

Cognitive Preconditions for Responses to Fairness: An Object Retrieval Test of Inhibitory Control in Capuchin Monkeys (*Cebus apella*)

Venkat R. Lakshminarayanan and Laurie R. Santos
Yale University

The authors explore the evolution of one cognitive mechanism required for altruistic behavior: the capacity to inhibit prepotent responses. Specifically, the authors used an object retrieval task to investigate whether capuchins (*Cebus apella*) can inhibit a prepotent strategy of reaching directly for a food reward. Success in this task varies across species and across development, but is also known to depend critically on the maturity of dorsolateral prefrontal cortex, the cortical area implicated in rejecting small payoffs in an Ultimatum Game. Capuchins easily inhibit the tendency to reach directly for food in the object retrieval task, successfully employing an alternative reaching strategy even in the first session of performance. This contrasts with the performance of closely related tamarin monkeys, who performed less well despite extensive training. These results provide the first evidence that capuchins likely exhibit human-like inhibitory control in tasks previously linked to the function of the dorsolateral prefrontal cortex, such as the Ultimatum Game.

Keywords: fairness, Ultimatum Game, prefrontal cortex, Capuchin monkeys

A standard assumption in classical economics is that human decision makers should consistently prefer options that maximize their own expected payoffs. In contrast to this standard assumption of self-interest, people regularly behave in ways that require inhibiting the desire to behave in a purely self-interested way—people volunteer for the military, donate blood, give money to charities, and so on. Humans also regularly violate their own self-interest to enforce similar fair behaviors in others, punishing those that act unfairly often at a substantial cost to themselves. In these and other examples, humans routinely provide striking examples of inhibitory control: we frequently are able to inhibit our own self-interest to behave both altruistically and in ways that uphold fairness norms (see Stevens & Hauser, 2004).

Over the last few years, neuroscientists and other researchers have become interested in the neural mechanisms required to inhibit prepotent responses in the service of altruistic and social

norms. In one recent paper, Sanfey and colleagues (2003) used functional imaging techniques to investigate the neural systems underlying performance on the Ultimatum Game, a one-shot game in which two anonymous players must decide how to split a cash offer. The first individual—the proposer—proposes a potential division of the offer. The second individual—the responder—then decides whether to accept or reject the proposed division. If the responder accepts the proposer's offer, then the money is split as proposed, but if the responder rejects the proposer's offer, then neither player earns any money. Sanfey and colleagues scanned participants during an Ultimatum Game and examined the systems that underlie responders' spiteful rejection behavior in this task. Curiously, a particular neural region involved in inhibitory control—the dorsolateral prefrontal cortex (DLPFC)—was selectively recruited when responders distinguished between fair and unfair offers, suggesting that this region is needed to inhibit the prepotent self-interested urge to accept the monetary reward. Testing this prediction more directly, Knoch and colleagues used transcranial magnetic stimulation (TMS) to disrupt DLPFC regions while responders were evaluating fair and unfair offers (Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006).

Venkat R. Lakshminarayanan and Laurie R. Santos, Department of Psychology, Yale University.

Correspondence concerning this article should be addressed to Venkat Lakshminarayanan, Department of Psychology, Yale University, 2 Hillhouse Ave., New Haven, CT 06520. E-mail: venkat.lakshminarayanan@yale.edu

tapraid5/nec-npe/nec-npe/nec00109/nec0002d09z | xppws | S=1 | 5/11/09 | 14:39 | Art: 2009-0001

INHIBITORY CONTROL IN CAPUCHINS

13

They observed that participants whose right DLPFC regions were disrupted were substantially more likely to accept low offers than matched controls. Interestingly, these participants accepted low offers despite the fact that they evaluated these low offers as unfair (Knoch et al., 2006). These results suggest that DLPFC-mediated inhibitory control systems play a key role in our human-like capacity to inhibit selfish choices, particularly in the context of foregoing self-interested payoffs in the service of altruistic acts.

At the same time that neuroscientists have begun unearthing the importance of inhibitory control systems in human altruistic behavior, comparative researchers have become interested in the evolutionary origins of our capacity for self-control. This comparative work has focused on the broad question of whether other primate species share the human capacity to inhibit prepotent responses in a variety of contexts and experimental tasks. To date, this work has revealed a complicated pattern of performance. Most primates have tended to perform poorly on one general test of inhibitory control known as the *reverse contingency task* (see Russell, Mauthner, Sharpe, & Tidswell, 1991). In this task, a subject is faced with a choice between either a large or a small food reward. The subject is then required to point to the smaller reward to obtain the larger one, to succeed at the task. To do so, subjects must inhibit the tendency to reach directly for the larger reward. Chimpanzees (*Pan troglodytes*) (Boysen & Bernston, 1995; Vlamings, Uher, & Call, 2006) and other primates (Kralik, Hauser, & Zimlicki, 2001; Silberberg & Fujita, 1996) have tended to perform extremely poorly on this kind of task, except in cases where symbols are used instead of food rewards (Boysen & Bernston, 1995). In contrast, another inhibitory control test—a discounting task in which subjects must forgo an immediate reward in favor of a larger delayed reward—has revealed an even more complicated pattern of performance. Recent work suggests that chimpanzees are willing to forgo an immediate small reward for a larger delayed reward even at very long delays (Beran, Savage-Rumbaugh, Pate, Rumbaugh, 1999; Beran, 2002; Rosati, Stevens, Hare, & Hauser, 2007), but a more distantly related New World monkey species—the capuchin, *Cebus apella*—appears to lack this patience, taking the imme-

diately reward even when offered a delayed food item that is nearly 40 times greater in value (Ramseyer, Pelé, Dufour, Chauvin, & Thierry, 2006). Taken generally, then, studies of inhibitory control conducted thus far with monkeys and apes paint a complicated picture of the inhibitory control capacities of nonhuman primates.

More recently, comparative researchers have begun examining the nature of primates' inhibitory capacities in the more specific context of social decision making, namely contexts much like the ones in which humans tend to altruistically forego their own self-interest. This work to date has focused mostly on two model primate species—the chimpanzee and the brown capuchin monkey. Both species possess relatively large brains, live in complicated social groups, and exhibit cooperative-like behaviors (e.g., food sharing and cooperative hunting) in the wild (see de Waal, 1996), which make them ideal model species for exploring inhibitory control specifically in the context of social decision-making situations. Unfortunately, however, there remains to date little consensus about the extent to which chimpanzees and capuchins can, like humans, forego their own self-interest in the context of social decision making (see reviews in Brosnan, 2006; de Waal, 1996, 2008; Silk, 2009). Both chimpanzees (Warneken & Tomasello, 2006; Warneken, Hare, Melis, Hanus, & Tomasello, 2007) and capuchins (Barnes, Martinez, Langer, Hill, & Santos, 2008; de Waal and Berger, 2000) sometimes behave in costly ways that serve to benefit other individuals. For example, these species will occasionally work to give a human experimenter an out of reach object (Barnes et al., 2008; Warneken & Tomasello, 2006; Warneken et al., 2007). In addition, there is evidence that chimpanzees (Braüer, Call, & Tomasello, 2006; Brosnan, 2006; Brosnan, Schiff, & de Waal, 2005) and capuchins (Brosnan & de Waal, 2003; van Wolkenten, Brosnan, & de Waal, 2007) will sometimes forego a food reward delivered by an unfair experimenter, one who has previously paid another monkey a better reward for the same amount of work (but see Dindo & de Waal, 2007; Dubreil, Gentile, & Visalberghi, 2006; Roma, Silberberg, Ruggiero, & Suomi, 2006 for failures to demonstrate this effect in the absence of any trading effort). In contrast, however, there are many more cases in which both of

these species fail to forego their own self-interest in altruistic tasks (e.g., Silk et al., 2005; Jensen, Hare, Call, & Tomasello, 2006; Silk et al., 2005; Barnes et al. 2008). To take one relevant example, Jensen, Call, & Tomasello (2007a) presented chimpanzees with a nonverbal Ultimatum Game and found that chimpanzee responders regularly accepted small or non-zero offers, suggesting that chimpanzees are unable to forego their own self interest (see also Jensen, Call, & Tomasello, 2007b). Taken together, the work exploring capuchin and chimpanzee inhibitory control within the context of social decision-making studies also paints a conflicting picture of these species' capacity to forego their own self-interest. Therefore, it remains unclear whether these primate species are capable of inhibiting their own self-interest in the specific social decision making contexts in which humans exhibit self-control.

One way to gain more insight into this confusing picture of primates' inhibitory performance is take a mechanistic hint from human neuroimaging work on social decision making. As reviewed above, neuroscientists have observed that one specific neural region—the DLPFC—is critically involved in our capacity to forego selfish payoffs in the context of altruistic games. Therefore, these findings suggest that comparative researchers may want to focus specifically on DLPFC-mediated capacities in other primates to gain insight both into the mechanisms underlying primate altruistic behavior and into the evolution of inhibitory capacities more generally. The goal of the present study is to do just this. Specifically, we examine how one model primate species—the brown capuchin—performs on a behavioral test of inhibitory control that has specifically been linked to the function of the DLPFC, a test known as the object retrieval task (see Diamond 1991a, 1991b; Diamond & Gilbert, 1989; Dias, Robbins, & Robertset, 1996; Santos, Ericson, & Häuser, 1999). In the object retrieval task, subjects are presented with the opportunity to retrieve a desired object from inside a small transparent box. The logic behind this task is that subjects must inhibit the prepotent tendency to reach directly for the object to successfully obtain it; rather than reaching straight for the box, the participant must first locate the opening (which is often on the side or top) and reach there instead. The object retrieval task has been

used both as behavioral test of inhibitory control in human infants (see reviews in Diamond 1991a, 1991b), and as a comparative measure of inhibitory control across many primate species (Diamond, 1990, 1991a, 1991b; Dias et al., 1996; Santos et al., 1999). This work has revealed striking differences in performance across different primates species; whereas some monkey species naturally perform at mature human levels (e.g., Diamond, 1990; Dias et al., 1996), other species perform at the level of the youngest human infants (Santos et al., 1999). Indeed, adult tamarins—a species closely related to capuchin monkeys—were only able to succeed on the object retrieval task when they were first trained to retrieve food from an easier, opaque box. Finally, and most importantly for the purpose of this paper, performance on the object retrieval test has been specifically linked to DLPFC function (Diamond 1991a, 1991b; Dias et al., 1996). Lesions of the DLPFC impede successful performance on the object retrieval task (e.g., Diamond 1991a, 1991b; Dias et al., 1996), and thus the object retrieval task can be considered a behavioral test of the inhibitory processes controlled by neural areas implicated in the Ultimatum Game.

To better explore the inhibitory constraints underlying primates' performance on experimental tests of cooperation, we extended the object retrieval test to the brown capuchin monkey, one of the two species most commonly tested not only in prior studies of inhibitory control, but also in previous studies of cooperation and social decision making. To facilitate the most accurate cross-species comparisons, we presented capuchins with a version of the object retrieval test previously used with tamarin monkeys (Santos et al., 1999). If capuchins are capable of inhibiting their prepotent reaching desires better than other primate species, they would be expected to perform better than tamarin monkeys on an identical version of this task. In contrast, if capuchins have only a limited capacity for inhibitory control, then they may perform similarly to tamarins, failing to inhibit the prepotent tendency to reach straight for the reward.

Method

We tested four male (N.N., F.L., A.G., J.B.) and two female (H.G., J.M.) brown ca-

INHIBITORY CONTROL IN CAPUCHINS

puchin monkeys, ranging in age from 1.5 to 10 years. Our capuchin participants were members of the capuchin colony at the Yale Comparative Cognition Laboratory. All capuchins were communally housed in a large enclosure equipped with toys and natural branches. Monkeys were provisioned with monkey chow after testing and had access to water ad libitum. All subjects had previously participated in experiments concerning object cognition, numerical representation, and tool use but had not yet been tested in an object retrieval task.

We conducted our test sessions in a cubical mesh testing enclosure (approximately 71 cm³) adjacent to the main cage. One enclosure wall, a locking Plexiglas slider, served as a door to the main cage. At the center of the opposite enclosure wall was an opening (of 3" diameter) through which a subject could manipulate experimental stimuli. Our studies involved an ob-

ject retrieval box (see Figure 1), a hollow plastic cube (about 13 cm × 13 cm × 13 cm) that had five solid sides and an open side (the position of which changed from trial to trial). We also used a separate box for the Opaque condition that was identical to the first box except that it was opaque rather than transparent. We were able to secure each experimental box to a plastic pedestal (15 cm high) to ensure that the box remained positioned in front of the reach-hole during test trials. We also positioned an opaque occluder against the reach-hole to regulate the monkeys' visual and physical access to the box during preparation and set-up. We used grapes, which had been sliced in half so that they would not roll once put into place, as a food reward throughout all sessions. All trials were videotaped by a Sony Handycam.

We closely mimicked the procedure used by Santos et al. (1999). All monkeys were required to retrieve a piece of food from inside one of the two

FI

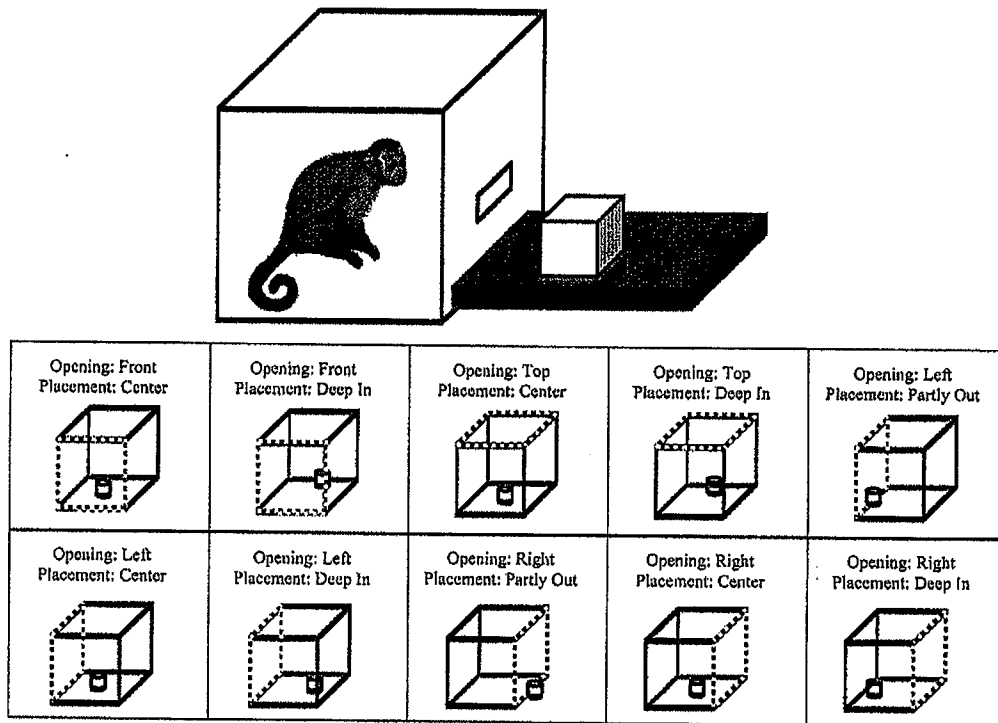


Figure 1. A diagram of the set-up. Monkeys were placed inside a testing box enclosure and could reach out into the box to obtain food. Across the session, we varied the location of the opening of the box ("Opening") and the location of the marshmallow within the box ("Placement").

tapraid5/nec-npe/nec-npe/nec00109/nec0002d09z	xppws	S=1	5/11/09	14:39	Art: 2009-0001
---	-------	-----	---------	-------	----------------

experimental boxes. Before beginning, we randomly classified our subjects into two test groups: a transparent box testing group (J.B., H.G., and F.L.) and an opaque box testing group (N.N., A.G., and J.M.). The two groups were tested using nearly the same procedure, except that the subjects in the opaque group were first trained with an opaque testing box before beginning their experimental trials with the transparent box.

The transparent box testing group began testing with the transparent box. Each session began when the experimenter isolated the monkey inside the testing enclosure. Monkeys received 24 trials in each session, just as in previous studies (e.g., Diamond & Gilbert, 1989; Santos et al., 1999). The experimenter began each trial by placing the occluder over the reach hole. He then placed the food inside of the transparent box and then secured the box onto the plastic pedestal. The experimenter then removed the occluder and the monkey was given 15 s to retrieve the food from inside the box. When 15 s had elapsed, or the monkey had retrieved the grape, the experimenter replaced the occluder and began set-up for the next trial. A correct trial was one in which the monkey was able to retrieve the reward within the first 15 s, which could only be achieved when the monkey successfully reached into the opening. Our logic was that subjects would need to quickly revise any incorrect reaching strategies to successfully retrieve food during a 15-s window; any preservative errors such as repeatedly reaching for a closed side of the box would

prevent subjects from obtaining food during this short window. Trials were initially scored online during testing and then were rechecked by a coder blind to the hypothesis. Interobserver reliability between the online code and the blind coding was extremely high (98%).

As in previous object retrieval studies (Diamond & Gilbert, 1989; Santos et al., 1999), the 24 trials within each session varied on the placement of the box's opening (front, top, left, right) and the position of the grape in the box (partly out, in the center, or deep inside). All monkeys received a preset order of these sessions as described in Table 1; this preset order was used to best match the way that previous object retrieval tests were conducted with other species. As in previous studies, we distinguished between two kinds of trials: more difficult trials in which inhibition was required, which we labeled *critical* and the rest of less difficult trials that we referred to as *normal* (Santos et al., 1999). Subjects continued with the transparent box until they performed above 90% on critical trials for two consecutive sessions with the transparent trial box.

The opaque box testing group received a task identical to the one presented to the transparent group except that the box used was made of opaque white plastic. As in previous studies, monkeys remained on this condition until they completed one perfect session of 24 trials with this opaque box. After reaching this criterion with the opaque box, monkeys moved onto subsequent sessions with the transparent box. Subjects con-

T1

Table 1
The Order of Object Retrieval Trials Presented to Capuchins, After Santos et al. (1999)

No.	Opening	Placement	Type	No.	Opening	Placement	Type
1	Front	Deep in	Normal	13	Left	Deep in	Critical
2	Front	Deep in	Normal	14	Right	Deep in	Critical
3	Right	Partly out	Normal	15	Top	Center	Critical
4	Right	Deep in	Critical	16	Top	Deep in	Critical
5	Right	Partly out	Normal	17	Top	Center	Critical
6	Right	Center	Critical	18	Top	Deep in	Critical
7	Right	Deep in	Critical	19	Top	Center	Critical
8	Left	Partly out	Normal	20	Top	Deep in	Critical
9	Left	Deep in	Critical	21	Left	Deep in	Critical
10	Left	Partly out	Normal	22	Right	Deep in	Critical
11	Left	Center	Critical	23	Top	Deep in	Critical
12	Left	Deep in	Critical	24	Front	Deep in	Normal

Note. Across this set order of trials, we varied the position of the opening and the placement of the marshmallow, resulting in a set order of "normal" and "critical" trials.

