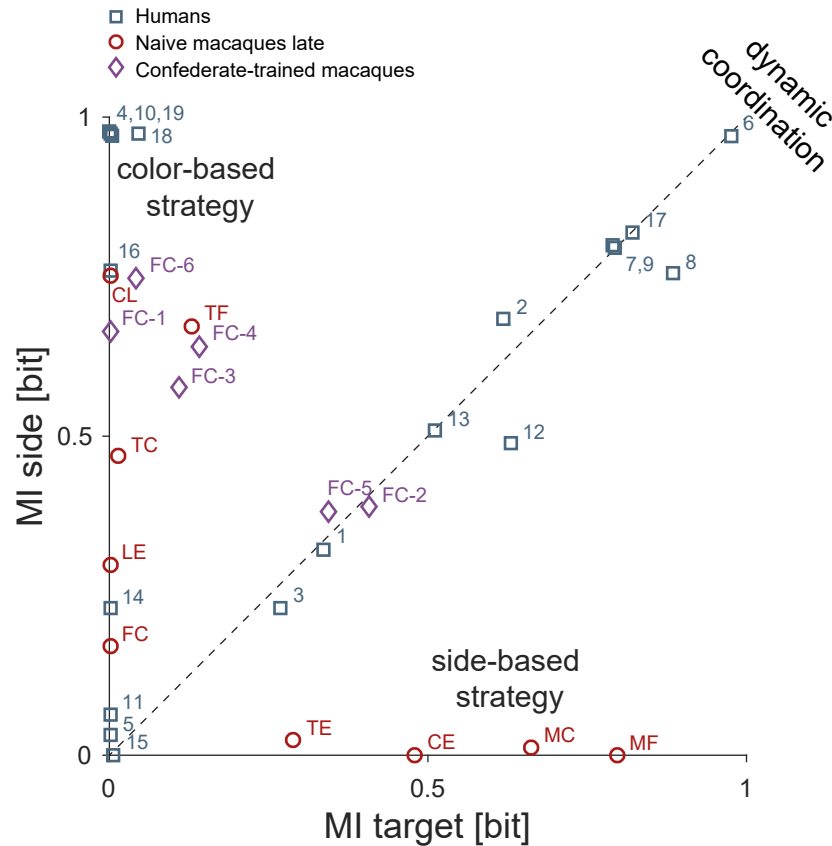
Figure 4. Comparison of coordination strategies in humans and macaques. Each symbol shows the mutual information between side choices (MIS) and target color choices (MIT) calculated over the last 200 trials of a session, for human and macaque pairs. Since mutual information for a non-varying sequence of choices is zero, sessions in which a pair statically converged on a fixed color have MIT around zero and a non-zero MIS, and sessions in which a pair statically converged on a fixed side have a non-zero MIT and around zero MIS. Pairs/sessions near the y-axis hence indicate static selection of one of the two target colors (i.e. either coordinating on A's or B's preferred option, or only selecting the own color), while pairs/sessions near the x-axis indicate static coordination on a side. Pairs/sessions with balanced side and color choices fall onto the main diagonal, indicating dynamic coordination or turn-taking. The distance between each symbol and the origin represent the (anti)coordination strength. Ten human pairs showed dynamic coordination, using trial-by-trial or block-wise turn-taking. Eight human pairs employed color-based strategies: 5 pairs only selected one of the two possible colors, i.e. one of the two possible coordinated options, and 2 other pairs did not select any coordination options (4, 19). The eighth pair (18) sampled both coordination options albeit in a much skewed 4 to 394 ratio. Interestingly, seven of these eight pairs (but only one of the 10 turn-taking pairs) did not explore all 4 possible choice combinations (i.e. payoff matrix outcomes) and hence could not have developed a full understanding of the rules of the game. In one human pair both MIS and MIT did not significantly differ from zero (p > 0.01; pair 15, in this pair one agent selected one side and the other - one color). In late sessions, naïve macaque pairs either converged on selecting the same color target (pairs FC, LE, TC, TF, CL) or converged on the same side (TE, CE, MC, MF). In addition to naïve humans and macaque pairs, diamonds show 6 sessions of the confederate-trained macaque pair (F-C), see subsequent sections.

Macaques paired with a human confederate learn to follow

The absence of trial-by-trial coordination in macaque pairs suggests that they did not monitor their partners in order to take their choices into account on a trial-by-trial basis. This might be a consequence of the animals' training history (starting with extended individual solo training to associate one color with higher reward) or a general inability to consider the partner's choice in making immediate choices, at least when a simpler strategy also helps to maximize the reward.

13

To test the latter, we trained two monkeys in the dyadic condition with a human confederate who followed a strict pattern of alternating between the two colors in blocks of mostly 20 trials ("confederate training"). If the rhesus monkeys were insensitive to partner's choices in our dyadic task, the pattern of confederate's choices should not affect monkey's behavior.

Figures 5 and **Figure S4** show examples from early and late confederate training sessions for two animals. In both cases the monkeys started with a strong bias for selecting their own preferred color but over time changed to a reliable coordination (cf. **Figure 5A** with **5C** and **S4A** with **S4C**). We call it the "following" behavior since in both animals reaction times initially were faster that the confederate's, and did not vary strongly depending on the confederate's actions, but in later sessions the animals reacted considerably slower in blocks where the confederate was selecting the animals' non-preferred color (compare **Figure 5B** with **5D**, and **S4B** with **S4D**), waiting for the confederate to commit to an action.

The aggregate measures of choice behavior (**Figure S5A,B, S5G,H**) and coordination development (**Figure S5C-F, S5I-L**) over multiple confederate-paired sessions showed similarities in monkeys' learning (cf. **Figure S5A-F** and **S5G-L**). Both animals reached close to 0.8 DCR units, which is similar to the values reached by humans employing turn-taking (cf. **Figure 2H**). These results indicate that macaques can take information from their partner into consideration when making decisions in the BoS paradigm. Reaction time dynamics hinted that macaques started to actively monitor confederate's actions and only selected their non-preferred color if the partner started to reach there first.

Due to the repeating block strategy employed by the human confederate it is not immediately clear however whether the monkeys made use of the action visibility and based their decision on the action observed within a trial, or rather on the reward history (e.g. switched after several uncoordinated trials at the beginning of a new block). To test how important the immediate access to the partner's choice is for driving the "following" behavior, we performed a control experiment in which we placed an opaque barrier on the confederate's side of the screen for roughly the middle third of a session. The barrier blocked the monkey's view of the confederate hand, while keeping the face visible. **Figures 5E,F** and **S4E,F** show the strong effect this manipulation had on the monkeys' behavior. In both monkeys the "following" behavior ceased, shifting to selecting the own preferred colors (**Figures 5E, S4E**). The reaction time difference histograms (**Figures 5F, S4F**) confirmed this observation: the significant difference between the selecting the preferred (blue) and the non-preferred (red) color in the upper half of the plot for

the transparent condition disappeared in the lower half of the plot showing the results for the opaque condition. In other words, monkeys failed to coordinate without seeing the confederate's immediate actions. These results imply that the immediate visual information about partner's actions and not merely monitoring own reward history drove and maintained the macaques' "following" behavior in transparent settings.
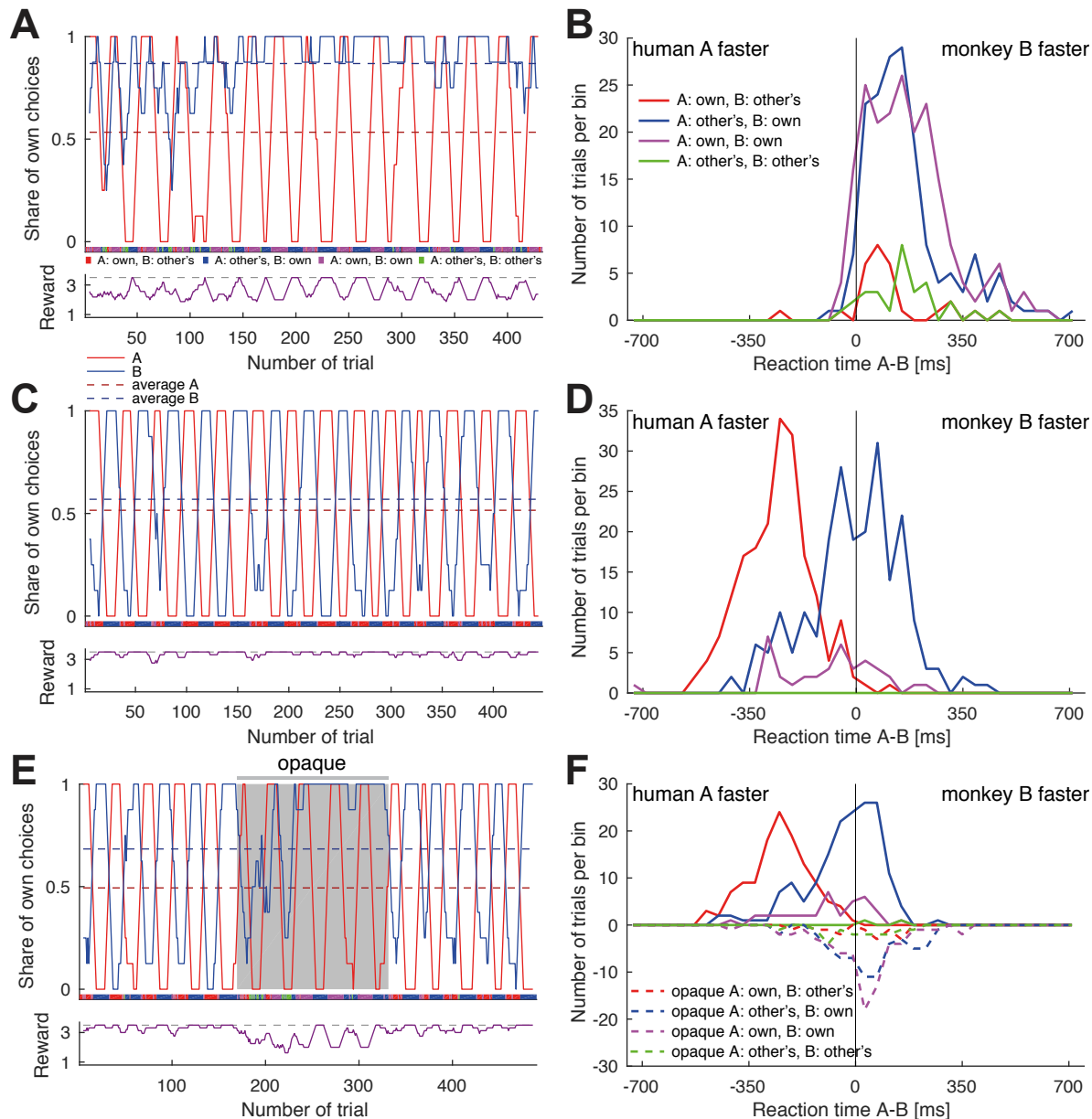

Figure 5. Macaques can be trained to take turns if paired with a human confederate. Pairing a macaque with a human partner that alternated between the two color targets in blocks leads to turn-taking. **(A)** In an early confederate training session the monkey F (agent B) mostly selected his own target, independent of the human's (agent A) choices. The lower sub-panel shows the running average (8 trials) of the average joint reward over both agents; the cyclic modulation shows that the macaque did not coordinate with the confederate. **(B)** The reaction time difference histograms (bin size 40 ms) show that in all four choices combinations the macaque acted faster than the

confederate. **(C)** A late confederate training session. After several sessions (Figure S5) the macaque had learned to coordinate with the confederate (note anti-correlation of the two SOC curves) and the average joint reward approaching maximum 3.5 value. **(D)** The reaction time difference histograms show that in blocks when the pair selected the macaque's non-preferred target (red curve), the monkey acted slower than the confederate, waiting to see the human's choice before committing to its own choice (t-test: coordination on red (mean: -259 ms, SD: 113, N: 193) vs. blue (mean: -7 ms, SD: 153, N: 217), t(395.5): -19.1, p: $4*10^{-58}$). **(E)** Macaque turn-taking with human confederate depended on direct visibility of the confederate's actions. The gray background denotes the period in which an opaque screen was placed over the region of the transparent display encompassing all touch targets; hence, the agents could not see each other's reach movements during that period but could still see the other's face. Note how the macaque F switched from turn-taking (as seen before and after obstruction) to mainly selecting own preferred target. **(F)** Positive histograms for the transparent condition (solid curves) and negative histograms for the opaque condition (dashed curves). In the transparent condition the behavior was similar to (D) in that the monkey waited for the human to act first in blocks of the monkey's non-preferred target (t-test: coordination on red (mean: -248 ms, SD: 101, N: 123) vs. blue (mean: -39 ms, SD: 124, N: 166), t(284.5): -15.9, p: $3*10^{-41}$). In the opaque condition the behavior looked similar to (B) the monkey mostly selected his preferred target and there was no difference dependent on the human's choices (note the overlap between dashed blue and magenta curves).

Confederate-trained macaque pair shows dynamic coordination

After two macaques had learned to monitor and follow a partner's actions, we tested how these two animals would perform when paired with each other again. At the end of initial naïve sessions of this pair (i.e. before training with a human confederate), monkey F (agent A) mostly selected his own preferred color while monkey C (agent B) either selected the left side or monkey F's preferred color. At the end of the confederate training, both monkeys followed the confederate with a block length of approximately 20 trials. Would they express turn-taking behavior or would they fall back to their initial static color-coordination behavior?

Figure 6 shows the share of own choices and the reaction time differences for the first three sessions after the re-pairing (**Figure 6A-C**), as well as the choices and aggregate measures of choice behavior and coordination over the course of all 6 sessions (**Figure 6D-I**). In the first session, after some initial back and forth, the pair converged on mostly selecting B's color, with B reacting faster (**Figure 6A**). Animal A occasionally selected his own color repeatedly, even in the latter part of the session. In the third session, the pattern is inverted, with the pair converging mostly on A's color, with A reacting faster (**Figure 6C**). This time animal B repeatedly tried to select his preferred color, and A actually followed for three short periods. The most interesting behavior however developed during the second session, in which the repeatedly animals alternated between the two colors in long blocks (**Figure 6B**). Block size was on average ~45 trials, i.e. noticeably longer than the ~20 trial blocks employed during the confederate training.

If action visibility is driving this form of dynamic coordination, than the question of who is leading ("insists" on own preference) and who is following ("accommodates" to the other's preferred target) should depend on the relative reaction times of the two agents. Indeed, in the session in

which the animals expressed turn-taking behavior (**Figure 6B**), it was the faster animal that selected the own preferred color, and the slower animal accommodated, demonstrating dynamic coordination and suggesting a competitive interaction. The mutual information measures as well as the DCR values (**Figure 6G-I**) indicate that in the 2nd session and the 5th session after confederate training, this pair exhibited robust dynamic coordination resulting in turn-taking behavior.
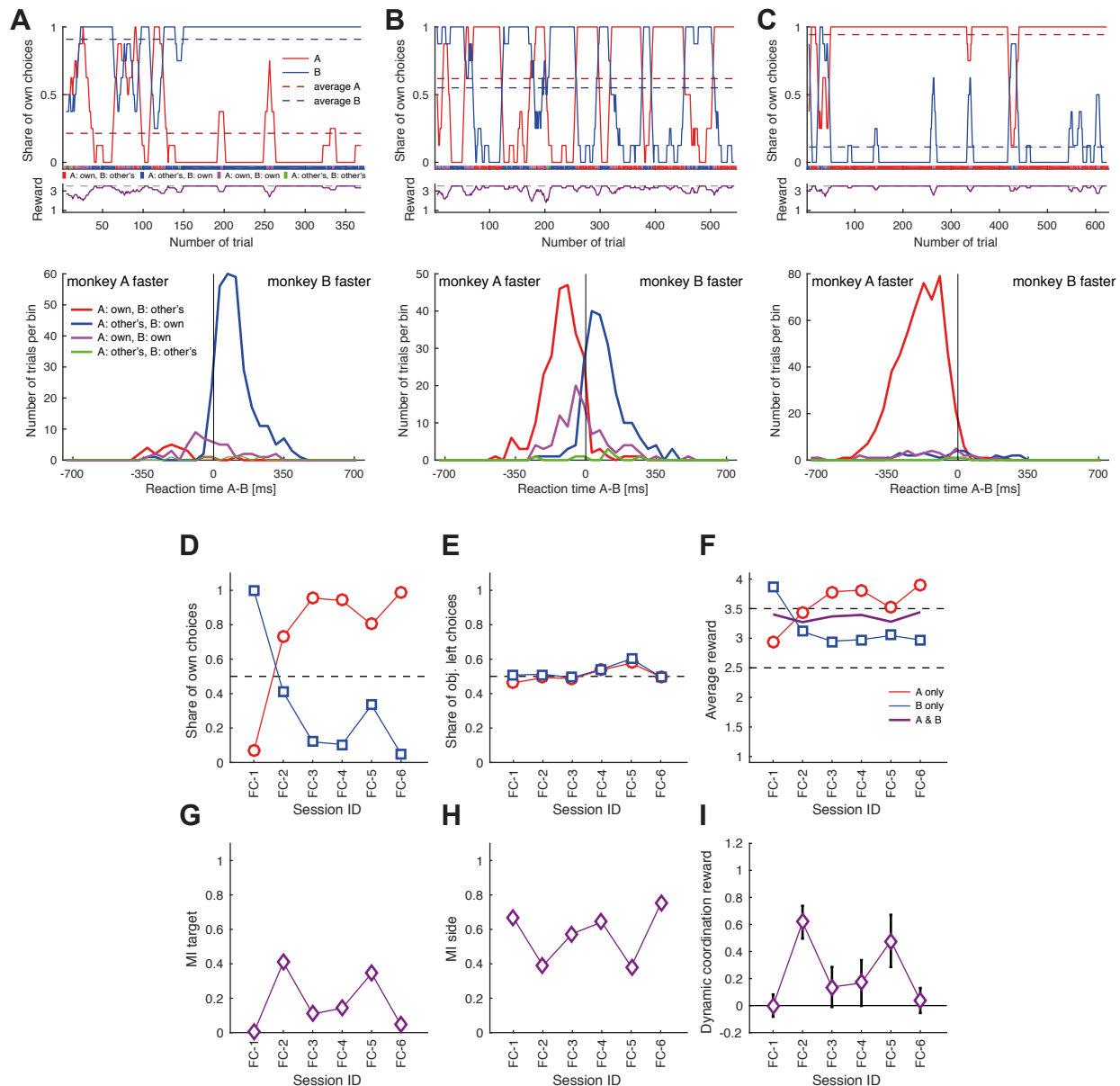
Figure 6. Confederate-trained macaques show competitive turn-taking behavior. (A) In the first paired session after confederate training some intial signs of turn-taking followed by a convergence on B's blue target can be seen in the SOC curce in the top sub-panel; but note how agent A (red) occasionally probes selecting his preferred target. The reaction time difference histograms in the lower panel (bin size 40 ms) show that agent B acted faster when both agents selected his preferred target (t-test: coordination on red (mean: -211 ms, SD: 112, N: 27) vs. blue (mean: 105 ms, SD: 98, N: 286), $t(29.9)$: -14.1, p: $9*10^{-15}$). **(B)** In the next session the pair developed spontaneous turn-taking;

the RT difference histograms show that the pair generally converged on the faster monkey's target (t-test: coordination on red (mean: -117 ms, SD: 96, N: 238) vs. blue (mean: 90 ms, SD: 104, N: 201), t(411.7): -21.4, p: 5*10^{-69}). In the following session **(C)** the pair mostly converged on A's preferred target, with A reaching faster (t-test: coordination on red (mean: -188 ms, SD: 115, N: 553) vs. blue (mean: -31 ms, SD: 192, N: 33), t(33.4): -4.6, p: 5*10^{-5}); note that as in the first session the agent B that mostly selected his non-preferred target occasionally went to his own preferred target (also note how two of these attempts were reciprocated by agent A). **(D-I)** The aggregate measures of choice behavior and coordination over 6 sessions (same conventions as in Figure 2C-H). Note that the MIs and the DCR values for the second and fifth session indicate dynamic coordination.

Comparison of dynamic coordination between species

At the first glance, the observed turn-taking in the confederate-trained macaques resembled the long turn-taking observed in some human pairs (cf. **Figure 4**, FC-2,5). Yet, the analysis of reaction time distributions suggests that macaques employed a competitive version of turn-taking in which the faster agent selects its own color (insists) while the slower agent follows (accommodates). Such behavior would be in line with our theoretical predictions [8], where we show with evolutionary simulations that competitive turn-taking provides the most effective strategy for a BoS type game when players have high probability to observe choices of the partner. To further test the hypothesis that this macaque pair developed competitive turn-taking we performed a correlation analysis between the modeled visibility of the faster agent's action by the slower agent and the observed likelihood of following to the faster agent's color (the faster animal nearly always opted to select his own color). **Figure 7A** shows agent A's probability to select his non-preferred color, the "share of other's choices" (the inverted share of own choices) and the modeled probability for A to see B's choice before A made his choice; **Figure 7B** shows the same for agent B (see Methods for details on modeling). In both monkeys there was a strong correlation between the model and the observation, ranging from 0.92 for smoothed data to 0.64 for unsmoothed data, all significant at p < 0.00005. This analysis further supports the notion that the confederate-trained macaque pair had developed a competitive turn-taking (alternating Leader-Follower) strategy. Strong positive correlations between action visibility and following behavior were present in all 6 post-confederate-trained sessions, indicating bouts of dynamic coordination (**Table S6**).

In contrast, human pairs did not show a strong relationship between partner's action visibility and choice. Even in the human pair with the strongest positive correlations (pair 12, **Figure 7C,D**), the "following" behavior was weak and resulted from a gradual transition from B initially following A to A following B later in the session, rather than from alternating turn-taking as seen in **Figure 6B**. This does not imply that reaction times did not play a role in human choices. Indeed, in 47% (9 of 19) human pairs the relative difference in reaction times between the agents significantly differed in coordinated trials compared to non-coordinated trials (**Table S3**).

18

This implies that similarly to macaques, seeing the faster agent's choice can help the slower agent to coordinate. But unlike the macaque pair, dynamic human coordination was not driven by the faster agent striving to select the own preferred color. Only one human pair out of 10 that exhibited dynamic coordination showed a significant difference in reaction times for coordinated choices favoring A vs. B, with each agent being faster when selecting the own preferred color (pair 12, coordination on A (mean: -88, SD: 129, N: 96) coordination on B (mean: 12, SD: 109, N: 180), t(168.8): -6.5, p: $10*10^{-06}$, t-test). See **Figure S6** and **Tables S4-S7** for details of correlation analyses in different groups.
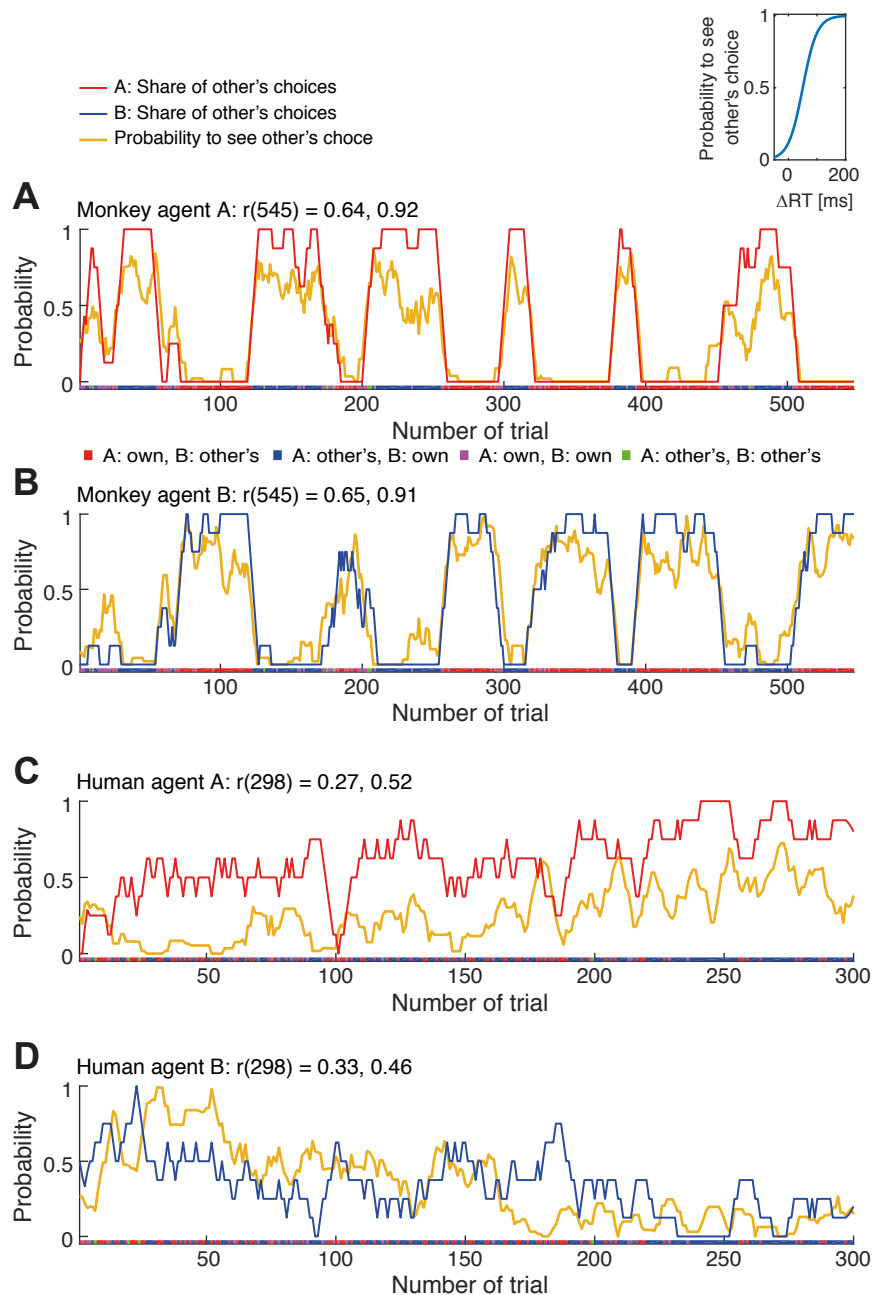
Figure 7. Confederate-trained macaque turn-taking is driven by action visibility. This figure shows the correlation between probability of seeing the other reach first (p_{see}) and following to the other's preferred target (p_{other}). **(A,B)** The 2nd session of confederate-trained macaque pair F-C (cf. Figure 6B). The inset (top right) shows the logistic mapping function from reaction time difference to the probability to see the partner's choice before acting. The "share of other's choices" is the inverse of the share of own choices measure used so far. All curves show running averages of the respective measure over 8 trials. For macaque A the correlation coefficient was $r(545): 0.92$ ($p <0.00005$) and for macaque B $r(545): 0.91$ ($p < 0.00005$) for smoothed data, and $r(545): 0.64$ ($p < 0.00005$) and $r(545): 0.65$ ($p < 0.00005$) for non-smoothed (trial-by-trial) data. **(C,D)** In comparison the human pair 12 with the highest significant positive correlation showed a starkly different pattern: a slow gradual shift from B following A to B leading A over the course of the session. For human A the correlation coefficient was $r(298): 0.52$ ($p: 2*10^{-22}$) and for human B (D) $r(298): 0.46$ ($p: 8*10^{-17}$) for smoothed data, and $r(298): 0.27$ ($p: 2*10^{-6}$) and $r(298): 0.33$ ($p: 5*10^{-9}$) for non-smoothed data.

The correlation analysis shows that for the confederate-trained macaques the visibility of the other's action before making own decision correlated with the probability to follow the choice of the other in a trial-by-trial fashion, with the faster monkey "selfishly" selecting its own preference. Turn-taking in humans, however, did not rely on temporal competition to systematically steer the pair towards the faster agent's preferred color. To highlight this difference, we asked how the two species established and maintained the dynamic coordination (in case it was present)? The trial-by-trial sequences of choices suggested that confederate-trained macaques transitioned between periods of coordination using short periods of non-coordination (**Figure 8A**, magenta arrows), while turn-taking humans tended to switch seamlessly between the two coordination options (despite verbal communication being prohibited). **Figure 8B** illustrates possible uncoordinated and seamless switch patterns, sorting them to "selfish" and "benevolent". **Figure 8C-F** provides a detailed quantification of different aspects of transitions as scatter plots of benevolent choices on the y-axis and selfish choices on the x-axis, in the two species. **Figure 8C** compares the frequency of benevolent and selfish choices after the faster agent switched seamlessly from one coordination combination to the other. Turn-taking humans show a high and roughly equal amount of benevolent and selfish seamless switches, indicating overall balanced switching behavior. The four naïve macaque pairs showing the highest number of seamless switches all used a side-based strategy, which due to the color-to-side randomization (Methods) trivially generated seamless color switches. Confederate-trained macaques and non-turn-taking humans showed only very few seamless switches.

Alternatively to seamless switches agents can also end the current coordination block by switching to non-coordination (**Figure 8A,B**). **Figure 8D** shows, for each pair, the total numbers of own-own vs. other-other non-coordinated trials. These trials can be used to initiate a switch of coordination (cf. **Figure 2A**, from ~trial 275 to the end: in this human pair three coordination switches were initiated by "selfish" own-own trials and four by "benevolent" other-other trials). All

macaque pairs/sessions are below the unity-diagonal, indicating selfish preference to select own color. Turn-taking humans tend to show fewer non-coordinated trials and lay along the diagonal, indicating that in a half of trials there was a benevolent intention to initiate the switch to other's preferred color.

The above analysis and the inspection of choice sequences (cf. **Figure 8A**) shows that in the confederate-trained macaques coordination epochs were separated by epochs of non-coordination that looked like each agent "challenged" the other to accommodate. This poses the question whether there was a temporal component in how a pair initiated the challenging epochs and how they resolved those challenges. Of particular interest are the confederate-trained macaques because they showed a very strong dependence between the relative reaction time difference and the choice. Focusing on the transitions from coordinated to non-coordinated epochs and vice versa, **Figure S7** shows that the faster agent slowed down and the slower agent sped up during the transition to non-coordination. Conversely, the agents sped up to initiate the coordination on the own color, and the accommodating agents slowed down.

Figure 8E compares, for all coordination-to-non-coordination transitions, the number of own vs. other's color selections by the faster agent. Here all macaque pairs/sessions are located below the diagonal indicating that the faster agent tended to act selfish, while turn-taking humans mostly lay along the diagonal, again indicating balanced switching behavior. Finally, **Figure 8F** compares the number of times the faster agent selected own or other's color in trials when the pair when they resolved a challenge by switching from non-coordination to coordination. Confederate-trained macaque sessions and 6 out of 9 naïve macaque pairs fall below the diagonal indicating selfish choices of the faster agent. Turn-taking humans, however, lay along the diagonal indicating balanced challenge resolution.

Taken together this shows that in turn-taking human pairs there are no indications of competitive turn-taking, as the faster agents showed balanced selection of both coordination options. Furthermore, in 4 out of the 10 turn-taking pairs one agent was significantly faster than the other in coordinated trials as compared to non-coordinated trials, regardless if it was coordination on its own or other's color (**Table S3**). This means that such agent by necessity initiated switches to and from the own color; while being able to see the faster agent's action likely helped the slower agent to maintain coordination and to accomplish the observed high rate of seamless switches. In contrast, the confederate-trained macaques showed no seamless switches, but displayed temporal competition in which the faster agent led to his own color.
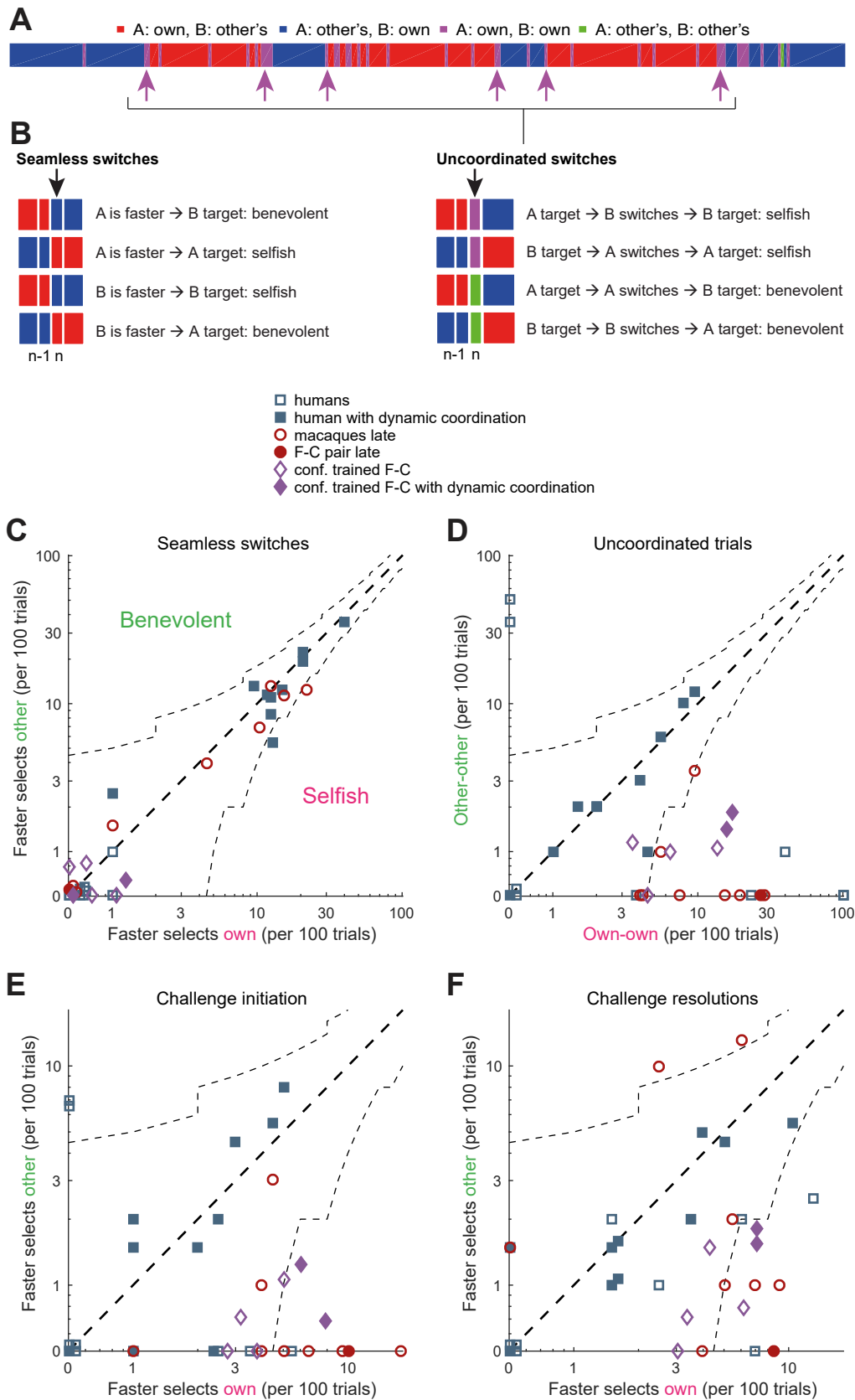
Figure 8. Humans and macaques show different switching behavior. (A) A magnification of the middle part of the macaque F-C session shown in Figure 6B (trials 211 - 495); color bars represent the choice combinations; the magenta arrows point to uncoordinated switches that resulted in a change of the coordination color. **(B)** 8 possible switch types, classified as selfish or benevolent based on which agent acted faster. Seamless switches involve uninterrupted change from one to another coordinated option; uncoordinated switches involve one or more uncoordinated trials. **(C)** The behavior of a faster agent for all seamless switches between the two coordination modes (red-to-blue and blue-to-red). Here and in E-F, the x-axis shows the number of own target selections (selfish), the y-axis the number of non-preferred target selections (benevolent). Turn-taking humans, as well as naïve monkey pairs that coordinated on one side, showed relatively high and balanced numbers of seamless switches, while the confederate-trained macaque pair and the non turn-taking humans showed only few seamless switches. Here and in D-F, the diagonal dashed line shows the line of equality, the black dashed contours show the 95% confidence interval around the diagonal. **(D)** The number of uncoordinated trials in which both subjects selected their respective preferred target (own-own) on the x-axis, and the number of trials in which both subjects selected their non-preferred target (other-other) on the y-axis. Turn-taking human pairs mainly cluster along the diagonal while naïve and trained macaque pairs cluster in the lower right half (mostly significantly below the diagonal); non turn-taking human pairs spread along the main axes. Thus in macaques uncoordinated trials that interrupt blocks of coordinated trials show selfish behavior while turn-taking humans show fair behavior with similar amounts of selfish and benevolent uncoordinated trials. **(E)** The switches from coordination to non-coordination, or "challenge initiations", for the faster agent. Monkey pairs fall below the diagonal (with the two competitive turn-taking sessions in pair FC even below the 95% confidence intervals around the diagonal). Turn-taking human pairs again cluster around the diagonal. **(F)** The behavior of the faster agent for all switches from non-coordination to coordination, or "challenge resolutions". Turn-taking humans cluster along the diagonal, while the trained monkey pair clustered below the diagonal (with the two competitive turn-taking sessions below 95% confidence interval around the diagonal).

Discussion

We studied macaque and human pairs in a coordination game, which offered higher rewards for selecting the same option, but entailed an inherent conflict about which of the two coordinated options to select. Both species largely converged on coordinated behavior but in a markedly different fashion. Many human pairs converged on nearly-optimal coordination and fair (cooperative) turn-taking that equalized the rewards of the two partners. Macaques, instead, tended to exploit simpler solutions that allowed them to maximize their reward without the need to track immediate actions of the partner, i.e. using static instead of dynamic coordination. The macaques' behavior however significantly shifted from such simple strategies to competitive turn-taking after they were trained to observe and attend to the other's choice with a human confederate. In post-confederate-training sessions, the choice behavior was highly correlated with trial-by-trial differences in the reach reaction times: the faster monkey chose his preferred option and the slower one followed. Our results show that both humans and macaques can take information about the other's action into account (when available) before making their own decisions. There was, however, a fundamental difference: when coordinating dynamically, monkeys showed competitive turn-taking while in humans the faster agent often "offered" switching to the partner's preferred color, exhibiting a form of benevolent and fair turn-taking.

Coordination in naïve humans and macaques

Half of the human pairs (10 of 19) converged on dynamic coordination that balanced the reward, developing fair, cooperative turn-taking, while another 5 pairs coordinated on the same fixed color throughout the session. Similarly to the approach of Brosnan and colleagues [28,37], the human subjects in the present study had to infer the underlying payoff rules while playing, to closely match the procedure in macaques. Even under such conditions, the majority of humans arrived at optimal behavior within a single session, in our study as well as in other coordination and anti-coordination games [28,31,37]. The observed turn-taking frequency was also close to other variants of BoS games with explicit instruction of the payoff matrix where approximately 60% of human subjects developed turn-taking [40–43], slightly higher than the 41% in the Conflict game [37].

In contrast to humans, macaque pairs converged on simpler strategies, either selecting the same color (56%) or selecting the same side (44%). These two strategies require less trial-by-trial coordination effort than the turn-taking. All but one macaque coordinated both via same color or same side selection depending on the partner, and no macaque persisted on the own color with all tested partners. For instance, monkey F insisted on his preferred color when playing with C, accommodated to the other's color when playing with T and converged on his right-side when playing with M. This indicates that macaques took the partner's actions into account, albeit not in the strict trial-by-trial fashion. Macaques started with insisting on their own color (reward size 2 drops), then one agent changed either to selecting the non-preferred color (increasing his reward to 3 and the other's to 4 drops) or to selecting one side (increasing his reward to an average of 2.5 and the other's to 3 drops). If the other monkey started favoring the same side, this increased both agents' reward up to 3.5. This sequence of intermediary steps illustrates how macaque pairs might develop their strategy over the course of several sessions.

While our study is to the best of our knowledge the first implementation of a classic BoS paradigm in macaques, other related 2×2 games have been investigated in different primates (humans, chimpanzees, macaques and capuchins). In some games human and nonhuman species show qualitatively comparable capabilities to coordinate choices and to maximize payoffs. This holds for simple coordination games such as Stag Hunt [28,31], as well as for the Prisoner's Dilemma, where macaques showed mutual cooperation significantly above chance and mutual defection below chance [36], similarly to humans [33]. Conversely, the behavior of humans and nonhuman primates differs when the competition can be resolved by turn-taking. In the anti-coordination Conflict game where maximum joint reward is obtained if each agent

selects a different icon [37], capuchins and macaques converged on only one of the two possible asymmetric equilibria (much like 4 of our 9 naïve macaque pairs), while 11 out of 27 human pairs balanced payoffs by alternating between the two anti-coordination equilibria (similar to the 10 turn-takers out of our 19 pairs).

Action visibility and dynamic coordination

The lack of spontaneous dynamic coordination in naïve macaques in our experiments cannot be attributed to their inability to infer the upcoming choices of the partner, because they could observe partner's actions through the transparent display. This setup is different from traditional economic games, which are either played simultaneously or sequentially. It has been demonstrated that coordination (or anti-coordination) typically is improved with increasing information about the other's choice by going from the strictly simultaneous to the sequential mode [53,54]. The transparent game approach we adopted here differs from those two classical modes in that each subject can decide independently when to act (within a certain time window). Brosnan and colleagues used a similar approach to show that compared to an opaque simultaneous setting, macaques and capuchins improved their anti-coordination in a transparent (called asynchronous) version of a Conflict game. In this study, subjects looked at the same monitor while sitting next to each other and could observe each other's decisions as cursor movements [37].

In our setup, agents sat opposite to each other and saw actual eye and hand movements, combining face-to-face and action visibility. In humans, face-to-face visibility in a simultaneous iterated Prisoner's Dilemma (iPD) and in the Ultimatum game significantly improves mutual cooperation, even without action visibility [55,56]. This supports the idea that non-verbal social signals and more generally, the observable presence of others influence decision-making. The ongoing action visibility per se also had a profound effect on mutual cooperation in iPD [57] and in a web interface-based anti-coordination game conceptually similar to BoS [47]. In a competitive reaching task, a face-to-face transparency allowed human subjects to glean useful information from observing the relevant hand effector and from seeing the face and the full body of an opponent [58]. Seeing actual movements rather than relying on abstract representations of others such as cursor motions might be especially crucial in nonhuman experiments.

The importance of action visibility is also supported by computational modeling. In the continuous-time cooperation games such as iPD and Stag Hunt where agents could observe and respond to each other's actions in real-time, cooperation by coaction is more easily

obtained and stabilized against exploitation than the cooperation that relies on delayed reciprocity [45]. For the iPD and BoS transparent games, we showed that different coordination strategies are preferable for different probabilities of seeing partner's choice [8,9]. When these probabilities are low (or when agents do not utilize the action visibility), simple strategies like "Win-stay, lose-shift" are most effective. For higher probabilities of seeing partner's choice more complex strategies emerge, such as coordinated turn-taking and temporal Leader-Follower, where the faster agent determines the choice of the slower agent. For the selfish agents this means that the faster one insists on its individually-preferred own target and the slower accommodates [8].

Given that many human pairs exhibited spontaneous dynamic coordination, while naïve macaques did not, we can ask if and how these two species utilized the action visibility. Our findings suggest that humans relied on action visibility for seamless switches from one color to another, and for maintaining the coordination within a block of trials. But besides one pair, there was no indication that humans employed the competitive variant of the Leader-Follower strategy. Instead, faster agents were as likely to switch from their own to the other's color, indicating benevolent turn-taking. Very different dynamics transpired in additional experiments with macaques. While naïve macaques did not seem to monitor each other's choices in each trial, pairing them with a human confederate playing a color alternation strategy in short blocks demonstrated that macaques are capable to closely follow the other's actions, and adopt the imposed "turn-taking". Moreover, temporarily blocking the view of the confederate's hands abruptly changed macaque behavior, disrupting the coordination. The macaques' coordination was thus driven by action visibility and not by simple strategies like win-stay-lose-shift or trial counting. In line with the Leader-Follower strategy, macaques were faster than the confederate when selecting own color, but waited for the confederate when selecting the confederate's color. Intriguingly, pairing the two macaques that completed the training with a human confederate resulted in behavior different from both naïve macaque behavior and from the "following behavior" with the confederate. The confederate-trained macaques competed, with the faster agent selecting the own color and the slower agent following, such that their choices depended on the relative difference between their reaction times. This behavior resulted in either sustained coordination on one fixed color or in turn-taking. Comparing the temporal signatures of the reaction time differences in trials around a switch to or from coordinating on a specific color further indicated that macaque turn-taking was competitive and dynamic in nature. The break of coordination was triggered by the faster agent slowing down and the slower one speeding up

and selecting the own color; conversely, the transition to coordination was associated with the speeding up agent selecting the own color and the slowing down agent accommodating.

In summary, both species exploited action visibility to achieve and/or maintain the turn-taking, emphasizing the importance of the "transparency" of interactive behavior as an important determinant of emerging strategies. Human agents initiated switches to and from their preferred colors to "fairly" balance the payoffs, similarly to the Conflict game [37], while macaques established competitive dynamics, as predicted by our evolutionary simulations [8]. These results add to the body of literature indicating that humans in a social setting might base their decisions not only on pure reward maximization [59]. It is also plausible that such normative behavior can reflect prospective planning to ensure a stable persistence of the individually-beneficial cooperation. The lack of the cooperative turn-taking in nonhuman primates, as compared to the humans' propensity to engage in it, might reflect cognitive limitations in long-term planning and perspective taking [60]. Species differences in general cooperative behavior may explain both the reluctance of naïve macaques to coordinate dynamically, as well as the competitive nature of dynamic interactions. The transparent BoS game requires the agents to cooperate in close proximity in the pursuit of an immediate reward which necessitates some degree of social tolerance by both partners. We can take natural food sharing behavior as a proxy for social tolerance. Food sharing behavior is not equally prevalent among primate species: and has been described for humans, apes, some baboon species [61], and some New World monkey species including capuchins share food even between adults, while rhesus macaques do not even share food with their offspring [62]. In line with these patterns, mutually beneficial alternating task performance and turn-taking has been observed in humans [23], apes [63–65], and capuchins [66], but not in more despotic, less tolerant rhesus macaques [5].

Limitations and future directions

There are several limitations in our study that have to be considered and possibly addressed in future experiments. Firstly, we only tested 6 macaques (and 9 pairs), all males, and only one pair was housed together (others were from neighboring but separate enclosures). Therefore, we cannot say much about the influence of social rank. Anecdotally, however, it did not seem that the more dominant agent always prevailed. For instance, one subordinate animal (monkey T), the smallest and likely to be subordinate to all other partners (cf. [67]), successfully "insisted" on his own color in 2 out of 3 cases. Furthermore, we only tested two confederate-trained macaques so far. While the results are exciting and show that macaques can engage in

dynamic competitive turn-taking, future experiments will need to test how generalizable this pattern is.

Secondly, due to the task target color/side randomization, selecting targets on the same side of the display resulted in efficient and fair coordination. We cannot say if fairness was a contributing factor in convergence of the four naïve macaque pairs to this strategy, although we deem it highly unlikely. A new experiment where the probability of a specific color to appear on the same side is parametrically modulated is needed to evaluate the conditions under which this strategy emerges (or what level of unfairness is accepted). Furthermore, it would be interesting to test human pairs in several sessions, similar to macaques, to see if humans might also converge on the fixed side strategy as means of fair and seamless coordination. It requires preventing the human participants from any verbal contact after and between the sessions however. Finally, it would be important to vary the ratio and the range of reward magnitude ("stakes"), as it has been shown that stakes may strongly affect the individual [68] and social [47] decisions.

Notwithstanding these limitations, our results contribute novel insights to understanding social decision processes as they unfold in real-time during transparent interactions, and offer a new route to further behavioral and neural investigations of dynamic decision-making in cooperative and competitive contexts.

Acknowledgements

Funding

Authors' contributions

JF, ST and IK wrote the key grants to support his study. SM, AMU, AG, ST and IK conceptualized and designed the task. SM implemented the task and conducted the experiments. SM and AMU analyzed the data. SM, AMU and IK prepared figures. SM, AMU and IK wrote the initial version of the manuscript. All authors discussed and interpreted the findings, and revised the manuscript.

Competing interests

The authors declare no competing interests.

Data availability

The datasets generated and analyzed in the current study are available from the corresponding author on reasonable request, and will be uploaded to Open Science Framework data repository (https://osf.io/).

Methods

Participants

Humans. 38 right-handed subjects (23 females, mean age: 26.1 $\pm$ 4.1 SD, range 20 to 41 years) participated in the study as paid volunteers. Subjects were tested as 19 unique pairs, i.e. each participant contributed only once. Instructions given to the subjects before the experiment are provided in **Supporting Information**. In short, subjects were instructed how to operate the setup and to interpret the auditory feedback as an indicator of the earned reward. They were not given an explicit description of the task's payoff structure beyond "Your reward will depend on your own and your partner's choice". Prior to the experiment, subjects were individually familiarized with the setup and practiced a single-player ("solo") version of the task. All subjects gave written informed consent for participation after the procedures had been explained to them and before taking part in the experiment. Experiments were performed in accordance with institutional guidelines for experiments with humans and adhered to the principles of the Declaration of Helsinki. The experimental protocol was approved by the ethics committee of the Georg-Elias-Mueller-Institute for Psychology, University of Goettingen (GEMI 17-06-06 171). We excluded 4 additional human subjects (2 pairs) from the analyses due to experimental differences (one pair did not perform the initial individual training; one pair ended the experiment prematurely).

Macaques. Research with nonhuman primates represents a small but indispensable component of neuroscience research. The scientists in this study are aware and are committed to the responsibility they have in ensuring the best possible science with the least possible harm to the animals [13].

Six adult male rhesus monkeys (designated by initials C, E, F, L, M, T) participated in the study, yielding 9 pairs (each monkey participated in 2 or more pairs). Animals were extensively trained with positive reinforcement to climb into and stay seated in a primate chair. The experimental procedures were approved of by the responsible regional government office (Niedersaechsisches Landesamt fuer Verbraucherschutz und Lebensmittelsicherheit (LAVES), permits 3392-42502-04-13/1100 and 3319-42502-04-18/2823). The animals were pair- or group-housed in facilities of the German Primate Center (DPZ) in accordance with all applicable German and European regulations. The facility provides the animals with an enriched environment, including a multitude of toys and wooden structures [69,70], natural as well as

artificial light and access to outdoor space, exceeding the size requirements of European regulations, and a rich diet including primate biscuits, fruit and vegetables. During the study the animals had unrestricted access to food and fluid, except on the days where data were collected or the animal was trained on the behavioral paradigm. On these days, the animals were allowed access to fluid through their performance in the behavioral paradigm. The DPZ's veterinarians, the animal facility staff and the lab's scientists carefully monitor the animals' welfare.

Experimental setup

For maximal comparability of the human and the monkey behavior, we developed a novel Dyadic Interaction platform in which two human or nonhuman primate subjects co-act in a shared workspace while sitting face-to-face (**Figure 1A, B**). Joint dyadic tasks have been previously implemented in a side-by-side setting with a shared or separate workspaces for each subject [3,36,37,71,72], or in a table-like settings with opposing subjects acting in a horizontal workspace that is not in the line of sight between the subjects' faces [73,74]. Only few studies utilized face-to-face arrangement of the subjects, using a video projector and two semitransparent mirrors virtually placing the stimuli into the shared plane [75], or incorporating physical targets into transparent Plexiglas screen [58]. Similarly, with our design we aimed for maximal availability of mutual social signaling in a face-to-face setting, a shared vertical workspace with computer-controlled stimuli in the direct line of sight, without risk of injuries due to physical contact, and suitability across species. We achieved this using a novel transparent display (1920 x 1080 pixels, 121 cm x 68 cm, 60 Hz, EYE-TOLED-5500, Eyevis, Reutlingen, Germany), amended for dual-side touch sensitivity (PQLabs, G5S, Freemont, CA, USA) with a custom-built "sandwich" construction. Sitting on either side of the display, both subjects saw each other and the same screen display. Two proximity sensors (Carlo Gavazzi CA18-CA30CAF16NA, Lainate, Italy) per subject for the left and right hand respectively, mounted below the screen ("home" buttons) and the two touch panels mounted on either side of the screen registered hand positions of both agents, at 240 Hz temporal resolution.

Experimental control and stimulus presentation were implemented using the EventIDE software package (Okazolab, Delft, The Netherlands). Liquid reward for monkeys was delivered via computer-controlled peristaltic fluid pumps, for every correctly performed trial.

Dyadic decision-making task

Bach-or-Stravinsky game. Since we are interested in the effect of mutual action visibility on coordination behavior, we implemented a transparent version of a Bach-or-Stravinsky (BoS) game, in which each player's time-continuous visuomotor behavior can be seen by the other player. Conceptionally, in the BoS game two agents are choosing between going to Bach or Stravinsky concerts. Agent A prefers Bach, Agent B prefers Stravinsky; yet, both prefer going to the concert together [76]. Thus, agents wish to coordinate their behavior but have conflicting interests. This is a classic 2×2 non-zero sum game (also known as the Battle of the Sexes) with two pure strategy Nash equilibria (Bach-Bach, or Stravinsky-Stravinsky), and one less efficient mixed strategy Nash equilibrium, where the agents go to their preferred event more often than the other [26,39].

In our implementation, the agents were choosing between two options represented by two differently colored targets placed left and right on the screen. Each target color was associated with a higher reward for one of the agents resulting in individually preferred targets. An agent selecting own preferred target was assured to get at least 2 reward units, while selecting the other target (individually non-preferred target) yielded at least 1 reward unit. Additionally, when both agents selected the same target, a bonus of 2 reward units was added to the payoff of each agent. Thus, the maximum average joint reward (3.5 units) was obtained on coordinated trials when one agent selected own preferred target (getting 4 reward units) while the other chose non-preferred target (getting 3 reward units). Note that on any given coordinated trial, i.e. when agents selected the same target for overall higher reward, this payoff matrix (**Figure 1C)** resulted in an unequal (i.e. unfair) reward distribution. Hence, the BoS paradigm probes both the ability to realize that coordinated target-selection results in higher rewards than non-coordinated target-selection, as well as the ability to perceive and counteract the unfairness/conflict situation.

The task. On every trial, each subject chose between the individually-preferred and non-preferred target on the touchscreen. Targets were light blue circles of 35 mm diameter with either a red (individually-preferred target for Agent A) or a yellow (individually-preferred target for Agent B) rings (for better visualization, we replaced yellow and red rings with blue and red solid targets in the figures and the text).

Subjects had to place both hands on the two home buttons for 500 ms to start a trial after the inter-trial interval (ITI). This allowed us to control which effector (acting hand) was used. (**Figure**

1D). Then an initial fixation target without a colored ring appeared on the screen (10 cm below eye level). Subjects had 1500 ms to touch this initial fixation target on the screen with the instructed hand. After both subjects touched the fixation target they had to hold it for 500 ms. Then two choice targets appeared at one of three different pairs of positions (140 mm to the left and right of the central fixation target and either at the same height as the central fixation target or 35 mm below or 35 mm above it). We randomized the "red" location equally over all six positions, balanced in sets of 18 trials with each of the six "red" positions appearing three times, and vertically mirrored the opposite location. Simultaneous with targets' appearance the initial fixation target disappeared, which served as a go signal. Subjects had 1500 ms to make their choice and touch one of the targets. After both subjects acquired their chosen target, selected target(s) brightened up, and both subjects needed to hold the hand on the target for another 500 ms. At that point, the choices were evaluated and rewards were dispensed according to the payoff matrix. The amount of reward earned by each subject was signaled by two sequential series of auditory pulses, with different pitch for each subject. Each pulse was constructed as a harmonic series with 12 overtones and a fundamental frequency of 443 Hz for side A and 733 Hz for side B to provide distinct sounds for each agent. For the monkeys, we immediately delivered water as reward (approximately 0.14 ml per pulse) concurrently with the auditory pulses for the respective monkey's side; humans were instructed to expect "a few" cents per pulse, the accumulated earnings were paid out as a lump sum after the experiment. After the reward period subjects had to wait for an inter-trial-interval of 1500 ms before they could initiate the next trial.

The side of the targets was randomized on each trial, i.e., in ~50% the red target was on the right, the blue on the left, and vice versa. This design can lead to the three following coordination patterns (or any mix of the three):

i)   Coordinating statically by each agent repeatedly selecting the same fixed color of the target, irrespective of unfair distribution of the rewards.

ii)   Coordinating statically by each agent repeatedly selecting the same fixed side. Due to color/side randomization, this pattern ensures a fair reward distribution.

iii) Coordinating dynamically by both agents selecting the same target, while picking from both colors and both sides trial-by-trial (e.g. trial 1 red right, trial 2 red left, trial 3 blue left, … etc.). This could result in a fair or unfair reward distribution depending on the ratio of red to blue color selection.

Human procedure. Human subjects were recruited via a university job website and pairs were selected based on matching schedules. Subjects were given a brief introduction to the experiment (see Supporting Information); this material included the information "Your reward will depend on your own and your partner's choice" and a very basic description of the task ("You will have to choose one of the two circles presented to you. [...] You will have to decide and respond quickly."), but did not include details of the payoff matrix. After the joint introduction, each subject alone performed 50 or 100 individual (solo) trials, to learn how to operate the touchscreen and to develop a preference for one of the two color targets (see SI for the exact verbal instructions given prior to solo training). Subjects were positioned ~50 cm from the display with the height of the chairs adjusted such that both subjects' eyes were ~121 cm above the ground. After the solo training, both subjects entered the setup for the main dyadic task which lasted for 300 or 400 trials. Participants had to infer the task rules by exploration, similar to the macaques [31]. Each session lasted approximately 1.5 hours. After the experiment we conducted an individual debriefing and paid the earned reward separately for each subject.

Macaque procedure. Macaques were brought to the setup in their individual primate chairs. The chairs were positioned such that the eyes were 30 cm from the display. The monkeys had previously been trained to perform the basic task structure (hands on proximity sensors, reach to the initial fixation target with the instructed hand, and select one of the presented choice targets by reaching to it). The animals performed the solo version of the task with differential rewards to develop a preference for one of the color targets and were only paired with a conspecific after selecting a higher rewarded target in ≥ 75% of the trials. Thereafter, pairs of macaques worked together in the dyadic version of the task, for 11 ± 7 sessions (range 4-25).

Data analysis

We computed the following six *aggregate measures of choice behavior and coordination,* explained below*: s*hares of "own" choices, shares of objective left choices, mutual information for target choice, mutual information for side choice, average reward, and dynamic coordination reward. We computed each of these measures for the last 200 trials of each session in order to assess the "steady-state" behavior after allowing for an initial period of exploration.

Shares of own and objective left choices. The share of own choices (SOC) is the fraction of trials where an agent has selected the individually preferred target. Similarly, the share of objective left choices (SLC) is a fraction of trials where an agent has selected the target on the "objective" left side of the screen which is the left side for Agent A and right side for Agent B.

Fractions range from 0 to 1. For figures showing SOC/SLC over the course of a session we also calculated both measures for the session as a whole, and in running windows of $w = 8$ trials. In the latter case, SOC and SLC can take the values of 0, 1/8, 2/8, …, 1. For instance, SOC = 0 means that an agent has selected individually non-preferred target for 8 trials in a row.

Mutual information. Mutual information (MI) represents the reduction of uncertainty regarding the values of one time series provided by knowing the values of the other time series. Here we consider mutual information for the color of the target (MIT) and side (MIS) choices of the two agents, showing how much information the target/side choices of one agent provide about the respective choices of the other. Mutual information is measured in bits. Since both target and side choices are binary (in each trial an agent selects either the preferred or non-preferred target and either the left or right side), both MIT and MIS range from 0 bit (the choices of one agent provide no information about the choices of the other) to 1 bit (the choices of one agent can be inferred precisely from the choices of the other agent).

For instance, if both agents select the objective left target in every odd trial and the objective right target in every even trial, both MIT = MIS = 1, since for every trial the target and the side selected by one agent entirely describes the target/side selected by the other. At the same time, the choices of each agent individually are highly uncertain: both sides and both targets are selected with the same probability of 0.5. If both agents constantly select the objective left side, this would result in MIT = 1 due to side-randomization of target color, but MIS = 0, since there is no uncertainty regarding the side selection and thus no additional knowledge about the other's choice can reduce the uncertainty.

Formally, mutual information of time series $X = (X_t)$ and $Y = (Y_t)$ is given by

$$\text{MI}(X,Y) = \sum_{x,y} p(X_t = x, Y_t = y) \log_2 \frac{p(X_t = x, Y_t = y)}{p(X_t = x) p(Y_t = y)}$$

where $p(X_t = x)$ is the probability of the value x in time series X, $p(Y_t = y)$ is the probability of the value y in time series Y, and $p(X_t = x, Y_t = y)$ is the joint probability to simultaneously have values x and y in time series X and Y, respectively, and x and y can be either 0 or 1, so that the sum is over all four combinations.

Since in our case time series X and Y have finite length, we simply replace probabilities by relative frequencies. This is known as a naïve estimation of mutual information, but it is sufficiently precise for binary time series [77,78].

To test whether the MI values were significantly different from zero, we generate Whittle surrogates for the given choice time series and estimate from them the threshold for the given significance level (p=0.01 in our case) [79].

Average reward. Average reward (AR) is computed as the average of an agent's payoff across the session. Note that the average reward of each individual agent can be in the range of 1 to 4 points, while the average joint reward of a pair cannot exceed 3.5 (since when one agent gets payoff of 4, the other agent gets only 3). The average reward for completely independent choices of two agents with 50% probability for either target is 2.5 (but note that an achieved reward of 2.5 is not a positive proof of independent choices).

Dynamic coordination reward. Dynamic coordination reward (DCR) is the surplus reward of the two agents compared to the reward they would get by playing randomly. By *playing randomly* we mean that choices of the agents in each round are independent of the history and of the current choices of the partner. The "random reward" is computed by selecting the two color targets with the same probabilities as actually observed in the two agents, but randomly permuting the choices over trials. For our payoff matrix, the range of DCR is [-1,1], with -1 corresponding to very inefficient playing (alternating selection of the two anti-coordination options), while 1 corresponds to very efficient playing with explicit coordination (for instance, turn-taking). DCR is hence a measure of dynamic (reciprocal) coordination. For instance, if both agents would coordinate statically by constantly selecting one and the same side, this would result in $DCR = 0$, even though this coordination pattern still yields the maximum average reward of 3.5.

Formally, DCR is defined as the actual average reward of a pair (R_{actual}) minus the reward the agents would get if they were playing randomly (R_{PR}). $DCR = R_{\text{actual}} - R_{\text{PR}}$. The reward for playing randomly (R_{PR}) depends only on eight probabilities, four for each agent. Below index i indicates the agent and stands for either A or B:

$P_{i,1,\text{left}}$ - probability to select non-preferred objective left target (which is the left side for agent A and right side for agent B),

$P_{i,1,\text{right}}$ - probability to select non-preferred objective right target,

$P_{i,2,\text{left}}$ - probability to select preferred objective left target,

$P_{i,2,\text{right}}$ - probability to select preferred objective right target.

36

Note that these probabilities are not independent on each other. First, for $i = A, B$, it holds

$$P_{i,1,\text{left}} + P_{i,2,\text{left}} + P_{i,1,\text{right}} + P_{i,2,\text{right}} = 1.$$

Second, it holds

$$P_{A,2,\text{left}} + P_{A,1,\text{right}} = P_{B,1,\text{left}} + P_{B,2,\text{right}} = Q_{\text{left}}$$

and

$$P_{A,1,\text{left}} + P_{A,2,\text{right}} = P_{B,2,\text{left}} + P_{B,1,\text{right}} = Q_{\text{right}}$$

where Q_{left} and Q_{right} are the probability of agent A's preferred target to appear on the left and on the right, respectively (and of B's preferred target too appear on the right and on the left). Given the independent random selection from trial to trial described above, Q-values should approximate 0.5 for larger N of trials.

Average reward of two agents for playing randomly is computed as follows:

$$R_{\text{PR}} = 3.5(p_{4,3} + p_{3,4}) + 2p_{2,2} + p_{1,1}$$

where $p_{a,b}$ is the probability that by random playing agent A gets reward a and agent B reward b. These probabilities are given by the following equations:

$$p_{1,1} = \frac{P_{A,1,\text{right}} \cdot P_{B,1,\text{left}}}{Q_{\text{left}}} + \frac{P_{A,1,\text{left}} \cdot P_{B,1,\text{right}}}{Q_{\text{right}}}$$

$$p_{2,2} = \frac{P_{A,2,\text{left}} \cdot P_{B,2,\text{right}}}{Q_{\text{left}}} + \frac{P_{A,2,\text{right}} \cdot P_{B,2,\text{left}}}{Q_{\text{right}}}$$

$$p_{4,3} = \frac{P_{A,2,\text{left}} \cdot P_{B,1,\text{left}}}{Q_{\text{left}}} + \frac{P_{A,2,\text{right}} \cdot P_{B,1,\text{right}}}{Q_{\text{right}}}$$

$$p_{3,4} = \frac{P_{A,1,\text{right}} \cdot P_{B,2,\text{right}}}{Q_{\text{left}}} + \frac{P_{A,1,\text{left}} \cdot P_{B,2,\text{left}}}{Q_{\text{right}}}$$

To see why this is the case, consider, for instance, $p_{1,1}$. Both agents get reward of 1 when they both select the other's preferred target, either when target of agent A appears on the left side (probability of this is encoded by the first term) or on the right side (second term).

To compute confidence intervals, we use the fact that SLC_i and SOC_i for an agent i can be considered as binomially distributed, thus radiuses of their confidence intervals ΔSLC_i and ΔSOC_i can be estimated by the classic method of approximating the distribution of error around binomially-distributed observation with a normal distribution. To obtain the confidence interval for DCR, it is sufficient to compute maximal and minimal possible DCR given that SOC and SCL of the two agents are within the respective confidence intervals. Note that $SLC_i = P_{i,1,\text{left}} + P_{i,2,\text{left}}$, $SOC_i = P_{i,2,\text{left}} + P_{i,2,\text{right}}$, thus all the probabilities necessary for calculating DCR can be computed from SLC_i and SOC_i given that Q_{left} and Q_{right} are fixed to 0.5 in the reported experiments. Simple analysis reveals that minimum and maximum DCR should be at the edges of the 4-D confidence interval formed by SLC_i and SOC_i of the two agents, which reduces the problem to testing 16 DCR values computed for $SLC_i \pm \Delta SLC_i$ and $SOC_i \pm \Delta SOC_i$.

For the main analysis presented in the Results, we used last 200 trials to compute DCR values. Using the last 150 or 250 trials for DCR analysis resulted in the exact same 10 human pairs with significant results as using the last 200.

Reaction and movement time measurements. We measured the time from the onset of the choice targets to the release time of the initial fixation target (t_{release}) and to the acquisition time of the selected target, $t_{\text{acquisition}}$ individually for each subject. We then calculated the movement time t_{movement} and reaction time t_{reaction} as follows:

$$t_{\text{movement}} = t_{\text{acquisition}} - t_{\text{release}}$$

$$t_{\text{reaction}} = t_{\text{release}} + \frac{t_{\text{movement}}}{2}$$

We use t_{reaction}, the half-way point between t_{release} and $t_{\text{acquisition}}$, i.e. the duration from target stimulus onset to the half-time of the reach movement, as a proxy for the estimated time at which the trajectory of each subject's reach movement should be evident for the other agent. Typical movement time values (mean ± SD across trials) in our experiment ranged from 314 ± 104 ms for humans, 171 ± 77 ms for macaques, and 180 ± 64 ms for the confederate trained macaque pair.

Estimated probability to see partner's choice. When an agent is acting slower than the partner, there is a chance for this agent to see the partner's choice and use this information for making the own choice. We therefore modeled the probability to see the partner's choice as a

logistic function of the difference of the agents' reaction times, in each trial. The logistic function $p_{\text{see}}(t)$ had an inflection point at 50 ms ($p_{\text{see}}(50) = 0.5$) and reached its plateau phase at 150 ms ($p_{\text{see}}(150) = 0.98$). Formally, the function was given by the following equation

$$p_{\text{see}}(t) = \frac{1}{1 + e^{-k(t-\Delta T_0)}}$$

with $k = 0.04$ (steepness of the slope) and $\Delta T_0 = 50$ ms (inflection point). The values of p_{see} were used for the analysis of correlation with probability of selecting the other's color (cf. **Figure 7**). Importantly, we also tested a wide range of these parameters (k fixed at 0.04, $\Delta T_0 = \{12.5, 25, 50, 75, 100, 200\}$ ms and ΔT_0 fixed at 50 ms, $k = \{0.01, 0.02, 0.04, 0.08, 0.16\}$), and confirmed that the values of correlation were robust in respect to these parameters; only the longest $\Delta T_0 = 200$ ms resulted in a noticeable drop of resultant correlations.

Abbreviations

AR   Average reward

BoS   Bach or Stravinsky game (also known as the Battle of the Sexes)

DCR   Dynamic coordination reward

MI   Mutual information

MIS   Mutual information between side choices

MIT   Mutual information between target color choices

(i)PD   (iterated) Prisoner's dilemma

RT   Reaction time

SOC   Share of own choices

SLC   Share of left choices

References

1. Byrne, R. W. & Bates, L. A. Primate Social Cognition: Uniquely Primate, Uniquely Social, or Just Unique? *Neuron* **65**, 815–830 (2010).

2. Jaeggi, A. V., Burkart, J. M. & Van Schaik, C. P. On the psychology of cooperation in humans and other primates: combining the natural history and experimental evidence of prosociality. *Philos Trans R Soc Lond B Biol Sci* **365**, 2723–2735 (2010).

3. Chang, S. W. C. *et al.* Neuroethology of primate social behavior. *Proc. Natl. Acad. Sci. U.S.A.* **110 Suppl 2**, 10387–10394 (2013).

4. Platt, M. L., Seyfarth, R. M. & Cheney, D. L. Adaptations for social cognition in the primate brain. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **371**, 20150096 (2016).

5. Joly, M. *et al.* Comparing physical and social cognitive skills in macaque species with different degrees of social tolerance. *Proc Biol Sci* **284**, (2017).

6. Waal, F. de & Waal, F. B. M. *Chimpanzee Politics: Power and Sex Among Apes.* (JHU Press, 2007).

7. Frith, U. & Frith, C. The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 165–176 (2010).

8. Unakafov, A. M. *et al.* Emergence and suppression of cooperation by action visibility in transparent games. *PLOS Computational Biology* **16**, e1007588 (2020).

9. Unakafov, A. M. *et al.* Evolutionary successful strategies in a transparent iterated Prisoner's Dilemma. *EvoApplications 2019, Lecture Notes in Computer Science* **11454**, 204–219 (2019).

10. de Waal, F. B. M. & Ferrari, P. F. Towards a bottom-up perspective on animal and human cognition. *Trends Cogn. Sci. (Regul. Ed.)* **14**, 201–207 (2010).

11. Santos, L. R. & Rosati, A. G. The Evolutionary Roots of Human Decision Making. *Annu. Rev. Psychol.* **66**, 321–347 (2015).

12. Brosnan, S. F. Insights into human cooperation from comparative economics. *Nature Human Behaviour* **2**, 432–434 (2018).

13. Roelfsema, P. R. & Treue, S. Basic Neuroscience Research with Nonhuman Primates: A Small but Indispensable Component of Biomedical Research. *Neuron* **82**, 1200–1204 (2014).

14. Chang, S. W. C. An Emerging Field of Primate Social Neurophysiology: Current Developments. *eNeuro* **4**, ENEURO.0295-17.2017 (2017).

15. Isoda, M., Noritake, A. & Ninomiya, T. Development of social systems neuroscience using macaques. *Proceedings of the Japan Academy, Series B* **94**, 305–323 (2018).

16. Wittmann, M. K., Lockwood, P. L. & Rushworth, M. F. S. Neural Mechanisms of Social Cognition in Primates. *Annual Review of Neuroscience* **41**, 99–118 (2018).

17. von Neumann, J. & Morgenstern, O. *Theory of Games and Economic Behavior*. (Princeton University Press, 1944).

18. Rilling, J. K. & Sanfey, A. G. The neuroscience of social decision-making. *Annu Rev Psychol* **62**, 23–48 (2011).

19. Smith, J. M. Game theory and the evolution of behaviour. *Proc. R. Soc. Lond., B, Biol. Sci.* **205**, 475–488 (1979).

20. Sanfey, A. G. Social Decision-Making: Insights from Game Theory and Neuroscience. *Science* **318**, 598–602 (2007).

21. *Neuroeconomics: decision making and the brain*. (Elsevier/AP, Academic Press is an imprint of Elsevier, 2014).

22. Tremblay, S., Sharika, K. M. & Platt, M. L. Social Decision-Making and the Brain: A Comparative Perspective. *Trends in Cognitive Sciences* **21**, 265–276 (2017).

23. Rapoport, A., Guyer, M. J. & Gordon, D. G. *The 2x2 Game*. (University of Michigan Press, 1976).

24. Rapoport, A. *Two-Person Game Theory*. (University of Michigan Press, 1973).

25. Bruns, B. R. Names for Games: Locating 2 × 2 Games. *Games* **6**, 495–520 (2015).

26. Rapoport, A. Exploiter, leader, hero, and martyr: The four archetypes of the 2 × 2 game. *Syst. Res.* **12**, 81–84 (1967).

27. Skyrms, B. *The Stag Hunt and the Evolution of Social Structure*. (Cambridge University Press, 2003).

28. Brosnan, S. F. *et al.* Responses to the Assurance game in monkeys, apes, and humans using equivalent procedures. *PNAS* **108**, 3442–3447 (2011).

29. Bullinger, A. F., Wyman, E., Melis, A. P. & Tomasello, M. Coordination of chimpanzees (Pan troglodytes) in a stag hunt game. *International Journal of Primatology* **32**, 1296–1310 (2011).

30. Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K. & Tomasello, M. Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings of the Royal Society of London B: Biological Sciences* **281**, 20141973 (2014).

31. Brosnan, S. F., Wilson, B. J. & Beran, M. J. Old World monkeys are more similar to humans than New World monkeys when playing a coordination game. *Proceedings of the Royal Society of London B: Biological Sciences* **279**, 1522–1530 (2012).

32. Parrish, A. E., Brosnan, S. F., Beran, B. J. & Würsig, M. J. Differential Responding by Rhesus Monkeys (Macaca mulatta) and Humans (Homo sapiens) to Variable Outcomes in the Assurance Game. *Animal Behavior and Cognition* **1**, 215 (2014).

33. Wedekind, C. & Milinski, M. Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 2686–2689 (1996).

34. Rilling, J. K. *et al.* A Neural Basis for Social Cooperation. *Neuron* **35**, 395–405 (2002).

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

36. Haroush, K. & Williams, Z. M. Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* **160**, 1233–1245 (2015).

37. Brosnan, S. F. *et al.* Human and monkey responses in a symmetric game of conflict with asymmetric equilibria. *Journal of Economic Behavior & Organization* **142**, 293–306 (2017).

38. Noë, R. Cooperation experiments: coordination through communication versus acting apart together. *Animal Behaviour* **71**, 1–18 (2006).

39. Kilgour, D. M. & Fraser, N. M. A taxonomy of all ordinal 2 × 2 games. *Theor Decis* **24**, 99–117 (1988).

40. Bornstein, G., Budescu, D. & Zamir, S. Cooperation in intergroup, N-person, and two-person games of chicken. *Journal of Conflict Resolution* **41**, 384–406 (1997).

41. Sonsino, D. & Sirota, J. Strategic pattern recognition—experimental evidence. *Games and Economic Behavior* **44**, 390–411 (2003).

42. Melis, A. P., Grocke, P., Kalbitz, J. & Tomasello, M. One for you, one for me: Humans' unique turn-taking skills. *Psychological science* **27**, 987–996 (2016).

43. Grueneisen, S. & Tomasello, M. Children coordinate in a recurrent social dilemma by taking turns and along dominance asymmetries. *Developmental psychology* **53**, 265 (2017).

44. Dugatkin, L. A., Mesterton-Gibbonsand, M. & Houston, A. I. Beyond the prisoner's dilemma: Toward models to discriminate among mechanisms of cooperation in nature. *Trends in ecology & evolution* **7**, 202–205 (1992).

45. van Doorn, G. S., Riebli, T. & Taborsky, M. Coaction versus reciprocity in continuous-time models of cooperation. *Journal of Theoretical Biology* **356**, 1–10 (2014).

46. McDonald, K. R., Broderick, W. F., Huettel, S. A. & Pearson, J. M. Bayesian nonparametric models characterize instantaneous strategies in a competitive dynamic game. *Nature Communications* **10**, 1–12 (2019).

47. Hawkins, R. X. D. & Goldstone, R. L. The Formation of Social Conventions in Real-Time Environments. *PLoS One* **11**, (2016).

48. Rizzolatti, G., Fogassi, L. & Gallese, V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature reviews. Neuroscience* **2**, 661–70 (2001).

49. Silk, J. B. The strategic dynamics of cooperation in primate groups. *Advances in the Study of Behavior* **37**, 1–41 (2007).

50. Sueur, C. & Petit, O. Signals use by leaders in Macaca tonkeana and Macaca mulatta: group-mate recruitment and behaviour monitoring. *Animal cognition* **13**, 239–248 (2010).

51. Conty, L., Dezecache, G., Hugueville, L. & Grèzes, J. Early Binding of Gaze, Gesture, and Emotion: Neural Time Course and Correlates. *J. Neurosci.* **32**, 4531–4539 (2012).

52. Fruteau, C., van Damme, E. & Noë, R. Vervet Monkeys Solve a Multiplayer "Forbidden Circle Game" by Queuing to Learn Restraint. *Current Biology* **23**, 665–670 (2013).

53. Rapoport, A. Order of Play in Strategically Equivalent Games in Extensive Form. *Int. J. Game Theory* **26**, 113–136 (1997).

54. Brocas, I., Carrillo, J. D. & Sachdeva, A. The path to equilibrium in sequential and simultaneous games: A mousetracking study. *Journal of Economic Theory* **178**, 246–274 (2018).

55. Tang, H. *et al.* Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Soc Cogn Affect Neurosci* **11**, 23–32 (2016).

56. Jahng, J., Kralik, J. D., Hwang, D.-U. & Jeong, J. Neural dynamics of two players when using nonverbal cues to gauge intentions to cooperate during the Prisoner's Dilemma Game. *NeuroImage* **157**, 263–274 (2017).

57. Friedman, D. & Oprea, R. A Continuous Dilemma. *American Economic Review* **102**, 337–363 (2012).

58. Vaziri-Pashkam, M., Cormiea, S. & Nakayama, K. Predicting actions from subtle preparatory movements. *Cognition* **168**, 65–75 (2017).

59. Clavien, C. & Chapuisat, M. The evolution of utility functions and psychological altruism. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* **56**, 24–31 (2016).

60. Stevens, J. R. & Hauser, M. D. Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci. (Regul. Ed.)* **8**, 60–65 (2004).

61. Goffe, A. S. & Fischer, J. Meat sharing between male and female Guinea baboons (*Papio papio*). *Primate Biology* **3**, 1–8 (2016).

62. Jaeggi, A. V. & Gurven, M. Natural cooperators: food sharing in humans and other primates. *Evol. Anthropol.* **22**, 186–195 (2013).

63. de Waal, F. B. Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution* **18**, 433–459 (1989).

64. de Waal, F. B. The Law of the Jungle: Conflict Resolution in Primates. in *Roads to Reconciliation: Conflict and Dialogue in the Twenty-First Century* 135–148 (Routledge, 2014).

65. Martin, C. F., Biro, D. & Matsuzawa, T. Chimpanzees spontaneously take turns in a shared serial ordering task. *Scientific Reports* **7**, 14307 (2017).

66. Parrish, A., Brosnan, S. & Beran, M. Capuchin monkeys alternate play and reward in a dual computerized task. *Animal Behavior and Cognition* **2**, 334–347 (2015).

67. Massen, J. J. M., van den Berg, L. M., Spruijt, B. M. & Sterck, E. H. M. Generous Leaders and Selfish Underdogs: Pro-Sociality in Despotic Macaques. *PLOS ONE* **5**, 1–5 (2010).

68. Schultz, W. Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews* **95**, 853–951 (2015).

69. Calapai, A. *et al.* A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. *Behavior research methods* **49**, 35–45 (2017).

70. Berger, M. *et al.* Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. *J. Neurophysiol.* **119**, 796–807 (2018).

71. Fujii, N., Hihara, S. & Iriki, A. Dynamic Social Adaptation of Motion-Related Neurons in Primate Parietal Cortex. *PLoS ONE* **2**, e397 (2007).

72. Falcone, R., Cirillo, R., Ferraina, S. & Genovesio, A. Neural activity in macaque medial frontal cortex represents others' choices. *Scientific Reports* **7**, (2017).

73. Yoshida, K., Saito, N., Iriki, A. & Isoda, M. Representation of Others' Action by Neurons in Monkey Medial Frontal Cortex. *Current Biology* **21**, 249–253 (2011).

74. Báez-Mendoza, R., Coeverden, C. R. van & Schultz, W. A neuronal reward inequity signal in primate striatum. *Journal of Neurophysiology* **115**, 68–79 (2016).

75. Ballesta, S. & Duhamel, J.-R. Rudimentary empathy in macaques' social decision-making. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15516–15521 (2015).

76. Osborne, M. J. & Rubinstein, A. *A Course in Game Theory*. (The MIT Press, 1994).

77. Grassberger, P. Finite sample corrections to entropy and dimension estimates. *Physics Letters A* **128**, 369–373 (1988).

78. Steuer, R., Kurths, J., Daub, C. O., Weise, J. & Selbig, J. The mutual information: detecting and evaluating dependencies between variables. *Bioinformatics* **18**, S231–S240 (2002).

79. Pethel, S. & Hahs, D. Exact test of independence using mutual information. *Entropy* **16**, 2839–2849 (2014).

Supporting Information

Supporting information contains

7 Supporting figures

7 Supporting tables

Instructions for human subjects

Supporting Figures
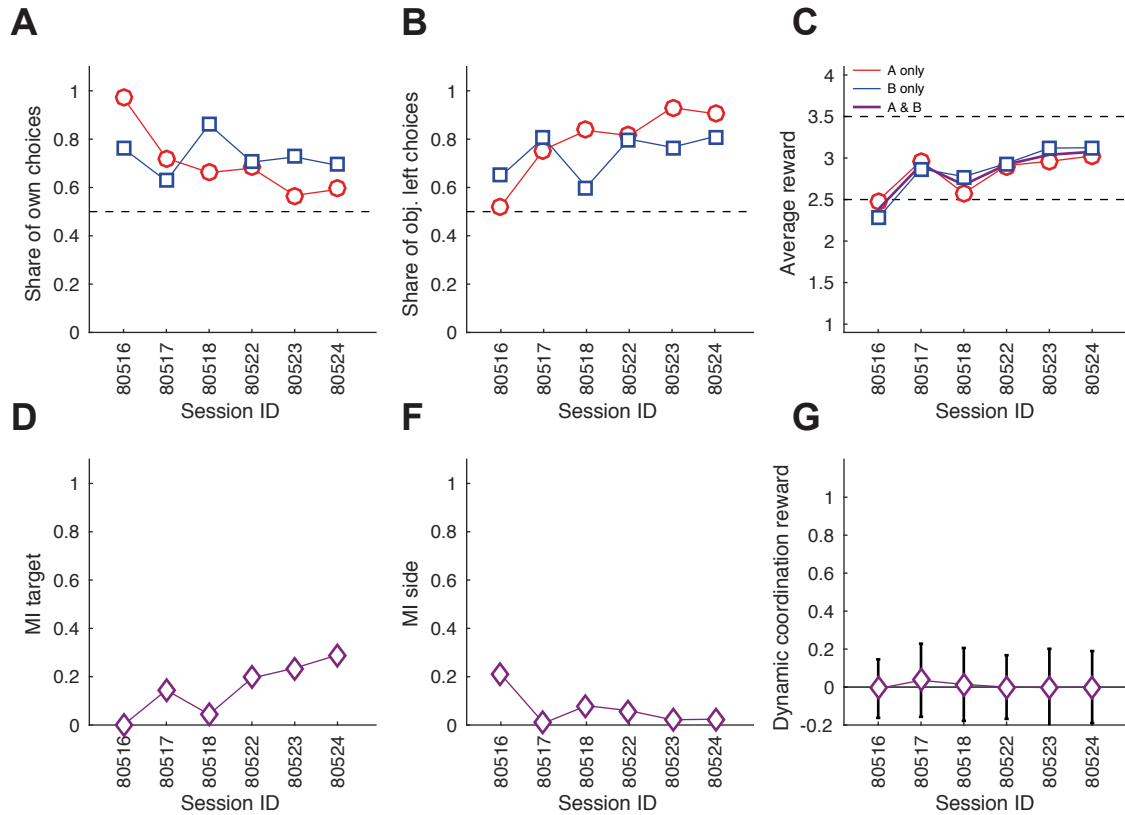
Supporting Figure S1. Exemplary development of choice behavior in one macaque pair TE (same pair as in Figure 2) over 6 sessions. (A-F) The change in aggregate measures of choice behavior and coordination over the course of 6 sessions for the last 200 trials of each session. Conventions as in Figure 2C-H. There was a clear development from selecting own targets to coordination by virtue of converging on mostly selecting the target on the same side of the transparent display. This is also reflected in the gradual increase of the target MI and the gradual decrease of the side MI. The zero DCR values indicate no significant dynamic coordination.
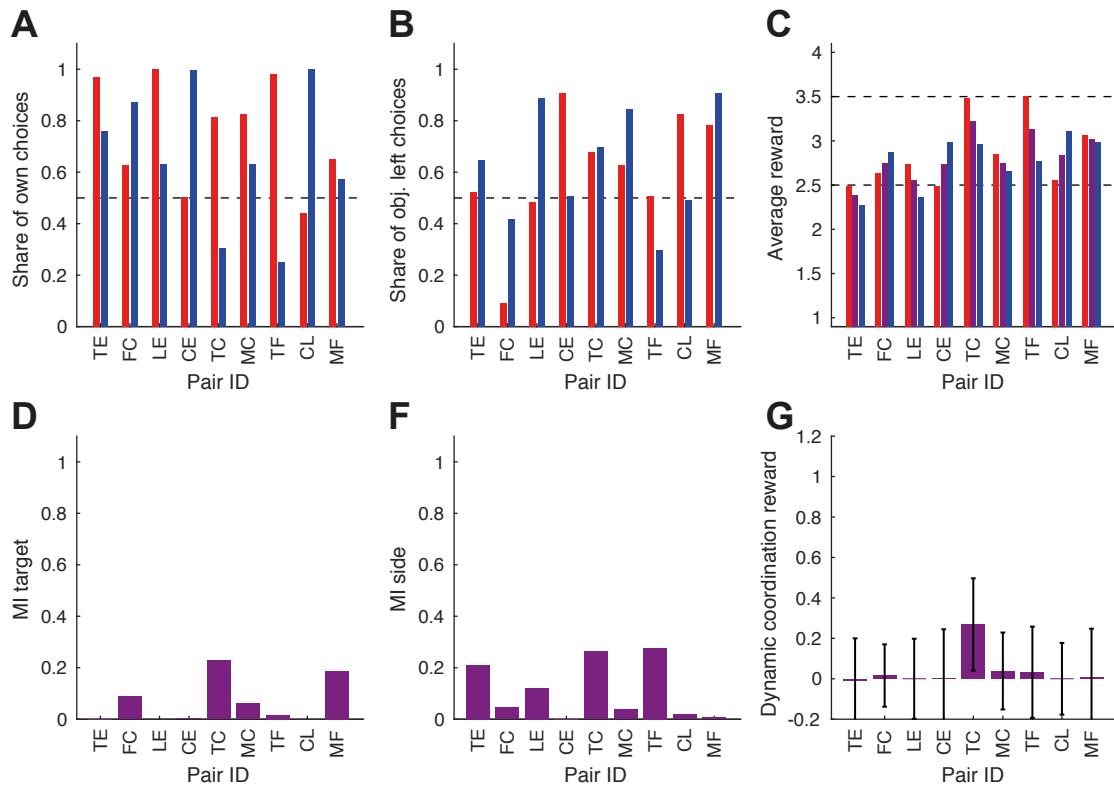
Supporting Figure S2. Performance of naïve monkey pairs sorted as in Figure 3E-J, in early sessions. Conventions as in Figure 3E-J.
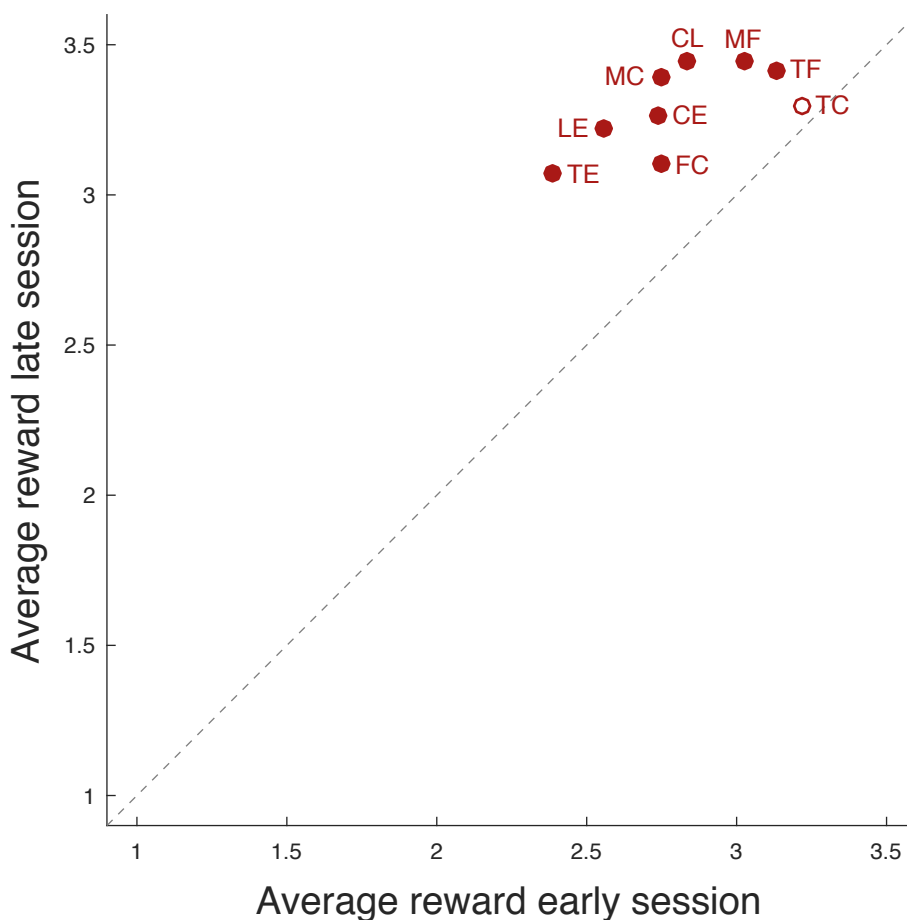
Supporting Figure S3. Paired training increases average joint reward in naïve monkey pairs. Scatter plot of the average reward for each macaque pair for an early versus a late session. Wilcoxon signed rank test showed that across all 9 pairs the average reward increased from early (median 2.75) to late sessions (median 3.29), Z: -2.67, $p < 0.0077$, r = -0.89. Note that 1 is the minimal possible average joint reward and 3.5 is the maximum. In all but one pair (TC, open circle), training increased the efficiency of coordination significantly ($p < 0.05$, Fisher's exact test of the proportion of coordinated versus uncoordinated trials in the early and the late session).
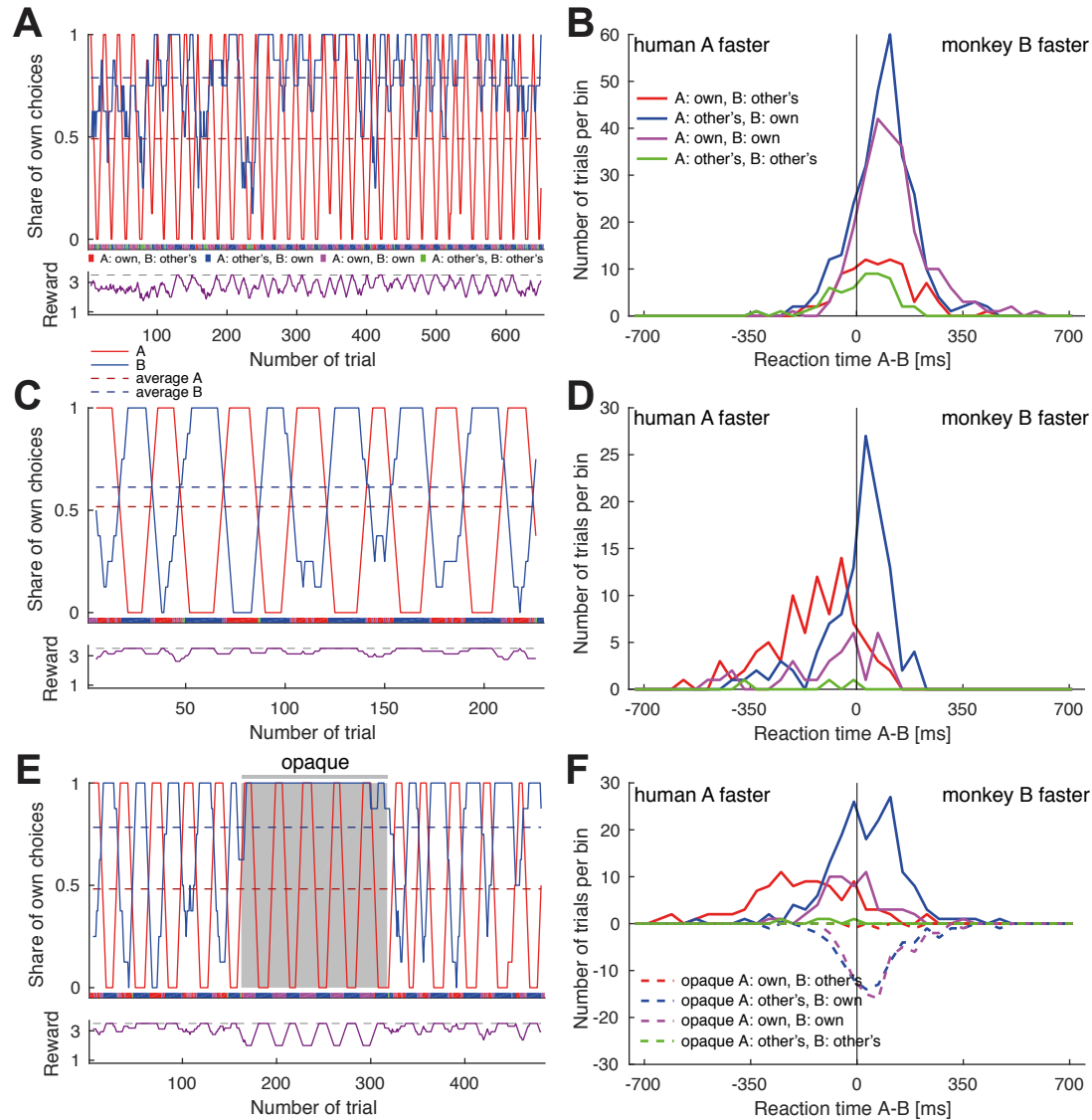
Supporting Figure S4. Macaques can be trained to take turns if paired with a human confederate. Same as the Figure 5, for monkey C. **(A,B)** Early confederate-training session, **(C,D)** late confederate-training session. The reaction time difference histograms of the late session show that in blocks when the pair selected the macaque's non-preferred target (red curve), the monkey acted slower than the confederate, waiting to see the human's choice before committing to its own choice (t-test: coordination on red (mean: -142 ms, SD: 138, N: 86) vs. blue (mean: 0.5 ms, SD: 119, N: 108), t(168.4): -7.6, p: 2*10[-12]). **(E,F)** Monkey C also ceased to take turns if an opaque screen was placed over the region of the transparent display containing all touch targets (gray background in E) and mainly selected his own target. The reaction time difference histogram in (F) show that the transparent condition the behavior was similar to (D) in that the monkey waited for the human to act first in blocks of the monkey's non-preferred target (t-test: coordination on red (mean: -192 ms, SD: 180, N: 98) vs. blue (mean: 27 ms, SD: 130, N: 169), t(156.3): -10.6, p: 5*10[-20]).
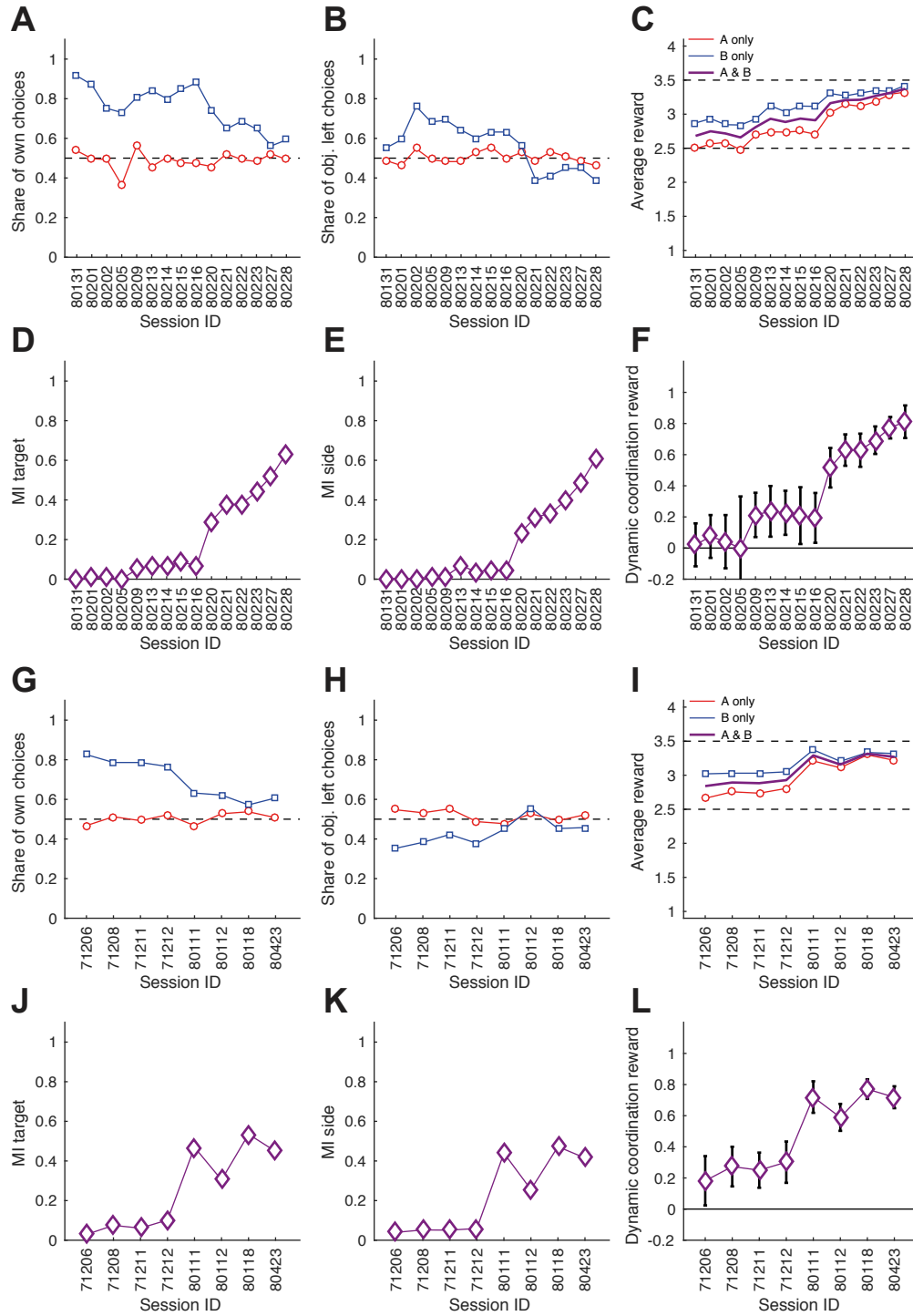
Supporting Figure S5. Monkeys gradually learn to follow a human confederate's turn-taking behavior. (A-F) The development of aggregate measures of choice behavior and coordination for monkey F (same conventions as Figure 2C-H). Note the gradual increase in average rewards, and the overall slow changes in the first 9 sessions as compared to the steeper increases in the MIs and the dynamic coordination reward (DCR) in the last 6 sessions. **(G-L)** Corresponding data for monkey C.
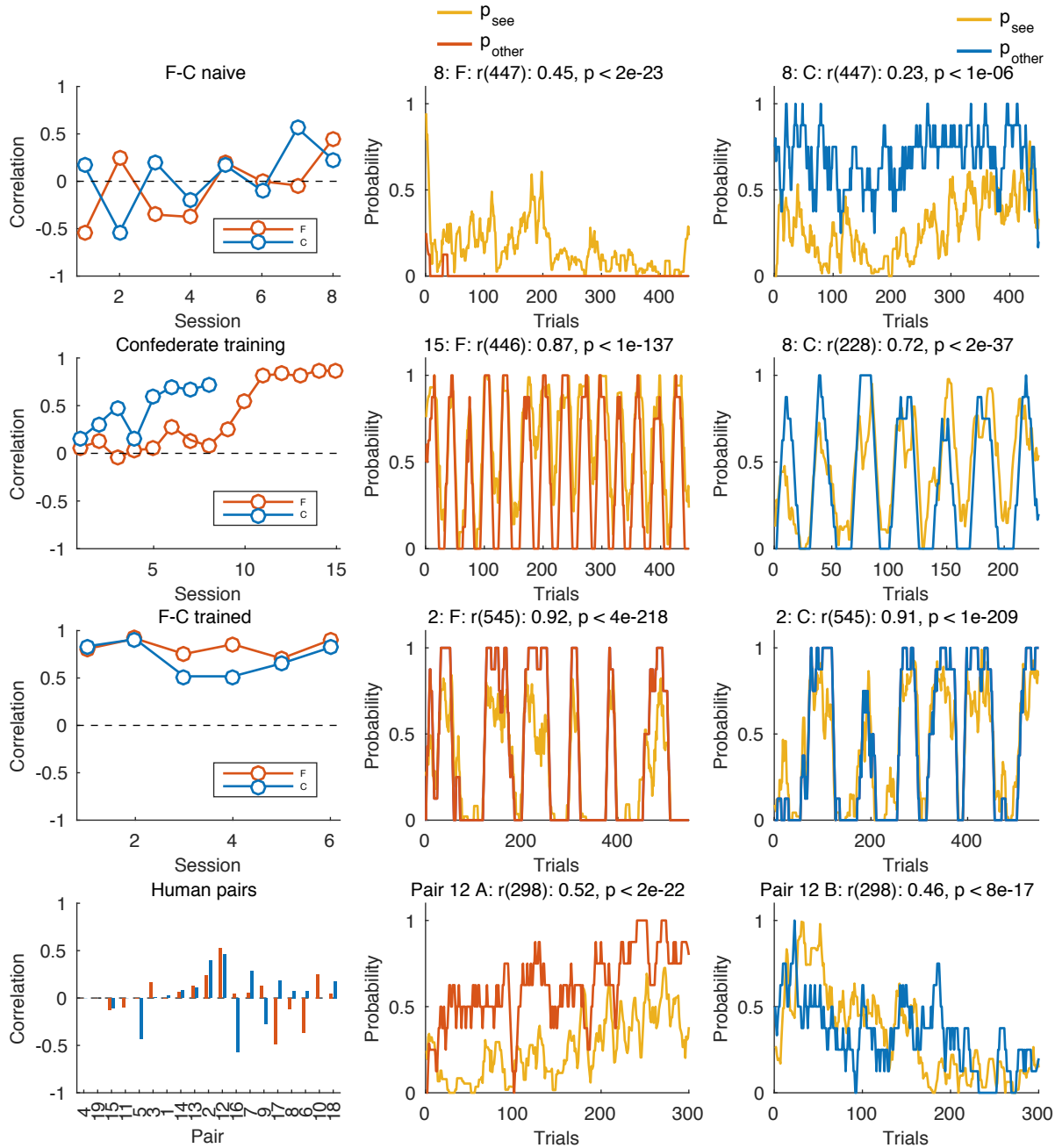
Supporting Figure S6. Correlations between the probability to see the partner's choice p_see and the likelihood to select the partner's target p_other. The left column shows the correlations for different sessions of the same pair of macaques (FC) before (F-C naïve), during and after the confederate training, as well as different human pairs. The middle and right columns show p_see versus p_other for one exemplary session for each group. Note that training with a confederate increased magnitude of the correlation (Confederate training) and how this carried into the post-confederate training sessions of the two macaques (F-C trained). Tables S4 to S7 contain the statistical information for each individual correlation.
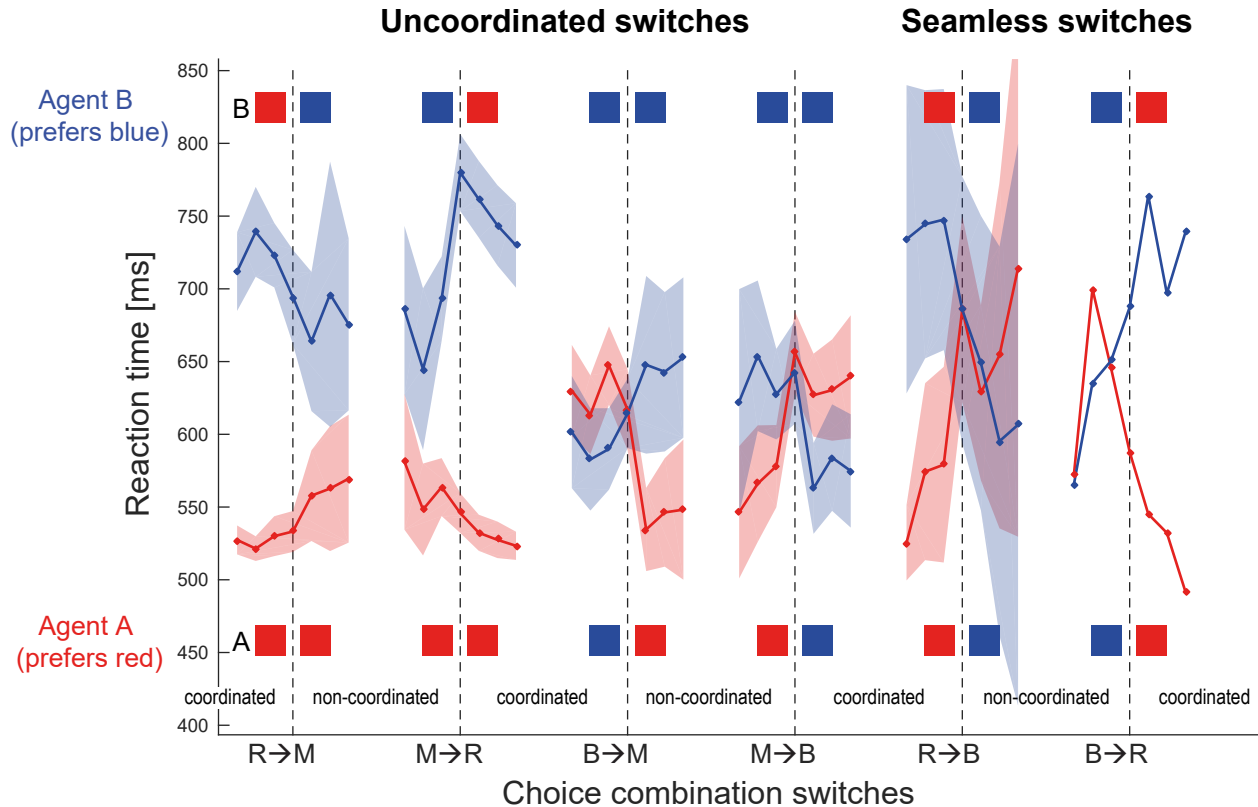
Supporting Figure S7. Reaction time changes around choice combination switches. The graphs show the reaction time transitions around 6 specific choice switches, for agent A (monkey F) in red and for agent B (monkey C) in blue. Each curve shows the mean reaction time for 3 trials before, at the actual switch trial and 3 trials after the switch, the shaded areas show 95% confidence intervals around the mean. R – red, B – blue, M – magenta. The choice transitions are R→M: from both agents selecting red to each selecting own color; B→M: from both selecting blue to each selecting own color; R→B and B→R: from both selecting one color to both selecting another color (seamless transition, rare in macaques, see Figure 8C); as well as the respective reverse transitions. The choices of agent B are indicated on the top, the choices of agent A on the bottom. Note that both macaques speed up when transitioning from coordinating on their non-preferred color to challenging (R→M for agent B, and B→M for agent A), and also when transitioning from challenging to coordinating on their preferred color (M→B for agent B, and M→R for agent A). Similarly, both macaques slow down when transitioning from coordination on their preferred color to challenging (B→M for agent B, and R→M for agent A) and when transitioning from challenging to coordination on their non-preferred color (M→R for agent B, and M→B for agent A). In rare seamless switches, monkeys speed up on coordination on the preferred color and slow down for transitions to the non-preferred color.

Supporting tables

Pair ID	N solo trials	Solo trials: share of own choices A [%]		Solo trials: share of own choices B [%]		Choice preference (own vs. other's) in last 200 dyadic trials	
		all trials	last 25	all trials	last 25	A	B
1	100	43	48	54	60	balanced	balanced
2	100	93	**100**	96	**100**	balanced	balanced
3	100	86	**100**	55	52	balanced	balanced
4	100	92	**100**	99	**100**	own	own
5	100	98	**100**	95	**100**	own	balanced
6	100	92	**100**	87	**92**	balanced	balanced
7	100	26	4	50	48	balanced	balanced
8	100	54	52	99	**100**	balanced	balanced
9	100	97	**100**	93	**100**	balanced	balanced
10	50	86	**96**	24	16	own	other's
11	50	46	48	2	0	balanced	other's
12	50	34	64	61	64	balanced	balanced
13	50	89	**92**	60	44	balanced	balanced
14	50	94	**100**	50	40	own	other's
15	100	1	0	53	40	other's	balanced
16	100	58	**80**	20	0	own	other's
17	100	94	**100**	94	**100**	balanced	balanced
18	100	96	**96**	31	24	other's	own
19	100	93	**100**	1	0	own	own

Table S1. No relationship between own target preference in solo trials and dyadic behavior in human pairs. Twenty of 38 human subjects preferred their own target at the end of the solo training trials (last 25 trials own color selection >75%, marked by bold font). Percentage of choice for the own preferred target for both agents of each human pair, for all solo trials and for the last 25 solo trials. Pairs displaying dynamic coordination (turn-taking behavior) are designated by a gray background. Choice preference in dyadic trials: other's (SOC<25%), balanced (25%≤SOC≤75%), own (SOC>75%). Out of 10 pairs (20 subjects) that showed balanced turn-taking, 11 subjects showed a strong preference for own targets in solo trials, and 9 did not. In the remaining 9 pairs (18 subjects), 7 subjects who had own preference in solo trials also showed it in dyadic trials, 4 subject who had other's preference in solo trials also showed it in dyadic trials, and 5 out remaining 7 subjects switched the preference from solo to dyadic trials, e.g. other's to own. Thus, the behavior in the solo trials is not predicting the behavior in the dyadic trials. This indicates that human participants did not simply carry their established preferences into the dyadic setting, but continued to explore and modify their strategies.

Monkey ID	N trials	Solo trials: share of own choices [%]	
		all trials	last 25 trials
C	360	100	96
E	270	99	100
F	70	97	100
L	370	97	100
M	700	78	76
T	250	100	100

Table S2. At the end of solo training all macaques preferred their own targets (SOC>75%). Percentage of choice for the own preferred target for all 6 macaques, for all trials of the session and for the last 25 trials. This was calculated for the last solo session before being paired with a partner.

Pair ID	RT A – RT B, coordinated trials			RT A – RT B, non-coordinated trials			Test on coordinated vs non-coordinated RT differences		
	mean	SD	N trials	mean	SD	N trials	df(corrected)	t	p
1	-110.4	90.8	306	-89.7	76.2	94	181.2512	-2.1947	**0.029453**
2	72.4	85.1	383	64.6	63.7	17	18.6252	0.48486	0.63343
3	130.5	116.1	263	60.2	114.6	90	155.9884	5.007	**0.0000014791**
4	n/a	n/a	0	-21.3	52.3	400	n/a	n/a	n/a
5	-7.0	57.4	121	-30.5	52.3	279	210.3226	3.8661	**0.0001474**
6	-7.3	77.9	396	104.3	70.0	4	3.0754	-3.1696	**0.0488**
7	-27.2	72.4	367	-32.8	96.3	33	35.3318	0.32657	0.74592
8	-177.6	80.3	377	-56.4	114.6	23	23.3354	-4.9973	**0.000045152**
9	5.9	82.9	320	-5.1	77.3	80	128.3393	1.1281	0.26137
10	-36.4	73.7	300	156.2	n/a	1	n/a	n/a	n/a
11	-43.2	66.4	176	-33.2	55.5	124	289.3612	-1.4156	0.15798
12	-22.5	125.9	276	-23.2	97.0	24	30.1775	0.028955	0.97709
13	106.2	109.2	239	75.1	84.6	61	116.5269	2.405	**0.017747**
14	-40.3	53.8	240	-14.2	61.3	60	83.0936	-3.0278	**0.0032803**
15	55.5	70.2	166	15.3	67.0	235	345.0293	5.7627	**0.0000000184**
16	0.2	66.5	388	-31.0	142.8	12	11.1479	0.75305	0.46703
17	114.5	85.4	337	153.5	87.1	63	85.7539	-3.274	**0.0015301**
18	-96.0	108.8	398	9.5	65.6	2	1.0279	-2.2605	0.25973
19	n/a	n/a	0	22.2	90.8	407	n/a	n/a	n/a

Table S3. Nine of 19 human pairs showed significant differences between coordinated and non-coordinated trials' reaction time difference, indicating that seeing the other's action can help maintain coordination. Comparison of trial-by-trial reaction time differences between the two agents in the pair (RT A – RT B), for coordinated and non-coordinated trials for all human pairs by t-test using Satterthwaite's approximation to allow for unequal variance. Note that pairs 4, 10, and 19 had too few trials for one of the trial types to calculate the test statistic (n/a). Pairs displaying turn-taking behavior are designated by a gray background. Bold font denotes correlations with a p<0.05. Green font shows pairs where absolute RT differences increased with coordination, purple – where absolute RT differences decreased with coordination.

Session	df	r(A)	r(B)	p(A)	p(B)	raw r(A)	raw r(B)	raw p(A)	raw p(B)
FC-1	304	-0.544	0.169	**5.12E-25**	**0.002952**	-0.068	0.043	0.23521	0.45327
FC-2	365	0.236	-0.536	**4.62E-06**	**1.17E-28**	0.388	-0.114	**1.27E-14**	**0.028859**
FC-3	183	-0.351	0.200	**9.42E-07**	**0.0062399**	-0.124	-0.062	0.092404	0.40478
FC-4	207	-0.376	-0.200	**2.00E-08**	**0.0037778**	-0.128	0.218	0.064765	**0.001527**
FC-5	171	0.194	0.161	**0.010527**	**0.034044**	0.111	-0.0136	0.14464	0.8586
FC-6	397	n/a	-0.096	n/a	0.054399	n/a	-0.039	n/a	0.44066
FC-7	314	-0.044	0.560	0.433	**1.68E-27**	0.024	0.144	0.67187	**0.010417**
FC-8	447	0.447	0.228	**1.67E-23**	**9.91E-07**	0.171	0.027	**0.00027874**	0.5692

Table S4. Prior to confederate training (8 naïve sessions), macaques F (agent A) and C (agent B) only showed relatively weak and inconsistent (both positive and negative) correlations between probability to see other's action (r(A) and r(B)) and the selection of the other's target. List of correlation degree of freedom, correlation coefficients and correlation p-values for the data as displayed in Figure S6 (running average over 8 trials) and for the raw non-smoothed data. Bold font denotes correlations with p<0.05. n/a – not applicable: the correlation is undefined because at least one variable was constant (e.g. the share of other's choice 0). The few high correlations on smoothed data were mainly driven by the slow changes in the timecourse of the session.

Session	df	r(B)	p(B)	raw r(B)	raw p(B)
F-1	431	0.063	0.19125	0.079	0.10121
F-2	714	0.125	**0.00080364**	0.108	**0.0037979**
F-3	578	-0.053	0.20033	0.140	**0.00069979**
F-4	80	0.040	0.72184	0.248	**0.024749**
F-5	547	0.044	0.30683	0.173	**4.56E-05**
F-6	317	0.280	**3.54E-07**	0.205	**0.00022491**
F-7	540	0.141	**0.0010228**	0.170	**6.82E-05**
F-8	582	0.077	0.063123	0.076	0.068242
F-9	547	0.260	**5.64E-10**	0.298	**9.42E-13**
F-10	479	0.544	**2.19E-38**	0.390	**6.99E-19**
F-11	473	0.820	**7.89E-117**	0.586	**3.44E-45**
F-12	525	0.829	**1.15E-134**	0.579	**1.40E-48**
F-13	594	0.815	**3.75E-143**	0.589	**7.59E-57**
F-14	554	0.867	**2.42E-169**	0.557	**1.10E-46**
F-15	446	0.868	**1.40E-137**	0.642	**2.38E-53**
C-1	652	0.161	**3.51E-05**	0.188	**1.37E-06**
C-2	544	0.311	**9.50E-14**	0.172	**5.18E-05**
C-3	320	0.482	**3.56E-20**	0.307	**1.78E-08**
C-4	455	0.145	**0.0018963**	0.188	**5.03E-05**
C-5	412	0.588	**7.67E-40**	0.285	**3.63E-09**
C-6	476	0.697	**7.88E-71**	0.408	**1.46E-20**
C-7	581	0.670	**3.11E-77**	0.388	**2.42E-22**
C-8	228	0.715	**2.36E-37**	0.471	**4.42E-14**

Table S5. During confederate training both macaques (F, C) developed strong and highly significant positive correlations between being able to see the confederate's actions and following the confederate. Correlation between the probability to see other's action and the selection of the other's target: 15 sessions of monkey F and 8 sessions of monkey C during confederate training. List of correlation degree of freedom, correlation coefficients and correlation p-values for the data as displayed in Figure S6 (running average over 8 trials) and for the raw non-smoothed data. Note the human confederate worked on position A, and is excluded here. Bold font denotes correlations with a $p<0.05$.

Session	df	r(A)	r(B)	p(A)	p(B)	raw r(A)	raw r(B)	raw p(A)	raw p(B)
FC-1	371	0.806	0.834	**2.45E-86**	**1.24E-97**	0.516	0.603	**8.13E-27**	**3.13E-38**
FC-2	545	0.916	0.909	**3.57E-218**	**1.28E-209**	0.635	0.655	**4.15E-63**	**2.93E-68**
FC-3	625	0.753	0.518	**6.88E-116**	**2.65E-44**	0.423	0.321	**1.19E-28**	**1.76E-16**
FC-4	471	0.856	0.517	**9.73E-137**	**9.49E-34**	0.420	0.305	**1.20E-21**	**1.22E-11**
FC-5	683	0.707	0.654	**1.16E-104**	**9.32E-85**	0.562	0.430	**2.71E-58**	**3.53E-32**
FC-6	370	0.901	0.822	**3.33E-136**	**1.29E-92**	0.601	0.580	**7.55E-38**	**7.34E-35**

Table S6. After confederate training macaques F (agent A) and C (agent B) continued to show strong correlations between seeing the other's choice and following. Correlation between the probability to see other's action and the selection of the other's target: 6 sessions of monkeys F and C after confederate training. List of correlation degree of freedom, correlation coefficients and correlation p-values for the data as displayed in Figure 7 and Figure S6 (running average over 8 trials) and for the raw non-smoothed data. Bold font denotes correlations with a $p<0.05$. Note that even in the sessions without significant DCR in the last 200 trials (i.e. FC-1, FC-3, FC-4, FC-6), where monkeys largely converged on one color, the high positive correlations reflect brief bouts of competitive turn-taking or challenging.

Pair ID	df	r(A)	r(B)	p(A)	p(B)	raw r(A)	raw r(B)	raw p(A)	raw p(B)
4	398	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
19	405	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
15	399	-0.123	-0.110	**0.014003**	**0.026945**	0.067	0.176	0.1802	**0.00039892**
11	298	-0.097	n/a	0.092301	n/a	0.040	n/a	0.48759	n/a
5	398	-0.009	-0.427	0.8564	**3.9344E-19**	-0.028	-0.171	0.58233	**0.0005981**
3	351	0.164	0.006	**0.0019414**	0.91763	-0.012	0.054	0.82342	0.30902
1	398	-0.009	0.025	0.85616	0.62476	0.027	0.026	0.59697	0.60243
14	298	0.060	0.079	0.29914	0.17498	0.010	0.154	0.86265	**0.0073713**
13	298	0.126	0.108	**0.029495**	0.061976	0.026	-0.037	0.65319	0.52015
2	398	0.243	0.394	**9.1268E-07**	**2.5461E-16**	0.095	0.124	0.057691	**0.012791**
12	298	0.523	0.456	**1.9807E-22**	**8.0894E-17**	0.268	0.330	**2.41E-06**	**4.90E-09**
16	398	0.046	-0.569	0.36308	**1.1837E-35**	0.142	-0.214	**0.0044703**	**1.64E-05**
7	398	0.053	0.287	0.29108	**4.9403E-09**	0.001	0.066	0.97791	0.18541
9	398	0.126	-0.271	**0.011469**	**3.5903E-08**	0.009	-0.025	0.85404	0.61258
17	398	-0.487	0.187	**3.1051E-25**	**0.00017068**	-0.054	0.038	0.28453	0.44647
8	398	-0.114	0.075	**0.022805**	0.13386	-0.036	0.055	0.47255	0.2706
6	398	-0.370	0.071	**2.0434E-14**	0.15713	0.040	0.017	0.41953	0.73857
10	299	0.251	n/a	**1.0701E-05**	n/a	0.221	n/a	**0.00010978**	n/a
18	398	0.048	0.176	0.3415	**0.00039929**	-0.000	0.094	0.99829	0.060414

Table S7. Human pairs only showed weak and inconsistent (positive and negative) correlations between being able to see the other's choice and following to the other's target. Correlation between the probability to see other's action and the selection of the other's target in 19 human pairs. List of correlation degree of freedom, correlation coefficients and correlation p-values for the data as displayed (running average over 8 trials) and for the raw non-smoothed data. (A) and (B) denote the positions of the agents. Pairs are sorted by the average joint reward in ascending order (cf. Figure 2). Pairs displaying turn-taking behavior are designated by a gray background. Bold font denotes correlations with a p<0.05. n/a – not applicable: the correlation is undefined because at least one variable was constant (e.g. the share of other's choice 0).

Written instructions read by human subjects

Description of the study

This study is the investigation of the behavioral correlates of social decision making while playing a game with a partner. Every day we have to make decisions that depend not only on our own needs and goals but also on the needs and goals of others: for instance, while working on a project with colleagues, planning vacation with friends and family, getting on the bus or shopping groceries, etc. With the help of the task presented in this study we will investigate how people make such decisions.

The course of the study

You will complete one session of a decision making task on the computer together with your partner. You will have to choose one of the two circles presented to you. Your partner will have to perform the same task on his/her side of the touchscreen. You will have to decide and respond quickly. If either you or your partner are too slow, the trial will be aborted without any reward. Your reward will depend on your own and your partner's choice. After the decision is made, you both will receive different auditory feedbacks, denoting your reward and the reward of your partner. Please do not talk to your partner during the session. After the session, we will ask you several questions about the experiment.

Instructions read to human subjects by experimenter

1) You rest both hands on the gray board at the two round objects (touch sensors).

2) A central touch target will appear.

3) Move your right hand to the target and hold. Use the right hand during the entire session.

4) While the target brightens up keep holding your finger on the target.

5) Then two colored choice targets appear and the central target disappears.

6) Make your choice and touch the chosen target within 1.5 seconds.

7) All touched targets will brighten up, keep holding until the targets disappear.

8) Please note that both selected targets will brighten up:

   In case both players selected the same target only that single target will brighten;

   In case both players selected different targets, both targets will brighten.

9) Now two streams of auditory beeps will signify the earned reward for each player: each beep corresponds to a few cents. Please try to learn the sound related to your reward during the training trials.

10) While the audio plays move the hand back to the two touch sensors.

11) Go to 1.