

# Old World monkeys are more similar to humans than New World monkeys when playing a coordination game

Sarah F. Brosnan<sup>1,2,3,\*</sup>, Bart J. Wilson<sup>4,†</sup> and Michael J. Beran<sup>1,2,†</sup>

<sup>1</sup>Language Research Center, <sup>2</sup>Department of Psychology, and <sup>3</sup>Neuroscience Institute, Georgia State University, PO Box 5010, Atlanta, GA 30302, USA

<sup>4</sup>Economic Science Institute, Chapman University, Orange, CA 92866, USA

There is much debate about how humans' decision-making compares with that of other primates. One way to explore this is to compare species' performance using identical methodologies in games with strategic interactions. We presented a computerized Assurance Game, which was either functionally simultaneous or sequential, to investigate how humans, rhesus monkeys and capuchin monkeys used information in decision-making. All species coordinated via sequential play on the payoff-dominant Nash equilibrium, indicating that information about the partner's choice improved decisions. Furthermore, some humans and rhesus monkeys found the payoff-dominant Nash equilibrium in the simultaneous game, even when it was the first condition presented. Thus, Old World primates solved the task without any external cues to their partner's choice. Finally, when not explicitly prohibited, humans spontaneously used language to coordinate on the payoff-dominant Nash equilibrium, indicating an alternative mechanism for converting a simultaneous move game into a sequential move game. This phylogenetic distribution implies that no single mechanism drives coordination decisions across the primates, while humans' ability to spontaneously use language to change the structure of the game emphasizes that multiple mechanisms may be used even within the same species. These results provide insight into the evolution of decision-making strategies across the primates.

**Keywords:** coordination; comparative; decision-making; Assurance Game; primates; monkeys

## 1. INTRODUCTION

Along with the dispositions to trust and reciprocate, and the propensity to exchange, the human ability to coordinate activities is a pillar upon which the flourishing of the species is built. The ability of two individuals to coordinate (literally to mutually arrange) an activity presupposes, first, that two individuals cognize that the outcomes of their actions are interdependent. Second, successful coordination assumes a shared attention and agreement on the ends to be achieved by mutually arranging a pair's activities. Within the Pleistocene tribe or the modern small group of family, friends and neighbours, these conditions are almost trivially met as personally known individuals share the habits, knowledge and beliefs about the methods and possibilities necessary to coordinate successfully. But what happens when modern strangers face a novel task of playing a simple  $2 \times 2$  normal-form game (NFG) of coordination? How well can the personally unknown extend to each other the assumptions of interdependent decisions and agreement on the ends? As van Huyck *et al.* [1] suggested, not as robustly as we might hubristically expect. The question then becomes more interesting when posed of our primate relatives. Do we share the ability to cognize

actions as interdependent and to share attention on the ends achieved?

There is little research addressing the question, but what there is indicates that several primates may share these abilities with humans. For instance, in one Ultimatum Game study with chimpanzees [2], the apes' decisions were dissimilar from human-typical behaviour [3]. However, a subsequent study found that the protocol designed for chimpanzees led to a similar outcome in humans [4]. Our previous work investigated the Assurance Game, which is a well-known model of social interactions [5]. We found that while humans found payoff-dominant outcomes more readily than did either capuchin monkeys or chimpanzees, pairs of all species were able to find these outcomes, indicating that selection has favoured similar outcomes (if not similar cognitive mechanisms) across these three primates [6]. Moreover, we found that chimpanzees with greater experience in cognitive testing found the payoff-dominant outcome far more readily than did chimpanzees with little or no cognitive testing experience, indicating that, as with humans, experience may play a role in outcomes.

For the current study, we chose to investigate the Assurance Game using a computerized methodology. This provided a number of advantages for further research. First, there is a long history in comparative research suggesting that the format in which one presents the same kind of task to non-human animals can have radical effects on performance. For example, spatial discontinuity between response loci and stimuli was

\*Author for correspondence (sbrosnan@gsu.edu).

†All three authors contributed equally to the study.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2011.1781> or via <http://rspb.royalsocietypublishing.org>.

recognized as an obstacle to learning in animals [7–9]. However, training primates to use a joystick gets past that problem of discontinuity and produces markedly different patterns of results [10,11], and the same may be true for performance in economic games. Thus, comparing our previous results to those from a computer task may help to highlight factors that affect decision-making.

Another advantage of a computerized task is the abundance of data with respect to other cognitive abilities that are relevant. This allows us to not only consider *a priori* whether species might be able to solve the task, but if our predictions are proved false to reconsider how the subjects might perceive the game. Considering cognitive mechanisms, success in the game seems to require, at minimum, an ability to respond flexibly (e.g. contingent upon one's partner's decisions) and, related to this, an ability to inhibit (e.g. avoid the temptation of a short-term payoff). Considering first the role of phylogenetically widespread learning mechanisms, a number of primate species, including the rhesus monkeys and capuchin monkeys who participated in the current study, have shown substantial behavioural flexibility in responding to game-like tasks presented on computer screens. These include tasks that involve behavioural inhibition, tracking of relative rates of reward for different responses, and even information-seeking behaviour [12–15]. These skills, along with the monkeys' clear interest in maximizing their food intake during these tasks, indicate that comparative assessments of cooperation using computer tasks are likely to provide compelling data for understanding the emergence of cooperation in humans.

However, despite equivalent performance on basic learning tasks (e.g. two choice discrimination and learning set tasks), not all primates are equally adept at performing higher-level cognitive tasks. Relevant to this task, rhesus monkeys, but not capuchin monkeys, show evidence for metacognitive monitoring during psychophysical judgement tasks [16] and information-seeking paradigms [17]. This suggests a species difference in monitoring ongoing performance. Thus emerges a potentially important way to determine which cognitive mechanisms are important; it is possible that rhesus and capuchins would do equally well on games of coordination where contingencies for responses are clearly presented (that is, in a situation similar to a basic learning task), while diverging in performance when immediate cues are not present. Finally, it is possible, although we think unlikely, that even higher-order cognitive mechanisms are involved, such as theory of mind, which could be activated in this case owing to the social nature of the task. If this is the case, we expect humans to outperform the other species [18], despite some basic perspective abilities that have been seen in rhesus monkeys [19]. Thus, testing these species on the Assurance Game may shed light on not only performance levels, *per se*, on the game, but also the nature of the game itself from the perspective of the individuals playing it.

To explore these issues in more detail, we here investigated the role of information in coordination decisions among three primate species: humans (*Homo sapiens*), rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). We redesigned the typical NFG methodology [1,20–22] specifically to work across species, holding the methodology as constant as possible (see §2

for details). We had two hypotheses for the current work. First, based on our prior results with the exchange version of the task, we expected all species to be capable of successfully navigating the task, but we predicted that a higher percentage of human pairs would find the payoff-dominant outcome when compared with the monkey species. Our second hypothesis was that outcomes would change as the task parameters varied. Specifically, we predicted that if pairs could see each other's choices prior to making a decision, the task could be solved by a cognitively simpler matching rule, meaning that all species would perform equally well. On the other hand, we predicted that in the situation in which their partners' choices were not available, higher-order mechanisms might be required, separating the species based on their aptitude at the tests of cognitive abilities we proposed above as relevant to this task (that is, humans performing better than the other primates).

## 2. METHODS

### (a) *General methods*

#### (i) *The Assurance Game*

The game we used was a common game of coordination called the Assurance Game, sometimes referred to as the Stag Hunt Game. The reward structure was such that mutual Stag play was the most beneficial (four units each), mutual Hare play resulted in a low payoff (one unit each) and the uncoordinated payoff of playing Stag when one's partner plays Hare was unrewarded, while the individual who played Hare received one unit. This game has two pure strategy Nash equilibria: (Stag, Stag), which is the payoff-dominant equilibrium (the outcome that maximizes payoff to both individuals), and (Hare, Hare), the outcome which is payoff-dominated. This well-known coordination game is interesting to economists because strategical uncertainty plays a key role in the selection of the equilibrium, yet the players' objectives are aligned (for a summary, see [21]). In the Assurance Game, the objectives may be the same (Stag, Stag), but the question of strategical interest is how sure a given player is that the other player will play Stag when he or she plays Stag. Evidence from coordination game experiments with humans indicate that the payoff-dominant equilibrium is not a focal point with repeated interactions, as in van Huyck *et al.* [1], nor with anonymous play with different individuals, as in Cooper *et al.* [23].<sup>1</sup>

We explicitly incorporated a number of features common to non-human studies but different from traditional NFG experiments with humans to facilitate cross-species comparisons. First, subjects received no verbal instruction or pre-testing so that individuals had to discover the payoff structure during the course of the game (note that there were only two options from which to choose, and thus four possible outcomes). Second, all subjects, including humans, had participated in other experiments in the laboratory prior to this study, so they were aware that decisions would result in tangible rewards (e.g. food or cash). Third, subjects were paid on a trial-by-trial basis, in case the immediacy of receiving rewards on each trial affected behaviour. Fourth, for most treatments (see exception below), subjects sat directly next to one another and were not anonymous, so subjects could potentially communicate [24,25] (we saw no attempts to communicate between the primates). Fifth, neither humans nor rhesus monkeys received any pre-tests

designed to assess Assurance Game understanding, so all subjects, including humans, had to discover the payoff structure during the course of the game. Capuchins had previously participated in an exchange-based version of the task [6], but had no additional training. Finally, we manipulated whether they could see their partners' decisions to investigate how information affected decision-making.

(ii) *General computerized design*

Decisions were made by choosing one of two icons on each side of a split computer screen, one of which represented Stag (a red square) and one of which represented Hare (a blue circle). Icons were presented in a vertical distribution, with the order of presentation randomized both across trials and across individuals within the same trial. Subjects of all species made a choice using a joystick. We chose to study their behaviour as naturally as possible, and so did not constrain the order of play or the timing of decisions. Once both subjects had made a choice, each subject received (or not) rewards dependent on both what they chose and what their partner chose, following the payoff structure of the assurance experiment.

There were two conditions: *Synchronous*, in which subjects did not know what their partner had chosen until both choices were complete, and *Asynchronous*, in which decisions were revealed as they were made (e.g. subjects potentially had information about their partners' responses). To block any information transfer in the Synchronous condition, the joystick itself was occluded and the cursor did not move; when the joystick was manipulated, both options and the cursor disappeared simultaneously, and both subjects' choices were displayed simultaneously once both decisions had been made. Thus it was, to our knowledge, impossible to determine the partner's behaviour by observation in this procedure, other than knowing that one's partner had made a choice. In the Asynchronous game, the procedure was identical, except that choices were displayed as they were made, so that their partner could see their choice and potentially use the information when making their decision.

(iii) *Non-human primates*

All non-human primates were socially housed at the Language Research Center of Georgia State University. Rhesus monkeys were all adult males who were moved to a specially designed paired testing area. Capuchin monkeys were socially housed in multi-male, multi-female social arrangements and voluntarily separated into an adjacent cage for testing, to limit distractions. Only adults were tested, always with members of their social group, and in multiple pairings from within the same social group whenever possible. No individual was ever food- or water-deprived for testing. During test sessions, pairs did not always finish a trial block, and pairs completed different numbers of blocks during each testing session. Thus, the number of trials varied across both pairs and sessions.

The capuchins and the rhesus monkeys were used to somewhat different testing schedules, so to avoid changing their schedules and causing unnecessary stress, we initially proceeded using their typical schedules. Rhesus monkeys were given 6 h testing sessions consisting of 60-trial blocks with a 30 min interval between blocks. Pairs could complete as many trial blocks as they chose. Capuchins were initially given a single 40-trial session per day, as per their norm, but of four pairs, only one achieved the Stag-Stag outcome

in the Asynchronous version, and even this pair did not maintain it. We then implemented a more rhesus-like schedule, except with 2 h test sessions (they became agitated if left in their testing cages for any longer), at which point all reached the Stag-Stag outcome. For more detail, see the electronic supplementary material.

The non-human primates did require some training to learn using the split screen (a novel experimental feature) and to make choices within the same time frame. This occurred through a two-stage training process. First, two monkeys worked together to learn that they had to both hit a single target on their side of the screen before both would receive a food pellet. Then, they had to progress to a point where they would make those same responses within a 5 s window from the initial presentation of a trial. Note that these training stimuli were not those used in the Assurance Game, and there was no choice behaviour on the part of the primates. There was only one icon on the computer screen, which could be contacted with the cursor. This training was used to teach them which half of the screen presented their choices (and outcomes), and to teach them that they needed to respond relatively quickly when a trial was presented. This training assured that, at minimum, subjects knew that (i) rewards were not given without both individuals making a response and (ii) they could control only their cursor. Although, of course, we cannot know how the monkeys actually interpreted the task, we do know they at least understood how to generate responses that might bring rewards, and what limitations had been imposed.

*Capuchin monkeys.* Based on our previous research [6], we knew that the capuchins would have more difficulty with the task than did the humans. Thus, we started all capuchins with the Asynchronous version of the task and, once they had learned it, gave them the Synchronous version. The subjects had far more difficulty with this, so to verify that they were still able to do the basic task, we repeated the Asynchronous version. Finally, we repeated the Synchronous version to see if the extensive experience had increased their skill level.

*Rhesus monkeys.* Half of the rhesus monkeys were started on the Synchronous and half with the Asynchronous version. All of the subjects on the Synchronous version succeeded, and so we did not return them to the Asynchronous version (see §3 and table 2). Subjects that started on the Asynchronous version were subsequently run on the Synchronous version.

(iv) *Humans*

Undergraduate subjects were recruited from the general student body at Chapman University, Orange, CA, USA. Subjects were randomly recruited via an email system and paid \$7 for showing up on time, plus what they earned in the experiment. Each subject had participated in at least one economic experiment sometime prior to this session (participating in a previous study involving NFG experiments or the Assurance Game disqualified individuals as a participant) so that they had experience with receiving actual payment for their decisions in this laboratory. No subject participated in more than one pairing or more than one version of the task.

The humans' only instruction on the Synchronous treatment was limited to the following six points:

- Have you participated in an economic experiment before? (Both must reply with a 'yes' to participate.)

- In this experiment, you will be making decisions using a joystick attached to a computer. Use the *left* thumb pad to make a decision.
- As the experiment progresses, you may be paid in quarters by the machines next to your computer.
- Please collect the coins in the yellow cups provided so as to not clog up the machines.
- These are the only instructions you will receive in the experiment. Once the experiment begins, the experimenter will not be allowed to answer *any* questions until the experiment is over.
- Do you have any questions before the experiment begins?

Subjects initially began with the Synchronous task, as described above. Pairs of participants, who were the only two individuals in the room, sat next to one another at a single computer and used a joystick. The lack of anonymity enhances the likelihood of achieving the Pareto-dominant outcome. Pairs received payment in quarters (from a coin dispenser, an analogue to the primates' pellet dispenser) and payoffs were in the same ratios as those of the monkeys (accumulated coins were converted into large bills at the conclusion of the experiment). However, the results of this game indicated that language was an important characteristic in determining the pairs' outcomes (see §3 for details), thus a true comparison between the conditions could not be done. While we could have simply asked participants not to talk with each other during the game, this differed from the other primates, who could communicate to the fullest extent of their abilities, and may have led to an awkward social environment. Thus, we instead investigated the Synchronous/Asynchronous comparison using an alternative procedure.

This procedure was based on a typical NFG procedure with a  $2 \times 2$  matrix of payoffs and strategies. For the NFG treatments, the game and payoffs were the same; however, participants were isolated at individual study carrels playing on their own computer against an anonymous opponent drawn from among the other participants in the room. To hide who was partnered with whom, rewards accumulated and were paid out at the end of the session, rather than using a coin dispenser after each decision (the noise from coin dispensers would have served as a cue). As with the monkeys, in the NFG Synchronous version, both partners' choices were displayed simultaneously to both players after both decisions had been made, while in the NFG Asynchronous version choices were displayed to both players as they were made. As an additional benefit, having results in a traditional NFG procedure allowed us to see how our Synchronous results with humans compared with typical NFG experimental procedures (e.g. involving instruction).

#### (v) *Synchronous version*

Fifty-two undergraduate subjects participated in the study in pairs (i.e. in 26 separate sessions).

#### (vi) *Normal-form game synchronous and normal-form game asynchronous versions*

One hundred and eighteen undergraduate subjects were recruited by the same protocols above, except that 12–24 people participated at the same time. Fifty-eight people in three sessions of 22, 24 and 12 participated in the NFG Synchronous treatment and 60 people in three sessions of 24, 24 and 12 participated in the NFG Asynchronous treatment. The subjects were simultaneously seated in a computer laboratory

at visually isolated carrels and instructed not to talk to one another. They then read self-paced instructions on how to participate in the experiment. These subjects were privately paid their total accumulated earnings at the conclusion of the experiment; they did not receive payment as they made each decision to avoid possible cuing to one's partner's identity.

#### (vii) *Statistics*

Statistics are non-parametric owing to small sample sizes. Primate results are based on individual analyses, while for humans we include both individual analyses and inferential statistics that allow generalizations about the population. This difference in approach is because fewer monkeys were available for the study, as few are sufficiently well trained for computerized testing. All statistics are two-tailed.

Note that in many cases both  $\chi^2$  and Fisher's exact tests were impossible owing to the large number of cells with zero values. Thus, to determine whether a pair showed a pattern in their decision-making, we considered it meaningful if the subject showed an 80 per cent or greater preference for one of the four options (for the pair; chance was actually 25% in this case) or one of the two options (for the individual; chance is 50%). This percentage is significant for a binomial test for 20 trials and, as all of our subjects had at least 40 trials (and for pairs, chance was 25%, not 50%), this represented a conservative estimate for what constituted a significant pattern to their decision-making.

### 3. RESULTS

#### (a) *Non-human primate results*

Both monkey species did very well in the Asynchronous task. All capuchins began with the Asynchronous version owing to previous results indicating that they would have trouble finding the payoff-dominant outcome [6]. One pair reached the payoff-dominant outcome in our initial 40-trial sessions and the other three did so when switched to 60-trial sessions (see §3). Three of the rhesus monkey pairs (composed of four unique individuals) first played the Asynchronous game. Two reached the 80 per cent Stag-Stag criterion within a single session (table 2) and the third pair did so in the second session.

Despite this similarity, the monkeys differed in their outcomes in the Synchronous task. None of the capuchin pairings showed any preference for playing Stag-Stag (or any other outcome) when tested in a novel pairing on the Synchronous task. To see if this was due to a lack of understanding of the task, we reran the Asynchronous version with the capuchins. Eight monkeys were paired in multiple pairings (with a range of 1–3 partners per individual; table 1) for a single session consisting of as many trials as they chose to complete in 2 h (mean = 253 trials). All but one pair chose Stag-Stag at least 80 per cent of the time (range: 80.6–96.1%; table 1) and the exceptional pair chose Stag-Stag 71 per cent of the time. Nonetheless, when retested on the Synchronous task, outcomes remained poor. Stag-Stag was maintained in only one of the five pairs that had been tested together previously (table 1). Intriguingly, the exceptional pair showed the lowest frequency of Stag-Stag outcomes in the Asynchronous test, and the pair that had not previously been tested together also more often played the Stag-Stag outcome (table 1). Note that capuchins' poor performance occurred despite both previous experience

Table 1. Comparing the capuchin pairs' performance on the second Asynchronous and Synchronous games (note some pairs had previous experience on both the Asynchronous and Synchronous games; see §3 for details).

pairings	Asynchronous game (second)		Synchronous game (second)					
	% Stag-Stag	no. trials (1 session)	overall		first session		second session	
			% Stag-Stag	no. sessions/trials	% Stag-Stag	no. trials (1 session)	% Stag-Stag	no. trials (1 session)
Drella + Wren	80.6	201						
Griffin + Drella	93.0	196						
Griffin + Wren	96.1	360	33.1	8/3042	34	300	24.2	355
Lily + Wren	85.2	240	38.9	7/2176	81.1	180	43.8	420
Griffin + Lily	81.7	68	28.4	6/2864	36.1	540	36.7	414
Liam + Logan	71.3	240	78.1	6/1467	88.2	170	93.1	29
Logan + Gabe	89.2	240						
Nala + Logan	93.5	420	34.6	5/2260	74.2	360	31.4	420
Nala + Liam			60.1	4/1437	78.3	300	64.0	417

Table 2. Comparing rhesus performance by session.

game		overall		% Stag-Stag				
		% Stag-Stag	no. sessions/trials	session 1	session 2	session 3	session 4	session 5
Obi + Han	Asynchronous	74.6	5/1483	53.7	77.9	84.6	76.7	81.0
Luke + Obi <sup>a</sup>	Asynchronous	94.5	5/1876	90.9	95.3	96.1	93.9	96.4
	Synchronous	65.7	2/824	76.7	54.8			
Chewey + Han <sup>a</sup>	Asynchronous	87.8	5/1440	85	94.2	85.3	87.2	87.3
	Synchronous	72.2	3/360	66.7	68.8	81.3		
Hank + Gale	Synchronous	84.0	2/720	74.5	93.4			
Willie + Murph	Synchronous	88.1	3/472	31.6 <sup>b</sup>	81.3	95.7		

<sup>a</sup>Both pairs that played both games played Asynchronous first, followed by Synchronous.

<sup>b</sup>This represented only a single trial block; in actuality Willie and Murph reached 90% Stag-Stag preference in their third trial block, which was in their second session. Hank and Gale did not reach 90% until their fifth trial block, which was in their first session.

with the Assurance Game and introductory experience in the presumably easier Asynchronous version of the game.

On the other hand, the two pairs of rhesus monkeys that were tested first on the Synchronous version quickly found the Stag-Stag outcome. Although it took them slightly longer to reach the 80 per cent criterion than the pairs that played the Asynchronous version first, both pairs did so within four trial blocks, and at the same frequency as those pairs that first played the Asynchronous version (table 2). Note that they accomplished this despite never having had the opportunity to match a partner's play after having seen it, ruling out this simple associative mechanism.

Finally, the two rhesus pairs that showed the highest frequency of Stag-Stag choices in the Asynchronous game were subsequently given the opportunity to play the Synchronous version. One pair maintained a preference for playing Stag-Stag, while the other started at a lower level of Stag-Stag choices and subsequently declined further. This seemed to be primarily owing to one individual who began preferentially choosing Hare. Thus at least one pair was able to maintain the Stag-Stag outcome when switched to the Synchronous task.

### (b) Human results

Among 27 human pairs in the Synchronous condition, 22 ultimately settled on Stag-Stag and five settled on

Hare-Hare. What was notable was the perfect correlation between outcomes and pairs' discussions. Although all pairs spoke to each other, textual analysis of video recordings revealed that not every pair spoke about the game. Among the five pairs who settled on Hare-Hare (range: 34–36 Hare-Hare choices in 40 trials; table 3), not a single pair spoke about the game (henceforth, non-communicators), while among the 22 pairs who spoke about the game (henceforth, communicators), every pair ultimately settled on Stag-Stag, choosing it in at least seven of the last 10 choices (13 pairs did so on every one of these choices). One-third of communicators chose Stag-Stag a minimum of 80 per cent of the time overall (range: 32–39 Stag-Stag choices) and more than half (55%) did so at least 70 per cent of the time. No non-communicators ever played Stag-Stag, indicating that they did not explore the decision space as thoroughly as did the other pairs (or the monkeys).

Given the larger human sample size, we can consider variation using inferential statistics. The overall payoff between communicators and non-communicators differed by a factor of 3 (independent samples Mann–Whitney *U*-test,  $p < 0.001$ , mean  $\pm$  s.e. payoff per trial for pairs of communicators:  $\$1.48 \pm 0.08$ ; non-communicators:  $\$0.47 \pm 0.002$ ). Moreover, among communicators, there was an increase in payoff between the first quartile and last quartile (Wilcoxon signed-ranks test,  $p < 0.001$ , quartile 1 mean  $\pm$  s.e.:  $\$0.98 \pm 0.11$ ; quartile 4:  $\$1.89 \pm 0.03$ ). On the other

Table 3. Human preferences across the three games. Earnings measure the mean earnings per trial for the pair, and so range from 0.25 (one individual earned a quarter from playing Hare and their partner earned nothing from playing Stag) to 2.0 (both earned a dollar from playing Stag). An average of 0.5 means that both played Hare on every trial. A mean of 1.825 indicates a single Hare play and a mean of 1.85 indicates a single Hare-Hare trial.

	talked	spoke about symbols	number of pairs	number of trials	ever played Stag-Stag	mean no. Stag-Stag	mean no. Hare-Hare	mean Q1 earnings	mean Q4 earnings
paired	yes	yes	22	40	all	27.7 (69%)	6.2	0.98	1.89
	yes	no	5	40	none	0 (0%)	35.2	0.44	0.5
Synchronous	no		30	40	25 (83%)	20.27 (51%)	10.83	0.93	1.43
Asynchronous	no		29	40	19 (66%)	13.48 (34%)	16.28	0.77	1.07

hand, among non-communicators, the payoffs increased across these quartiles only because the decrease in number of Stag choices stabilized payoffs at one quarter for each individual by the fourth quartile (Wilcoxon signed-ranks test,  $p = 0.039$ , quartile 1 mean  $\pm$  s.e.:  $\$0.44 \pm 0.01$ ; quartile 2:  $\$0.50 \pm 0.00$ ). Thus, sociality is not synonymous with coordination; humans had to actually discuss the game in order to benefit from language. Moreover, humans were not intrinsically better than the other primates at the Synchronous condition. They appeared to use communication to turn the Synchronous game into an asynchronous game, giving humans an additional mechanism for coordination.

Based on these results, humans played the NFG Synchronous and NFG Asynchronous versions using a more traditional NFG set-up that prohibited discussion (see §2). One-third (10 of 30) of the pairs in the NFG Asynchronous game chose Stag-Stag at least 80 per cent of the time. Intriguingly, an additional four of these pairs chose Hare-Hare 80 per cent or more (and one other did so 78% of the time), an outcome that we never saw in either of the monkey species. This could indicate that the non-human primates are more likely to explore the decision space than are humans, or that humans are more likely to persevere on responses that are rewarding. In the NFG Synchronous game, only four (14%) pairs chose Stag-Stag this often, while 10 (35%) pairs chose Hare-Hare. Thus, the frequency of Stag-Stag and Hare-Hare choices flipped between the two conditions.

Considering the data quantitatively, the overall payoffs do not differ between the games (Mann-Whitney  $U$ -test,  $z = 1.47$ ,  $p = 0.14$ ), although there is a non-significant trend for the players in the NFG Asynchronous game to earn more by the fourth period than do players in the NFG Synchronous game (Mann-Whitney  $U$ -test,  $z = 1.85$ ,  $p = 0.0643$ ). Moreover, the number of Stag-Stag choices was greater in the NFG Asynchronous treatment than in the NFG Synchronous treatment (Mann-Whitney  $U$ -test,  $z = 1.9$ ,  $p = 0.0574$ ), and twice as many partnerships failed to ever play Stag-Stag in the NFG Synchronous when compared with the NFG Asynchronous treatment (10/29 versus 5/30). Considering the games individually, participants showed a trend towards playing Stag-Stag more often than Hare-Hare in the NFG Asynchronous game (Wilcoxon signed-ranks test,  $z = 1.86$ ,  $p = 0.0624$ ), but not the NFG Synchronous version (Wilcoxon signed-ranks test,  $z = 0.46$ ,  $p = 0.6818$ ). Subjects improved over the course of both treatments, earning more money in the fourth quartile than in the first (NFG Asynchronous: Wilcoxon signed-ranks test,  $z = 4.26$ ,  $p < 0.001$  NFG Synchronous:  $z = 3.42$ ,  $p = 0.0006$ ).

Nonetheless, as with the other two primate species, humans were better at finding the payoff-dominant outcome when an opportunity to coordinate was presented (e.g. sequential play).

Finally, comparing humans' results from the non-human primate format (which was the Synchronous treatment) to the NFG Synchronous format, the pairs' payoffs were higher when they could talk (mean payoffs for the Synchronous treatment:  $\$1.63$ ; NFG Synchronous treatment:  $\$1.07$ ; Mann-Whitney  $U$ -test,  $z = 2.62$ ,  $p = 0.0087$ ). On the other hand, by the end of the game, subjects did equally well when they could see each other's decision as it was made as when they could talk (comparing fourth quartile payoffs in the Synchronous and NFG Asynchronous treatments; Mann-Whitney  $U$ -test,  $z = 1.00$ ,  $p = 0.3179$ ; table 3). Thus, subjects were equally able to use language or sequential moves without language to coordinate on the Stag-Stag outcome and achieve the same level of payoffs across procedures.

### (c) Does finding Stag-Stag mark a change in behaviour?

The above results assess individuals' tendency to find the coordinated, Stag-Stag solution. That is, to what degree do they explore the problem space and encounter Stag-Stag as an option? A second way to consider these data is to determine the frequency with which they played Stag-Stag after having found it for the first time. Considering first the monkeys, there is no evidence that finding Stag-Stag was sufficient to alter their behaviour. Among capuchins, no pair ever had a single trial block (60 trials) in which they did not play Stag-Stag at least once (see electronic supplementary material, results for details), yet despite this not a single pair was able to successfully solve the Synchronous version of the task. Rhesus monkeys similarly played Stag-Stag in every trial block, yet it was rare for a pair to settle on Stag-Stag before the third trial block (see electronic supplementary material, results for details).

Humans showed a similar pattern. In the Asynchronous task, most players who found Stag-Stag did so quickly (within the first four trials), yet only three (12%) of these pairs played nothing but Stag-Stag after finding it. In the Synchronous task, no pair ever exclusively played Stag-Stag after finding it (for details of both conditions, see electronic supplementary material, results). Thus, we find clear indication that in neither of the two computerized conditions do pairs' payoffs change meaningfully after finding the first Stag-Stag outcome, indicating that these results cannot be explained by a lack of knowledge about the payoff matrix. Similarly, sitting next to one's partner did not affect the frequency

with which individuals stayed on the Stag-Stag payoff. Of the 22 pairs that played Stag-Stag at some point, only one played nothing else after finding that outcome. Four additional pairs played Stag-Stag all but one or two times, possibly indicating that they were either jointly exploring the parameter space or individually exploring potentially increased outcomes (see electronic supplementary material, results).

#### 4. DISCUSSION

Our results indicate that generalizing the decision-making outcomes of primates into one in which species are 'better' or 'worse' than others is not sufficiently nuanced. Instead, we find that the results vary in interesting ways depending on context. Considering our second hypothesis first, all species were able to more easily solve the task in the Asynchronous than in the Synchronous conditions. Although these studies were explicitly designed to investigate outcomes, not mechanism, we are able to use these results to posit necessary mechanisms. Thus, it seems probable that this boost in performance was due to the availability of simple rules, such as matching-to-sample [26], which could not be used in the Synchronous task since the partner's behaviour was hidden. This indicates that it is possible to solve this task with a suite of fairly simple behavioural mechanisms. We also note that, without other data, this would indicate phylogenetic continuity among the primates, and thus, when considered with the Synchronous results, emphasizes the utility of using multiple methods to assess behaviour and cognition.

On the other hand, not all species were able to solve the Synchronous task. No pair of capuchin monkeys ever solved this task, even after experience (successfully) solving the Asynchronous task. Thus, the capuchins' inability in the Synchronous pairing was not due to a failure to understand the task outcomes. Instead, we think that the best explanation for the decline in performance in the Synchronous task is probably related to the fact that they could no longer see their partners' decisions and use a simple strategy such as matching. These results also imply that the one pair's success in the previous, exchange version [6] was probably due to the fact that they could see their partners' choices. On the other hand, some rhesus monkeys and humans found the payoff-dominant outcome in the Asynchronous task, and required very few trials in order to do so. Thus, Old World primates outperformed New World primates, rather than humans outperforming non-humans.

This has several very important implications when considering the mechanisms required to solve the Assurance Game. First, of course, the task can be solved using fairly simple learning-based mechanisms when appropriate cues are available (e.g. the Asynchronous task). Second, the high performance of both rhesus and humans implies that the presence of theory of mind is not required. Although one can see how the ability to predict one's partner's behaviour would be useful, it seems the task can be solved without it. It will be interesting to see whether this remains true in more complex games that lack a mutually beneficial payoff-dominant outcome. Finally, in the absence of cues, it seems likely that other more complex cognitive abilities are required to solve the task. We note that these results are

particularly strong given our choice of a New World monkey species. Capuchins have an unusually large (for a monkey) brain-to-body ratio [27], a general high ability in cooperative tasks [28,29], and equal rhesuses in many learning tasks (see §1). Any of these might have led to the prediction that capuchins would cluster with the Old World monkeys, even if uniquely among New World primates. On the other hand, despite the capuchins' apparent advantages (for solving this task) over other New World monkeys, cognitive differences favouring rhesus monkeys over capuchins have been seen in other recent tasks involving higher-order cognitive abilities, such as metacognition [16,30].

Our final intriguing result is that outcomes in these games (and, in particular, the inefficiencies in decision-making) were based on an inability to coordinate on the payoff-dominant outcome, not an inability to locate it owing to a lack of exploration of the problem space. None of our subjects, including humans, immediately reverted to the payoff-dominant choice after having first experienced it. This was true whether they could or could not see their partners' choice when it was made and, in the case of humans, whether or not they could talk to their partner. This observation is clearly contrary to what would be expected if finding the payoff-dominant solution is the only thing required to solve this game. Moreover, this result indicates that there are really two aspects involved in successful performance. First is exploring the parameter space and locating the payoff-dominant outcome. In this, we saw great variability within each species, as indicated in our analysis of their overall results. Second is the ability of pairs to coordinate on that outcome once they have found it. This is explored in our analysis of the choices following the first Stag-Stag play. The data indicate that both of these aspects of the Assurance Game present challenges to all three species; obviously individuals who do not fully explore the parameter space may never find the payoff-dominant outcome, but even after this has been found, coordination is not assured.

While we are strongly in favour of testing additional primates (and non-primates) of all taxa, given these data, we predict that evidence will continue to favour a New World/Old World split in decision-making outcomes. We are also enthusiastic to see how the inclusion of different game features that may require other abilities, such as theory of mind, will affect species' play. We predict that such studies will serve to illuminate not only the phylogeny of decision-making, but also the cognitive requirements of different decision-making situations, as illustrated by various economic games. In particular, some rhesus monkeys were notable for their success, a pattern of results that clearly deserves additional investigation. One possibility for investigating these mechanisms is to see whether humans and rhesus monkeys differ in the flexibility with which they can switch between strategies. This question is more than academic; if humans and rhesus monkeys are using similar mechanisms (e.g. a cognitive homology), then we know that humans' abilities are built on a shared foundation that extends back at least as far as the split with Old World monkeys. On the other hand, if humans and rhesus monkeys have converged on similar outcomes despite using different cognitive mechanisms to reach those outcomes, then we can begin to investigate what

shared social or ecological factors in humans' and rhesuses' environments selected for these outcomes.

Finally, uniquely among the primates, language is an important mechanism for solving coordination tasks in humans, much as additional information was used by all species in the Asynchronous task. In other words, we hypothesize that humans may use communication to transform a Synchronous task into an Asynchronous one by providing information about future moves. This may indicate that some underlying mechanisms, such as the ability to increase performance when cues are present, are conserved between humans and other primates, albeit with differing specific mechanisms. We also think the most likely explanation for the humans who did not communicate about (and thus did not solve) the task is not that they were incapable of doing so. Even subjects who did not solve the task spoke to one another. Instead, we propose that they thought that they had solved the task, did not see the necessity of exploring other options, and so did not converse about it. Future research aimed at determining which factors cause these different reactions may help to clarify both individual (e.g. personality) and contextual (e.g. social) factors that affect humans' ability to coordinate in such situations.

The picture that is emerging makes it clear that humans are not alone in their ability to find efficient, payoff-maximizing outcomes in a coordination game. Moreover, other species' behaviours can be measured using the same experimental mechanisms common in humans—that is, economic games—and results can be made comparable when procedures are equalized across species, including humans. Future work should be done to investigate how cognitive mechanisms interact with game structure, and whether there is homology in the underlying cognitive mechanisms or instead whether similar outcomes are reached in analogous ways. Finally, it will be interesting to see whether this similarity in outcome remains in other, more challenging decision-making situations.

This research was funded by NSF HSD grant SES 0729244 to S.F.B., B.J.W. and M.J.B. S.F.B. was additionally funded by NSF CAREER award SES 0847351; M.J.B. was additionally funded by NSF grants BCS-0924811 and BCS 0956993, and NICHD grants HD-38051, HD-061455 and HD-060563; and B.J.W. was additionally funded by NSF SES 0833310. We thank Betty Chan, Theodore Evans, Audrey Parrish and Kate Talbot for assistance with data collection at the LRC, Jennifer Cunningham for subject recruitment and data collection at Chapman University, and Jeffrey Kirchner for software programming at Chapman University. We additionally thank the animal care and enrichment staff for maintaining the health and wellbeing of the primates and making this research possible. All authors contributed equally in experimental design, data collection, analysis and writing. The authors declare no competing interests.

## ENDNOTE

<sup>1</sup>Though we should note that the former experiment involves more than two players with more than two potential actions, and the latter experiment uses a  $3 \times 3$  coordination experiment.

## REFERENCES

- van Huyck, J., Battalio, R. & Beil, R. 1990 Tacit coordination games, strategic uncertainty, and coordination failure. *Am. Econ. Rev.* **80**, 234–248.
- Jensen, K., Call, J. & Tomasello, M. 2007 Chimpanzees are rational maximizers in an ultimatum game. *Science* **318**, 107–109. (doi:10.1126/science.1145850)
- Camerer, C. 2003 *Behavioral game theory: experiments in strategic interaction*. Princeton, NJ: Russell Sage Foundation/Princeton University Press.
- Smith, P. & Silberberg, A. 2010 Rational maximizing by humans (*Homo sapiens*) in an ultimatum game. *Anim. Cogn.* **13**, 671–677. (doi:10.1007/s10071-010-0310-4)
- Skyrms, B. 2003 *The stag hunt and the evolution of social structure*. Cambridge, UK: Cambridge University Press.
- Brosnan, S. F., Parrish, A. R., Beran, M. J., Flemming, T. E., Heimbauer, L., Talbot, C. F., Lambeth, S. P., Schapiro, S. J. & Wilson, B. J. 2011 Responses to the Assurance game in monkeys, apes, and humans using equivalent procedures. *Proc. Natl Acad. Sci. USA* **108**, 3442–3447. (doi:10.1073/pnas.1016269108)
- McCleam, G. & Harlow, H. F. 1954 The effect of spatial contiguity on discrimination learning by rhesus monkeys. *J. Comp. Physiol. Psychol.* **47**, 391–394. (doi:10.1037/h0059728)
- Murphy, J. & Miller, R. 1959 Spatial contiguity of cue, reward, and response in discrimination learning by children. *J. Exp. Psychol.* **58**, 485–489. (doi:10.1037/h0041157)
- Stollnitz, F. & Schrier, A. M. 1962 Discrimination learning by monkeys with spatial separation of cue and response. *J. Comp. Physiol. Psychol.* **55**, 876–881. (doi:10.1037/h0044549)
- Beran, M. J., Rumbaugh, D. M. & Washburn, D. A. 2007 Rhesus monkeys (*Macaca mulatta*) maintain learning set despite second-order stimulus-response spatial discontinuity. *Psychol. Record* **57**, 9–22.
- Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S. & Hopkins, W. D. 1989 Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *J. Comp. Psychol.* **103**, 32–38. (doi:10.1037/0735-7036.103.1.32)
- Beran, M. J., Decker, S., Schwartz, A. & Shultz, N. 2011 Monkeys (*Macaca mulatta* and *Cebus apella*) and human adults and children (*Homo sapiens*) compare subsets of moving stimuli based on numerosity. *Front. Psychol.* **2**, 61. (doi:10.3389/fpsyg.2011.00061)
- Beran, M. J., Harris, E. H., Evans, T. A., Klein, E. D., Chan, B., Flemming, T. E. & Washburn, D. A. 2008 Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*), and rhesus monkeys (*Macaca mulatta*): the effects of differential and nondifferential reward. *J. Comp. Psychol.* **122**, 52–61. (doi:10.1037/0735-7036.122.1.52)
- Evans, T. A. 2007 Performance in a computerized self-control task by rhesus macaques (*Macaca mulatta*): the combined influence of effort and delay. *Learn. Motiv.* **38**, 342–357. (doi:10.1016/j.lmot.2007.02.001)
- Flemming, T. M., Thompson, R. K. R., Beran, M. J. & Washburn, D. A. 2011 Analogical reasoning and the differential outcome effect: transitory bridging of the conceptual gap for rhesus monkeys (*Macaca mulatta*). *J. Exp. Psychol. Anim. Behav. Process.* **37**, 353–360. (doi:10.1037/a0022142)
- Beran, M. J., Smith, J. D., Coutinho, M. V. C., Couchman, J. J. & Boomer, J. 2009 The psychological organization of 'uncertainty' responses and 'middle' responses: a dissociation in capuchin monkeys (*Cebus apella*). *J. Exp. Psychol. Anim. Behav. Process.* **35**, 371–381. (doi:10.1037/a0014626)
- Beran, M. J. & Smith, J. D. 2011 Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition* **120**, 90–105. (doi:10.1016/j.cognition.2011.02.016)



- 18 Penn, D. C. & Povinelli, D. J. 2007 On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Phil. Trans. R. Soc. B* **362**, 731–744. (doi:10.1098/rstb.2006.2023)
- 19 Flombaum, J. I. & Santos, L. R. 2005 Rhesus monkeys attribute perceptions to others. *Curr. Biol.* **15**, 447–452. (doi:10.1016/j.cub.2004.12.076)
- 20 Cooper, D. J., Feltovich, N., Roth, A. E. & Zwick, R. 2003 Relative versus absolute speed of adjustment in strategic environments: responder behavior in ultimatum games. *Exp. Econ.* **6**, 181–207.
- 21 Ochs, J. 1995 Coordinate problems. In *The handbook of experimental economics* (eds J. Kagel & A. E. Roth), pp. 195–252. Princeton, NJ: Princeton University Press.
- 22 Smith, V. L. 1982 Microeconomic systems as an experimental science. *Am. Econ. Rev.* **72**, 923–955.
- 23 Cooper, R., DeJong, D., Forsythe, R. & Ross, T. 1990 Selection criteria in a coordination game: some experimental results. *Am. Econom. Rev.* **80**, 218–233.
- 24 Brosig, J. 2002 Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. *J. Econom. Behav. Organ.* **47**, 275–290. (doi:10.1016/S0167-2681(01)00211-6)
- 25 Fetchenhauer, D., Grootuis, T. & Pradel, J. 2010 Not only states but traits—humans can identify permanent altruistic dispositions in 20 s. *Evol. Hum. Behav.* **31**, 80–86. (doi:10.1016/j.evolhumbehav.2009.06.009)
- 26 Cumming, W. W. & Berryman, R. 1961 Some data on matching behavior in the pigeon. *J. Exp. Anal. Behav.* **4**, 281–284. (doi:10.1901/jeab.1961.4-281)
- 27 Rilling, J. K. & Insel, T. 1999 The primate neocortex in comparative perspective using magnetic resonance imaging. *J. Hum. Evol.* **16**, 191–233. (doi:10.1006/jhev.1999.0313)
- 28 de Waal, F. B. M. & Berger, M. L. 2000 Payment for labour in monkeys. *Nature* **404**, 563. (doi:10.1038/35007138)
- 29 Mendres, K. A. & de Waal, F. B. M. 2000 Capuchins do cooperate: the advantage of an intuitive task. *Anim. Behav.* **60**, 523–529. (doi:10.1006/anbe.2000.1512)
- 30 Hampton, R. R. 2001 Rhesus monkeys know when they remember. *Proc. Natl Acad. Sci. USA* **98**, 5359–5362. (doi:10.1073/pnas.071600998)