

The Evolution of Morality: Which Aspects of Human Moral Concerns Are Shared With Nonhuman Primates?

Mark Sheskin *and* Laurie Santos

Abstract

Morality is a critical part of human society. This chapter explores the origins of human morality by examining whether nonhuman primate species share aspects of five domains thought to be important in human moral behavior—concerns involving harm, fairness, hierarchy, ingroup allegiance, and purity. Behaviors in the harm domain have received the most attention from researchers, and converging lines of evidence suggest that some primates express harm concerns. The domain of fairness has become a recent focus of primate research, with active debate about whether closely related primates share human-like concerns. Moral behaviors regarding ingroup allegiance, authority, and purity have received the least attention in nonhuman species, though recent work suggests that research with primates might productively pursue the ingroup allegiance and authority domains. Future primate research will continue to elucidate the nature of human morality, and should include an increased focus on the previously neglected domains of ingroup and hierarchy.

Key Words: morality, nonhuman primates, fairness, cognitive origins

Introduction

It was feeding time and, as usual, Felix was the first in line. As the head of his group, he got to eat before all others. He stepped up to the bounty and began leisurely eating more than he needed, never looking back to the others in his group who watched him from behind, waiting silently. Suddenly, there was a crash out of view. Felix paused for a second and headed off to look at what the commotion was. Most of the others soon followed Felix, but Ric, the lowest-ranking member of the group, stayed behind. As soon as Felix was out of view, Ric quickly ran up to the food and grabbed as much as he could, shoving it into his mouth quickly. When Felix finished looking in the direction of the crash, he turned back, and saw Ric feeding. Felix retaliated immediately. He screamed loudly, grabbed the food from Ric's hands, and reached for his throat. Ric ran off crying, as the others watched the entire altercation in silence.

As humans, we often experience social interactions as complicated as the one just described. Despite their complexity, we quickly and easily make sense of such scenarios—understanding who did what to whom in a way that lets us both interpret these events and make predictions about what will occur next. As humans, our comprehension of complex social scenarios goes beyond the mere surface properties of the events we witness. When reading the above scenario, for example, we spontaneously interpret the agents' behaviors not merely in terms of what each agent did—how they behaved—but also in terms of what the agents intended, thought, and experienced. We quickly recognize that Felix *intended* to keep the food all to himself, that Ric was *hungry* and *trying* to deceive Felix, and that Felix was *outraged* when he realized what happened and *wanted* to punish Ric. In this way, humans naturally interpret social events

in ways that go beyond the mere behavior witnessed, inferring the mental states that are causally responsible for the agents' actions.

However, humans take a step even beyond these mentalistic causal inferences when we interpret social events. Just as we spontaneously interpret agents' behaviors in mentalistic terms, we also spontaneously *evaluate* those behaviors, deeming them good or bad, acceptable or unacceptable, fair or unfair, moral or immoral, and so on. When we witness a scenario like the one described, we naturally form evaluations of the actors' behaviors, thinking for example that it's *unfair* for Felix to keep the food to himself, *dishonest* for Ric to take food from his group mate when he's not looking, or *wrong* for group mates not to step in during a physical altercation. Such moral evaluations share several features with our mentalistic interpretations of others' behavior. First, moral evaluations require us to go beyond a simple behavioral interpretation of the events we have witnessed. Indeed, even scenarios that are simple at the behavioral level can involve nuanced moral evaluations. We might evaluate Felix's behaviors in the preceding scenario differently if we knew more about the history of interactions between him and Ric, the way that Felix first came to his high status, and so on. In addition, like our omnipresent mentalizing, our tendency to evaluate actions in moral terms is ubiquitous. People universally make moral evaluations of others' behavior. Moreover, we apply our evaluations not just to the actions of others, but also to our own behaviors. Such self-evaluations mean that people often evaluate whether their own behaviors are right or wrong, which much of the time seems to motivate them to behave in ways that are good and avoid actions that would be evaluated as wrong. Our human tendency to spontaneously evaluate actions and act on the basis of such evaluations—our human moral cognition—is a fundamental aspect of human social life, one that plays out universally in our species and dictates much of our social interactions.

Of course, humans are not the only species forced to navigate complicated social situations. Like humans, nonhuman primates (hereafter, primates) face a complex array of social events. Consider, for example, the events described in the opening scenario. Although you would be forgiven for assuming that this story involved human agents, the scenario described actually involved a group of capuchin monkeys living in our colony at Yale University. Capuchin monkeys, like other primates, consistently deal with individuals who willfully try to deceive, unfairly attempt to harass, and so on. Given that other primates face

a social environment almost as complicated as our own, do they also bring to bear the same cognitive machinery to process these events?

Primate researchers have made considerable headway in exploring at least some of these questions. Over the past few decades, much empirical work has examined the question of whether primates also interpret their complex social world in the same mentalistic way as humans, namely, by representing agents in terms of their unseen mental states (e.g., Rosati, Santos, & Hare, 2009; Tomasello, Hare, & Call, 2003a, 2003b; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Although there is still considerable controversy about the extent to which primates represent others' behavior mentalistically (Hare, Addessi, Call, Tomasello, & Visalberghi, 2003; Penn & Povinelli, 2007; Povinelli & Vonk, 2003; Tomasello et al., 2003a; 2003b; Tomasello et al., 2005), a considerable body of work has demonstrated that many primates behave in ways that are consistent with an understanding of others' perceptions (Flombaum & Santos, 2005; Hare, Call, Agnetta, & Tomasello, 2000) knowledge (Hare, Call, & Tomasello, 2001; Kaminski, Call, & Tomasello, 2008; Santos, Nissen, & Ferrugia, 2006), and intentions (Call, Hare, Carpenter, & Tomasello, 2004; Phillips, Barnes, Mahajan, Yamaguchi, & Santos, 2009)

Less work, however, has investigated another aspect of primate social cognition—whether primates also share human-like moral considerations when watching and acting in social activities. Do primates, like humans, evaluate others' actions as moral or immoral? Do they represent actions as fair or unfair? Do primates decide how to behave based on notions of right and wrong? Primate researchers have been able to gain new insight into these questions in just the last few years. In this chapter, we review this recent empirical work in an attempt to address which features of human moral cognition might be shared with nonhuman primates. We first outline the kinds of domains in which moral cognition has been examined in our own species and then turn to what is known about how primates reason in these domains. Although there are still many questions to be addressed, recent empirical work provides new insight about the kinds of foundational moral capacities that are and are not shared broadly within our evolutionary order.

Carving Out the Domains of Moral Cognition

Before launching into a review of the foundation of moral cognition in primates, we must first discuss

two caveats about our approach. The first caveat involves the question of how to carve up human moral cognition to investigate the possibility that similar capacities exist in other species. Given that there are centuries of philosophical inquiry into the nature of human morality, finding a simple definition of human moral cognition that could apply to primates is undoubtedly going to be tricky. However, a number of psychologists have successfully argued that human moral concerns—even when viewed cross-culturally—tend to fall into a relatively small number of specific domains (Haidt & Joseph, 2004, 2007; Hauser, 2006; Shweder, Much, Mahapatra & Park, 1997).

Although there is superficial societal variation in the nature of people's moral concerns, people universally tend to consider the same *types* of behaviors when they evaluate others' actions as "good" or "evil." Haidt and his colleagues (Graham & Haidt, 2010; Graham, Haidt, & Nosek, 2009; Haidt, 2008; Haidt & Graham, 2007; Haidt & Joseph, 2004, 2007; Haidt & Kesebir, 2009) have divided these types of human moral concerns into a set of five "foundational" domains of morality. These domains include: harm (concern for the welfare of others), fairness (concern for equitable outcomes), ingroup allegiance (concern for the welfare of the group), authority (concern for maintenance of group hierarchy), and purity (concerns for the sacredness of certain objects and actions). The content of each of the five domains can be illustrated with example items from Haidt's "Moral Foundations Sacredness Scale" (Graham & Haidt, 2010), which asks how much money you would require to agree to do various unsavory things. The harm domain includes actions such as kicking a dog in the head and taunting an overweight person. The fairness domain includes actions such as cheating in a game of cards with strangers and agreeing to secretly hire only same-race applicants for a job. Violations of ingroup allegiance include actions such as breaking off communication with all family members for a year and changing citizenship to another country. Examples of immoral actions in the authority domain include throwing a rotten tomato at a disliked political leader and slapping your father as part of a rehearsed comedy skit. Finally, violations in the purity domain include getting a two-inch tail surgically added for three years and getting a blood transfusion from a child molester.

Haidt's five domains of moral consideration have provided a promising approach to studying moral cognition in humans for several reasons (Graham et

al., 2009). First, these five domains capture the moral intuitions not just of participants from Western educated backgrounds (who tend to focus mostly on harm and fairness violations), but also the kinds of intuitions commonly observed in non-Western cultures, in which people tend to focus more on issues related to ingroup allegiance, authority, and purity (Haidt & Graham, 2007; Haidt & Joseph, 2007). In this way, Haidt's foundational domains capture the universal aspects of human moral concerns, exactly the ones we might want to focus on when taking an evolutionary approach to similar concerns in primates. Second, Haidt and colleagues developed these five foundational domains with an eye for the different kinds of evolutionary selection pressures that may have led to human moral intuitions (see Haidt and Joseph, 2007, for a review). For example, Haidt and colleagues hypothesize that harm concerns emerged as a result of selection pressures to protect vulnerable yet closely related kin, while concerns about purity emerged based on pressures to avoid microbes and other pathogens. In this way, Haidt's organization of human moral concerns is based on the idea that our moral intuitions evolved in response to selection pressures in exactly the same way we might expect if some aspects of these concerns were shared in other primate species.

The second caveat, however, concerns how we can determine whether primates share a human-like moral concern in these domains. For a verbal species like humans, determining a subject's moral concerns can often be addressed simply by asking: human participants can be asked whether they consider the actions verbally presented in a scenario to be "right" or "wrong." The situation is much trickier when you're dealing with nonverbal subjects like primates. Given that primates cannot verbally express whether they consider any actions to be good or bad, what evidence can we use to determine whether they share human-like moral considerations? Although this question is sure to generate much debate among philosophers and animal cognition researchers, in this review we have chosen to use two kinds of evidence to argue for similarity across humans and primates. The first piece of evidence we discuss concerns whether primates *behave* in ways that are consistent with possessing a given moral concern. When humans think an action is wrong, they tend to avoid doing it. Similarly, when people think actions are permissible or obligatory, they tend to engage in them. In this way, we can use evidence about whether primates behave in ways that are consistent with certain moral considerations as

evidence for possessing such concerns. For example, do primates engage in actions that violate authority or fairness considerations? Or do they instead make choices in ways that are consistent with obeying these considerations? By examining whether primates themselves behave in ways that are consistent with different considerations, we can see—at the very least—whether they behave as though they represented that such a moral norm was in place.

A second form of evidence concerns whether primates *evaluate* the actions of others who do and do not behave in ways that are consistent with different moral considerations. When humans watch an individual performing an action that they think is wrong, they tend to evaluate that individual negatively. Such negative evaluations in humans can take different forms, including punishment, shunning, a refusal to interact with the individual in the future, and so forth. As such, we review whether primates show evidence of *evaluating others' actions* when those others violate moral considerations. Do primates refuse to interact with those that violate harm constraints or ignore ingroup allegiances? Do they punish such individuals when given the opportunity? Throughout the review, we also pay attention to whether primates' evaluations depend on *who* is being affected by the moral violation. Do primates react negatively only when *they themselves* are negatively affected by the moral violation (e.g., cases in which their own place in their hierarchy is not respected, when they are harmed personally, etc.), or do they also show similar evaluations when they are not involved in the violation, such as when only a *third party* is negatively affected (i.e., someone else is treated unfairly, a third party individual is harmed, etc.). If primates' moral evaluations operate like those of humans, then primates should negatively evaluate immoral agents based on transgressions against a wide range of targets (and certainly not just transgressions against the evaluator).

Focusing on these two kinds of evidence, we next examine whether primates seem to possess moral considerations within each of Haidt's five domains. Behaviors in the harm domain have received the most attention from researchers, and converging lines of evidence suggest that some primates express harm concerns as well. The domain of fairness has also become a recent focus of primate research, with some debate regarding whether closely related primates share human-like fairness concerns. On the other hand, although some nonhumans have social hierarchies and participate in intergroup conflicts, there is little study of primate moral concerns

regarding hierarchy or ingroup. That is, although some primates act differentially toward others based on hierarchy and ingroup status, there is little evidence that they positively or negatively evaluate others for violations of the hierarchy or ingroup. Finally, little primate work has explored concerns that fall within the purity domain, although our intuition is that purity concerns are a domain of human moral concern that might be unique to our species.¹

Evolution of Harm/Care Behaviors and Concerns

Behaviors in the harm domain focus on the physical welfare of others. Put in the most general terms, moral concerns in the domain of harm stipulate that it is moral (and sometimes obligatory) to increase others' physical welfare and often immoral to decrease or harm it. Prototypical cases of harm violations are hypothesized to involve bodily injury, but other cases of welfare removal also fall under this definition. Moral concerns in the harm domain also seem to recruit a standard set of emotions (e.g., compassion) as well as characteristic motivations for increasing others' welfare (kindness, prosocial preferences, etc.). Several researchers have hypothesized that intuitions about harm represent the most developmentally basic moral evaluations (Haidt & Joseph, 2007; Hauser, 2006; Mikhail, personal communication), and there is evidence that such intuitions come on line in our own species in the first few months of life (Hamlin, Wynn & Bloom, 2007).

The harm domain is a foundational area of moral cognition that has been extensively studied in a wide range of species. This is not surprising, given the hypothesized origin of harm concerns—namely, a motivation to help kin (see Haidt & Joseph, 2007). Indeed, many organisms increase their fitness indirectly by acting in ways that help closely related kin, either by behaving in ways that increase kin's welfare (e.g., feeding children) or decrease the risk of harm to kin (e.g., alarm calling to prevent predation; see Hamilton, 1963). A more interesting case of harm considerations, however—and one that maps most directly onto human moral concerns—involves attitudes and behaviors that are directed toward the welfare of *nonkin*. Human harm concerns move beyond closely related family members. Actions as varied as offering directions to lost tourists to donating blood anonymously suggest that humans are motivated to increase welfare and decrease harm toward unrelated (and sometimes unknown) individuals.

Do primates show similar concerns when dealing with nonrelatives? Early research with primates identified some situations in which individuals respond negatively toward the distress of other individuals. In one famous study, rhesus monkeys refrained from pulling a chain that delivered food when the chain was linked to an unrelated conspecific experiencing a painful electric shock (Masserman, Wechkin & Terris, 1964). This early study provided some of the first evidence that primates may avoid actions that cause unrelated individuals to experience pain. Unfortunately, however, this early work was also consistent with a number of other more deflationary alternatives. For example, witnessing the distress of others can often provide a good indicator of danger, and so individuals may find others' pain aversive simply because of this association. Put more basically, the motivation to avoid aversive signals in one's environment is different from the motivation to increase the welfare of other social agents; only the latter motivation falls under the moral domain of harm avoidance.

Economic Games in Humans and Nonhuman Animals

To get around the kinds of interpretational problems that plagued conspecific distress studies, researchers developed new methods to examine nonhuman harm concerns. Such new methods typically mimic the scenarios in which humans are motivated to increase others' welfare and test whether primates are willing to do the same. The most prolific of these new methods, *primate economic games*, are based on the kinds of economics games typically used to test human moral intuitions. In human versions of these games, participants are asked to allocate different kinds of resources between themselves and other individuals. Such tasks can, therefore, provide a window into the situations under which participants are and are not willing to increase others' welfare (i.e., give them more or less money).

One of the most famous of such human economic games is known as the Dictator game, a resource allocation task that takes place between two players, a proposer and a receiver (Henrich et al., 2005; Kahneman, Knetsch, & Thaler, 1986). The proposer is given control over the division of an endowment of money (e.g., \$10) that must be split between the two players. In cases in which the players do not know each other and play only once, a rational and self-interested proposer should keep all the money and give none to the receiver. However, humans tend not to behave in such a self-interested

way. Instead, proposers across many cultures tend to give at least some money to the recipient (Henrich et al., 2006), suggesting that humans seem to have preferences to behave *prosocially*, sometimes benefiting others' welfare at an immediate personal cost.²

Primate researchers have developed a set of non-verbal economic games that are conceptually similar to the human Dictator game, but are simplified in several ways (Burkart, Fehr, Efferson, & van Schaik, 2007; Jensen, Call, & Tomasello, 2007a; Jensen, Call, & Tomasello, 2007b; Jensen, Hare, Call, & Tomasello, 2006; Lakshminarayanan & Santos, 2008; Silk et al., 2005; Vonk et al., 2008; de Waal, Leimgruber, & Greenberg, 2008). First, in primate studies, the proposer is usually offered a dichotomous choice between two different resource distributions, making the cognitive demands of this task simpler than that of the human Dictator game. In addition, the primate choice task is often set up in such a way that there is no direct cost to acting prosocially: although the receiver's reward differs between the two options, the proposer typically obtains the same reward no matter which option is given to the recipient. Thus, acting prosocially in the nonhuman Dictator game is easier than in the human Dictator game in the sense that it does not demand that primates accept a cost to themselves to increase the welfare of another individual.

Despite the attempts to "stack the deck" in favor of finding prosocial tendencies in primates, the results using economic games have been mixed. Silk et al. (2005), for example, presented chimpanzees with a choice between pulling a handle that delivered one piece of food to the proposer and one piece to the receiver (a 1/1 option) and pulling a handle that delivered one piece of food to the proposer and nothing to the receiver (a 1/0 option). Chimpanzees showed no preference for pulling the handle that provided the receiver with food, choosing the 1/0 option as often as they chose the 1/1 option. Jensen and colleagues (2006) replicated this effect, and confirmed that chimpanzees understand the differences in payoff to the receiver position. They confirmed that chimpanzees understood the receiver payoff by including a condition in which the proposer had access to the receiver position, and finding that in this condition proposers chose the option that sent food to the receiver position. When another chimpanzee had access to the receiver position, they replicated the results from Silk et al. (2005), finding that chimpanzees acted with indifference to increasing the welfare of the receiver in both a 1/0 vs. 1/1 condition and a 0/0 vs. 0/1 condition. In a further

extension, Yamamoto and Tanaka (2010) found that chimpanzees' selfish focus is resistant not only to concerns of kin (in this case mother-offspring pairs) but also to the potential cooperation-inducing structure of reciprocation (in which chimpanzees took turns playing the proposer and receiver roles).

One initially plausible explanation for chimpanzee's selfish behavior in these tasks is that they are so distracted by their own rewards that they fail to notice how their behaviors influence the welfare of the receiver. To test this possibility, Vonk and colleagues (2008) provided chimpanzees with separate choices to deliver food to themselves and a receiver. Subjects were armed with a stick that could dislodge food set to roll toward the subject position and food set to roll toward the receiver position. Thus, the chimpanzees could dislodge and receive their own reward and then, no longer distracted, provide a reward to the receiver. Although chimpanzees reliably dislodged their own reward, they were not more likely to dislodge the receiver reward when a receiver was present as opposed to absent. This result suggests that, in the previous tasks, chimpanzee lack of concern for others was not due merely to distraction with their own reward.

Although chimpanzees, one of our closest evolutionary relatives, do not show human-like prosocial concerns in a dictator-style economic game, some more distantly related primates do (Burkart et al., 2007; de Waal et al., 2008; Lakshminarayanan & Santos, 2008; Takimoto, Kuroshima, & Fujita, 2009). Burkart and colleagues (2007) found that common marmosets, a New World monkey species, were more likely to pull a 0/1 option than a 0/0 option when a receiver was present than when the reward was delivered to an empty chamber. Similar results have been reported in another New World monkey, the brown capuchin (de Waal et al., 2008; Lakshminarayanan & Santos, 2008; Takimoto et al., 2009). Takimoto and colleagues (2009), for example, found that capuchins were willing to provide conspecifics with high payoffs, especially when recipients were visible and low ranking. Capuchin prosociality in these tasks is remarkably robust; Lakshminarayanan and Santos (2008) found that capuchin proposers provide another monkey with a reward that is greater than its own reward, willingly delivering a high-value treat (marshmallow) to a receiver even when they themselves got a low-value treat (cucumber). However, not all New World primates show prosocial preferences on this task. Cronin and colleagues found that cotton-top

tamarins are not more likely to provide a 1/1 outcome over a 1/0 outcome when a receiver is present (Cronin, Schroeder, Rothwell, Silk, & Snowdon, 2009, see also Stevens, 2010).

In sum, nonhuman performance on simple economic games provides evidence of prosociality in some species, but the failure of chimpanzees (among our closest evolutionary relatives) to behave prosocially indicates that the evolutionary history that leads humans to behave prosocially in a donation task may not be as simple as we originally thought. Before attempting to dissect the pattern any further, however, it will be useful to look at primates' performance on another measure of prosocial preferences.

Instrumental Helping Tasks in Humans and Primates

Humans also demonstrate a motivation to increase others' welfare, helping others achieve a variety of goals; helping a friend move their couch or handing someone an out-of-reach pen are situations in which people are willing to incur a slight cost to help someone achieve a goal and thus increase another's welfare. Do primates also show a willingness to increase others' welfare through instrumental helping? Several studies have examined this issue, observing that primates, like humans, are willing to act in ways that instrumentally help others. In contrast to their indifferent performance on other economic games (Jensen et al., 2007a; Silk et al., 2005; Vonk et al., 2008; Yamamoto & Tanaka, 2010), Warneken and Tomasello (2006) found that chimpanzees were willing to go out of their way to help humans and conspecifics achieve goals. Chimpanzees, for instance, are willing to help a person reach an out-of-reach object and also help a conspecific enter a room to obtain food. Indeed, chimpanzee subjects acted in ways that increased others' welfare even when it required them to take a slight cost, climbing into a raised raceway to retrieve the object (Warneken, Hare, Melis, Hanus, & Tomasello, 2007). Although chimpanzees failed to help in some types of tasks (removing physical obstacles and completing a failed action through either imitated or novel means), chimpanzees showed helping behavior in a number of situations. Such results suggest that this species possesses prosocial motivations that would fall under the definition of human-like harm concerns (Warneken et al., 2007).

Extending this line of work to other primate species, Barnes, Hill, Langer, Martinez, and Santos,

(2008) found that capuchins were minimally interested in helping a human retrieve an out-of-reach object, although this effect was not observed as robustly as in chimpanzees. When helping was difficult (the capuchin had to reach outside the cage to manipulate the object), capuchins showed almost no helping behavior. When helping was less costly, capuchins attended to a person's goals and exhibited higher rates of helping. Thus, in contrast to performance in economic games, in which capuchins show prosocial preferences and chimpanzees do not, instrumental helping tasks are more likely to elicit welfare concerns from chimpanzees than they are from capuchins.

Accounting for Species Differences in Harm and Care Behaviors

Taken as a whole, the results of these experiments indicate that primates show *some* prosocial motivations—and, therefore, behave in ways that are consistent with some harm-related moral concerns—but tend not to express such behaviors as consistently as humans. One plausible explanation for the lack of prosocial behavior in some experimental contexts is that aspects of the context in which researchers have tested these capacities hinder the expression of prosocial concerns. One issue, for example, concerns the ecological validity of the tests typically used to study these considerations—namely, food donation tasks (see also chapter 20 of this volume). In their natural behavior, primates rarely directly share with or donate food to other individuals in the way they are required to do in donation tasks. As such, donation tasks may represent a situation that primates are unlikely to encounter in the wild. Indeed, the species with the most robust evidence for prosocial preferences on donation tasks—the brown capuchin monkey—is also the one best known for its tolerated food sharing (de Waal, 2006). An additional issue concerns the researchers' ability to account for the subjects' expectations of the experimenters. In tasks that elicit helping behavior, certain populations may have stronger expectations that helping behavior may earn a treat. Under this view, what appears to be species differences in prosocial behavior may actually be population-level differences in performance that arise due to differences in rearing and reinforcement histories.

Researchers will likely profit from developing new tasks, ones that better mimic the kinds of situations in which primates might naturally express their own harm concerns. One such situation would involve testing primates on a more prototypical

harm violation, namely, the act of physically hurting someone (see Haidt & Joseph, 2007). Indeed, human infants understand such harm violations when they are only three months of age (Hamlin et al., 2007; Hamlin et al., in press). In addition, developmental psychologists have developed a set of nonverbal methods (e.g., looking measures: Hamlin et al., in press; choice measures: Hamlin et al., 2007) to test these intuitions that could be applied to nonverbal primates.

Another profitable next step in this area of research might involve investigating how primates *evaluate* individuals that violate norms against harm. Most of the work addressing primates' concerns has focused on exploring whether primates behave in ways that are consistent with prosocial norms. Few researchers have explored whether primates react negatively to those who act harmfully toward others. Recently, Subiaul, Vonk, Okamoto-Barth, and Barth (2008) found that chimpanzees will preferentially beg for food from a human that has been seen giving (rather than refusing to give) treats to either a human or another chimpanzee. Similarly, Russell, Call, and Dunbar (2008) also found that chimpanzees preferred humans seen being generous rather than stingy, while finding that bonobos, gorillas, and orangutans did not form preferences in this situation. Although these results hint that some primates may evaluate others' harmful actions, more work is clearly needed on this issue. In particular, additional methods would be useful for investigating primate responses to third-party harm evaluations. For example, would an individual behave differently on a prosocial food donation task after first witnessing the prospective recipient doing a harmful act toward another individual? If given a choice to donate food to one of two individuals, would an individual who obeyed harm norms be preferred over one who did not? By exploring these questions, researchers may gain new insight into not just how primates behave when dealing with potentially harmful situations but also how they evaluate others who do and do not act in the same ways. Such new tasks may also allow researchers to get around some of the task demands that plague other studies in this line of work.

Evolution of Fairness/Reciprocity Behaviors and Concerns

Behaviors in the fairness domain focus on equity among social partners. Haidt and colleagues hypothesize that fairness concerns evolved in part to cultivate and maintain mutually beneficial partnerships

with other cooperators by establishing a norm of equitable treatment (Haidt & Joseph, 2007). The prototypical fairness violations, then, occur in cases in which a social partner is cheated by not receiving an equal share. Under this definition, fairness concerns can include cases in which the unfairness has a direct negative effect on the self (what researchers have referred to as “disadvantageous inequity”), cases in which the self benefits from the unfair split (“advantageous inequity”), and cases in which the self is not involved (which we will refer to as “third-party inequity”). In any case of unfairness between two individuals, one is experiencing disadvantageous inequity and the other is experiencing advantageous inequity (and an uninvolved observer may notice and evaluate the third-party inequity).

Human Behavior in Fairness Tasks

There are many real-world examples to suggest that humans act on the basis of fairness considerations. Just as in the case of harm considerations, however, researchers have investigated people’s fairness norms empirically using a number of economic games that tap into the tendency to avoid inequity. One of the most famous methods involves an economic scenario known as the Ultimatum game. This game is identical to the Dictator game, except that the second player (the receiver) has the opportunity to reject the first player’s (the proposer’s) division of the payoff. If the receiver chooses to reject, both players get nothing. As in the Dictator game, people tend not to play this game in ways that maximize self-interest, which would involve the receiver accepting any nonzero offer. Instead, people tend to play based on considerations of fairness. Henrich et al. (2006), for example, tested participants across many different cultures and found that over half the players were willing to reject some nonzero amounts offered to them. Thus, people incur costs to avoid certain unfair situations.

Nonhuman Behavior in Fairness Tasks

Do primates share the fairness concerns of humans? The question of whether primates care about fairness has recently become a focus of research in primate cognition. As in the case of studying harm concerns, researchers have begun examining this issue by developing primate versions of human economic games that tap into inequity aversion. Jensen and colleagues (2007a), for example, developed a version of the ultimatum game for chimpanzees. In this experiment, a chimpanzee proposer and receiver were tested in adjacent enclosures. The proposer

was allowed to select one of two different reward distributions for themselves and the receiver. However, like the human ultimatum game, the proposer was able to obtain his payoffs only if the receiver chimpanzee “accepted” the offer—in this case by completing a pull required to bring the rewards into reach. In contrast to human performance, Jensen and colleagues found that chimpanzees rarely rejected low offers. Instead, chimpanzees behaved like rational payoff maximizers. Proposer chimpanzees chose the distribution that provided themselves selfishly with more food, and receiver chimpanzees only rejected in cases in which they were offered a payoff of zero. This result suggests that chimpanzees are unwilling to reject offers that are unfair, at least in cases when such rejections are costly.

Recent research has also examined whether primates exhibit negative reactions when they receive disadvantageous treatment relative to a conspecific. In an influential early study, Brosnan and de Waal (2003) allowed capuchin monkeys to trade tokens in exchange for cucumber slices. Despite being a low-value treat, the monkeys were almost always willing to trade for and eat the low-value treat when no other interactions were taking place. Rejections of the low-value treat rose to almost 50 percent, however, when the subject first saw another monkey receive a high-value grape treat. When another monkey received a grape without needing to trade a token, the rejection rates rose to about 80 percent. Similar results were reported for at least some populations of chimpanzees tested on an identical task (Brosnan, Schiff, & de Waal, 2005). These results were initially interpreted as evidence that some primates are willing to reject unfair offers, and thus that these species evaluate unfair payoffs by choosing not to trade with unfair experimenters.

The rich interpretation of Brosnan and colleagues’ findings has been challenged by several researchers. Dubreuil, Gentile, & Visalberghi (2006) argued that primates may reject low-valued foods on this task merely because they are frustrated at not getting a high-value reward. To test this, they examined whether capuchin monkeys would show similar rejection effects when preferred food was present but out of reach. They found that capuchins exhibit the same behaviors when the preferred food was given to another monkey as they do when the preferred food is simply out of reach. On the basis of these data, Dubreuil and colleagues (2006) argued that Brosnan and colleagues’ findings were due to simple frustration effects rather than “true” fairness concerns. Using a similar logic, Silberberg, Crescimbeni,

Addressi, Anderson, and Visalberghi (2009) performed a direct test of the frustration account of capuchin rejections. Their subjects did not get the high-value treat during previous trials and, therefore, they could not be frustrated at getting the low-value treat during test trials. Silberberg et al. (2009) found that capuchins accepted nearly every offered trade and showed no differences based on observations of high- or low-value treats being delivered to another monkey. Silberberg thus argued that previous research interpreted as evidence of fairness concerns in primates were the result of unaccounted for frustration effects. Other researchers have also reinterpreted Brosnan and colleagues' findings by arguing that these results reflect only how primates react when their expectations are violated. Chen and Santos, for example, have argued that capuchins may react negatively in Brosnan and colleagues' studies merely because they got less of a reward than they originally expected (see Chen and Santos, 2006, for a more detailed discussion of these issues). Finally, Bräuer, Call, and Tomasello (2006) attempted to replicate Brosnan and colleagues' (2005) chimpanzee version of the study with another group of chimpanzees. Their study observed a conflicting pattern of performance; chimpanzees in Bräuer and colleagues' study became more involved when experimenters delivered better rewards to other chimpanzees, becoming more motivated to beg for food when better food was present. Taken together, these conflicting results suggest that fairness concerns observed in primates are, at best, extremely fragile: they may or may not be elicited, depending on minor variations in experimental design.

Given the difficulty in establishing fairness concerns in cases of disadvantageous inequity, it is not surprising that there are no reports of primates reacting negatively to cases of advantageous inequity, in which the subject himself benefits from the unequal distribution of payoffs. Indeed, Brosnan (2006) anecdotally noted that monkeys who were unfairly paid the higher reward in the original Brosnan and de Waal (2003) study never spontaneously shared their own reward with the subject who received less. (In fact, they sometimes stole the subject monkey's rejected cucumber!). In this way, primates' fairness concerns seem to emerge only in cases of disadvantageous inequity; there is no evidence that primates show any fairness concerns in cases of advantageous inequity. Similarly, little work has explored whether primates also attend to fairness considerations in third party cases, cases in which they themselves are not directly involved.

Re-evaluating Primate's Performance in Fairness Studies

In sum, there is much less conclusive evidence regarding primate fairness concerns as compared to the evidence that primates possess concerns in the harm domain. To the extent that primates have any concerns about being personally cheated, they seem to lack a general fairness concern that would respond to *others* being cheated. Even when focusing on disadvantageous inequity aversion, there are suggestions that the observed effects may result more from general frustration than from a concern for equity.

It is, however, worth considering whether the general-frustration effect described earlier might be relevant for the evolution of fairness concerns. Put differently, a frustration effect *could* in fact be a critical part of the fairness response in humans. Humans, of course, tend to exercise fairness concerns quite specifically, only when dealing with distributions involving other social agents. Nonetheless, one could imagine how a simple frustration-based mechanism—one that responded negatively to *any* payoff that was lower than one expected—could perform in much the same way as a true fairness consideration—it too would allow an organism to respond negatively when it received a smaller-than-expected reward in interactions with other social agents. In this way, a simple mechanism that allows primates to treat *all* instances of smaller-than-expected rewards as “unfair” exchanges might be an evolutionarily advantageous one, if the benefits of displaying negative reactions to unfairness in social situations outweigh the costs of displaying those reactions in nonsocial settings. In this analysis, the original fairness concern, from an evolutionary perspective, may result from a more general frustration effect, exactly the kind of effect that many of the preceding experiments were designed to factor out. In this way, we argue that researchers may want to reinterpret some of the so-called frustration effects in the primate fairness experiments (Dubreuil et al., 2006; Silberberg et al., 2009) and think more critically about whether a general-frustration mechanism might be more relevant to fairness concerns than previously thought.

Evolution of Ingroup/Loyalty Behaviors and Concerns

Moral concerns in the third domain—that of ingroup allegiance—focus on productive cooperation within the group and appropriate reactions to challenges by other groups (Haidt & Joseph, 2007). Some standard moral violations in the ingroup domain would include breaking off contact with

your family or changing citizenship (Graham & Haidt, 2010). Although harm and fairness concepts are considered important in all human cultures, the moral significance of ingroup concerns varies a bit across cultures, often taking on more importance among non-Western than Western people (see Haidt and Joseph, 2007). Nevertheless, as we review in the next section, all human cultures appear to represent social groupings and favor ingroup members when interacting with others in their social world.

Human Ingroup Loyalty

A long research tradition in social psychology has focused on the ease with which humans form social groups and act in ways that favor their own over other groups (see reviews in Fiske, 1998; Sidanius, 1993; Tajfel & Turner, 1979). Much of the early work in this area observed that humans are naturally predisposed to see the world in terms of social groups, spontaneously segregating themselves into groups based on the most minimal of grouping dimensions. In a classic experiment, Tajfel (1970) had participants indicate their liking for various paintings and then told each participant that they belonged to a group of people who, on average, preferred the paintings of Klee or of Kandinsky. When participants subsequently played an economics game in which they could maximize payoffs to members of their own group at the expense of maximizing payoffs to everyone, participants showed a clear ingroup bias, selectively increasing the welfare of their own group members. This preference for one's ingroup is a bias known to emerge relatively early in human development. By only a few months of age, human infants have already developed preferences for individuals of their own race and native language (Bar-Haim, Ziv, Lamy & Hodes, 2006; Kelly et al., 2005; Kinzler, Dupoux & Spelke, 2007).

Nonhuman ingroup favoritism

Social groups also play an important part in the lives of many primates. Nearly all primates form social groups (Pusey et al., 1987). Primate groups are typically based on kin lines, but most primate groups also contain some unrelated individuals who also engage with each other as part of long-term social interactions. As in humans, primates often behave more nicely toward ingroup members than toward outgroup individuals. Most primate groups are characterized by high levels of intergroup aggression. For some species, such intergroup hostility can translate to harmful behaviors directed at outgroup members, such as physical aggression and vocal

conflict. In an extreme case, chimpanzee populations are known to take part in outgroup raids, entering neighboring communities to hunt down and kill members of other groups.

Only recently have researchers begun to explore what primates actually know about their own social groups, and whether they represent ingroup and outgroup members in ways that are similar to humans. Pokorny and de Waal (2009) examined whether capuchin monkeys could be trained to categorize members of ingroups and outgroups using a touch-screen task. They presented monkeys with an "odd-ball task" in which arrays of ingroup or outgroup monkey faces appeared on a touch screen. In this task, subjects must select the individual who is not a member of the same group as the others. Capuchins easily succeeded on this task, demonstrating that, with training, capuchins could learn to discriminate faces of ingroup and outgroup members. Mahajan et al. (2011) explored whether monkeys show a similar level of discrimination in the absence of training. They presented free-ranging rhesus macaques with faces of ingroup and outgroup members and explored which faces captured the monkeys' attention. The monkeys spontaneously discriminated ingroup from outgroup faces, looking longer at the outgroup member with increased vigilance. In addition, the monkeys' categorizations were likely not driven merely by familiarity; Mahajan and colleagues observed that monkeys showed more vigilance toward recent transfers out of the group (who are very familiar yet still outgroup) than recent transfers into the group (who are relatively unfamiliar yet newly ingroup). These results demonstrate that monkeys spontaneously categorize conspecifics as members of ingroups and outgroups, even in the absence of training. Moreover, these findings suggest that macaques may naturally devote more vigilance toward outgroup than ingroup individuals.

The work just described suggests that some primates spontaneously recognize ingroup members, but do primates have a *preference* for ingroup members, as would be consistent with some kinds of ingroup moral considerations? To explore this issue, Mahajan et al. (2011) developed a looking-time test of monkeys' attitudes toward different social groups. The logic of their attitude measure was similar to that of the famous implicit association test (IAT), which is often used to assess category associations in humans (Greenwald & Banaji, 1995; Greenwald, McGhee & Schwartz, 1998): subjects should habituate more quickly to sets of images that have consistent valence than to sets of images that

have inconsistent valence. In this way, Mahajan and colleagues could explore whether monkeys associated ingroup members with positively valenced stimuli and outgroup members with negatively valenced stimuli. They presented free-ranging macaques with sequences of pictures that alternated between monkey faces (either ingroup or outgroup members) and positive/negative objects. When the picture sequences alternated between ingroup faces and positive images or outgroup faces and negative images, monkeys habituated quickly. In contrast, when presented with inconsistent sequences of ingroup faces and negative images or outgroup faces and positive images, monkeys failed to habituate, suggesting that these sets of images do not have the same valence for monkeys. In this way, Mahajan et al.'s results suggest that monkeys naturally perceive ingroup members positively and outgroup members negatively, suggesting that monkeys share a human-like favoritism toward ingroup members and dislike toward outgroup members.

Do primates' differential attitudes toward ingroup and outgroup members translate into moral intuitions about how these different groups should be treated? Do primates have preferences toward selectively increasing the welfare of ingroup members? Do they negatively evaluate those who do not show favoritism toward the group? Little work has addressed these issues directly, but one hint comes from a recent experimental economic study by de Waal and colleagues (2008). Frans de Waal and colleagues presented capuchin monkeys with an experimental economic task in which proposers could donate food to other recipient monkeys. When the recipient was an ingroup member, capuchin monkeys reliably chose to donate the prosocial option (see Lakshminarayanan and Santos, 2005 for similar results in this species). However, de Waal and colleagues observed a different pattern of performance when the recipient was an outgroup monkey. When proposers had the option to deliver a piece of food to a monkey from a different group, their performance fell to chance; proposers were indifferent to outgroup monkeys' welfare. These results provide an important hint that capuchins' prosocial motivations may be specific to ingroup members, consistent with the view that an individual's group may affect the extent to which they are seen as part of the moral circle.

The preceding results are consistent with the idea that the human ingroup moral domain could be shared with other primates. Unfortunately, however, there are still many questions regarding ingroup

considerations in primates. First, there appear to be several important ways in which primate intergroup cognition may diverge from that of humans. For humans, group identity, even when trivially determined, can be a critical part of self-construal. Given the limitations on nonhuman self-perception and identity formation, such elements would likely be missing from any nonhuman ingroup cognition. Second, human groups have several unique mechanisms for promoting group affiliation that primate groups lack, such as linguistic labels for different social groups. Finally, human ingroup concerns go beyond personal preferences to act positively toward the ingroup—humans also evaluate others based on their own actions toward ingroup and outgroup members. For example, third-party punishers are much more aggressive toward norm violators that harm an ingroup (rather than an outgroup) member (Bernhard, Fischbacher & Fehr, 2006). Would nonhuman primates positively evaluate individuals who promote their ingroup and negatively evaluate those who violate ingroup solidarity? Again, answers to these questions will allow us to better understand the extent to which primates' ingroup preferences actually map onto the kinds of moral concerns that humans experience when dealing with the domain of ingroup allegiance.

Evolution of Authority/Respect Behaviors and Concerns

Moral concerns in the fourth domain—the domain of authority—focus on negotiating the hierarchy within one's group, so that conflict can be avoided, by subordinates showing deference to superiors (Haidt & Joseph, 2007). Thus, harm against political leaders or elder family members are in violation of the authority domain (Graham & Haidt, 2010). Whereas the ingroup domain focused on favoritism toward one's own group over others (an attitude that would be important during, for example, intergroup conflict), the hierarchy domain focuses on respecting the divisions within a group.

There is much evidence that primates attend to authority issues, at least in their natural behavior. Many primate groups exhibit dominance hierarchies, a stable hierarchical pattern of dominance that dictates many aspects of an individual primate's daily behaviors. High-ranking individuals enjoy a number of privileges in primate groups—they often have first access to food, mating, and grooming opportunities, as well as other valued resources (e.g., Saunders & Hausfater, 1988). In this way, primates' natural social behavior appears to follow authority-related

constraints—lower-ranking primates tend to toe the line, allowing preferential resource access to individuals who outrank them.

There is also a growing body of research suggesting that primates carefully represent and attend to the order of individuals within their own hierarchy. Using an observational approach, Silk (1999) found that male bonnet macaques selectively recruited alliance members that outranked not only themselves but also their opponent, suggesting that recruiting macaques have some knowledge of their own and others' relative positions in the hierarchy. Similarly, Slocombe & Zuberbühler (2007) found that chimpanzees produce more pronounced screams when an individual of higher rank than their aggressor is present than when only lower-ranked individuals are present. Such decisions require third-party knowledge of rank relationships between others, rather than merely first-party knowledge of one's own rank relationship to others. To test this possibility using an experimental approach, Cheney and colleagues (Cheney, Seyfarth, & Silk, 1995) examined, using playback methods, whether baboons form expectations about an individual's position in the dominance hierarchy. They capitalized on a natural feature of vocal communication between dominant and subordinate female baboons: dominant female baboons tend to address subordinate females with grunts, but subordinate females tend to respond with a different vocalization, the fear bark. Cheney and colleagues recorded individuals making these vocalizations and then played baboon subjects sequences that were either expected (dominant individual grunted to a subordinate) or unexpected (dominant individual fear barked to a subordinate) based on the hierarchy. They observed that baboons responded more strongly to the unexpected rather than expected sequences, demonstrating the female baboons represent the relative ranks of other members of their group. In another study, Kitchen, Cheney, and Seyfarth (2005) found that male baboons responded more strongly to playbacks that simulated vocal contests between individuals ranked far apart as opposed to close to each other in the group hierarchy, suggesting that they recognized which individuals should and should not be engaging in dominance contests. Taken together, this work suggested that primates may represent the order of other individuals in their hierarchy and expect such individuals to interact in prescribed ways.

Although there is much evidence to suggest that primates recognize the rank ordering of individuals in their group, there is less direct evidence that

primates obey authority norms, such as deferring to the knowledge or skills of a high-ranking individual. One hint, however, comes from a study of primate social preferences. Deaner and colleagues (Deaner, Khera, & Platt, 2005) presented macaque monkeys with a choice between obtaining different amounts of juice and the opportunity to look at different images of other monkeys. Monkeys were willing to give up juice rewards to see images of high-ranking individuals, but had to be paid in juice to stare at images of low-ranking monkeys. This result suggests that monkeys value the opportunity to interact with high-ranking individuals, even when such interactions involve merely observing images of individuals in authority. Another hint comes from a recent study by Horner and colleagues (Horner, Proctor, Bonnie, Whiten, & de Waal, 2010). In this study, researchers allowed chimpanzees to socially learn how to obtain food by placing tokens in one of two containers. Observer chimpanzees saw two different demonstrators: a high-ranking individual who was trained to use one container and a low-ranking individual trained to use the other. When later given the chance to put tokens in the containers themselves, the observers were more likely to copy the high-ranking individual than the low-ranking individual. This result provides a hint that chimpanzees may selectively follow the behaviors of individuals in authority. Unfortunately, this result remains silent regarding the moral implications of this behavior. Would, for example, chimpanzees punish others who chose to copy the low-ranking individual?

In summary, then, there are many hints that primates are likely to have moral concerns relevant to the domain of authority, but little work has addressed this possibility directly. Nevertheless, the empirical stage is set for just such an investigation—researchers now have methods in place to examine who primates choose to interact with, how they choose to distribute rewards, and so on. We predict that using these techniques to examine primates' intuitions in the authority domain will be fruitful. Indeed, the available literature suggests that the domain of hierarchy concerns is likely to be one in which primates may show strong moral concerns.

Evolution of Purity/Sanctity Behaviors and Concerns

Behaviors in the fifth and final domain, the purity domain, focus on avoiding contaminants. The original targets of such behaviors were avoiding ingestion of physical contaminants, but the emotional reaction associated with violations in this

domain—our disgust reaction—has since expanded to include more symbolic contaminants, such as contact with unsavory individuals or behavioral practices (Haidt & Joseph, 2007). Thus, surgically adding a tail or getting blood from a child molester elicit disgust and also constitute violations of our purity norms (Graham & Haidt, 2010).

Rozin, Haidt, and Fincher (2009) hypothesized that natural selection may have shaped our original purity response—which at first was specific to physical contamination—in ways that allowed it to incorporate symbolic purity violations. Rozin and colleagues propose that the original physiological distaste reaction was initially elicited directly by certain tastes (e.g., bitterness) that indicated food that should be avoided. Once such a distaste reaction existed, it could be harnessed by a later-developing disgust-evaluation system that could be used to evaluate more cognitively elaborated inputs (e.g., to avoid even putting certain types of food in the mouth). Finally, they argue that the system may have been further leveraged by a later-developing moral-evaluation system, one that required other contaminating actions or people (especially those that remind us of our animal nature) to be avoided.

Primates, like many nonhuman animals, have the physiological distaste response that Rozin and colleagues hypothesize led to more complex purity norms. More impressively, many animals have the ability to learn distaste for novel foods. Garcia, Hankins, and Rusiniak (1974) report on a wide range of experiments on learned food aversion that include multiple species and methods for causing illness, including the famous example of rats avoiding foods they were introduced to shortly before being induced to feel nausea from radiation. This distaste reaction does not, however, seem to be elaborated into a *disgust reaction*. Indeed, Rozin and colleagues have proposed that distaste is shared across animals, but true disgust may be unique to humans.

Given that disgust reactions may be unique to humans, it is reasonable to assume that moral considerations in the purity domain may also be limited to humans. This leads to the prediction that primates should not show moral considerations related to issues of purity. To our knowledge, little work has explored this issue directly. Would primates, like humans, prefer not to interact with stimuli previously associated with a low-ranking individual, or something associated with a conspecific who has committed an immoral act? Do they, like humans, negatively evaluate individuals who engage in

impure actions, such as eating contaminated food? The answers to these questions would seem to be no, but little empirical work has directly addressed this issue.

Before ending our review of work in the purity domain, it's worth noting that the absence of disgust in primates may have broader implications for the evolution of moral cognition. In humans, the moral emotion of disgust can be leveraged to strengthen moral reactions that extend far beyond purity and into other domains. A growing body of work in human moral cognition has led to the view that eliciting disgust can be used to strengthen people's reactions to other non-purity-based moral violations. For example, Moretti and di Pellegrino (2010) found that experimentally induced disgust increased rejections of unfair offers during an ultimatum game. Importantly, the negative emotion of sadness did not have this effect. Thus, disgust (and not just negative emotion) seems to increase disadvantageous inequity aversion—a fairness-based moral consideration. This result is mirrored by studies that show disgust-related physiological reactions to unfairness, specifically activation of similar muscles in the face (Chapman, Kim, Susskind, & Anderson, 2009). If disgust serves to increase reactions across all five domains of morality, then the lack of disgust in nonhumans may partially account for many of the species differences in moral behaviors in the other four domains as well.

Conclusions

In this chapter, we have explored the evolution of moral concerns in five domains, focusing on the extent to which these concerns are present in primates. Clear evidence exists for primate harm concerns, though harm concerns do not show up as consistently in primates as they do in humans. The evidence for primate fairness concerns is less conclusive, though we argue in favor of some degree of fairness concerns, perhaps based on a more general frustration response. There is far less evidence regarding primates' concerns in the moral domains of ingroup or authority, though many primates behave in ways that appear to be consistent with such concerns. Finally, we believe there is little reason to suspect that primates will engage in moral evaluations similar to those of human in the purity domain.

Thus, there are a few clear conclusions from our review regarding primate moral cognition. First, some of our moral concerns (e.g., against harm) are present, to some degree, in primates. Second, many

of the cognitions that underlie our moral concerns (e.g., preference for our ingroup) are also present in primates. Similar to the state of morality research with humans, there is a critical lack of primate research in the moral domains of ingroup and hierarchy. Third, there appear to be several features of human cognition that allow our species to reason about moral considerations in ways that are not shared with other species. We have already discussed the role of an elaborated disgust response, but other obvious candidates are human language and the cognitive abilities it entails (Spelke, 2003; Tse, 2008) and human culture and learning within cultures (Chater, Vlaev & Grinberg, 2008; Henrich et al., 2010). Without disgust, language, and culture, primates may lack the capacity to introduce novel moral norms via the process of moralization (Rozin, 1999).

The goal of this chapter was to provide a current snapshot of what is known in the area of primate moral cognition using a framework based on human moral cognition work. Although our review revealed numerous open questions and areas in need of empirical work, it also reviewed a variety of new methodical techniques that can readily be applied to address these open issues. Thus, we look forward to the upcoming years of empirical work in this area. Such new work will provide further insight into the nature of primates' moral considerations and new hints into the evolutionary origins of our own moral capacities.

Future Directions

Throughout the chapter, we focused on two kinds of evidence for primate moral considerations—first, whether primates themselves behaved in ways that obeyed different moral norms, and second, whether they evaluated others who failed to obey these norms. To date, researchers have mostly focused on the former kind of evidence. In the harm domain, for example, researchers have devoted considerable effort to examining whether primates act in ways that increase others' welfare, but little work has investigated how primates feel about other individuals when they violate these norms. Turning to the question of primates' evaluations in this domain may provide some insight into the confusing pattern of findings regarding how different primate species evaluate the welfare of others. In the domain of fairness, some work has begun to explore how primates react to others when they violate fairness norms (e.g., Brosnan and de Waal, 2003), but nearly all of this work to date has focused on cases of disadvantageous norm violation, in which the subject loses

out because of the violation. To better understand the nature of primates' considerations, researchers would profit from exploring cases of advantageous norm violations, when the violation benefits the subject, and cases of third-party violations, when the subject is not involved. It is possible that primates lack moral concerns in cases in which they themselves are not negatively affected. On the other hand, it could be that researchers simply do not yet have evidence of such unselfish considerations.

Another current weakness in the literature on primate moral considerations stems from the disproportionate focus on certain domains. As reviewed here, nearly all the work on primate moral cognition to date has focused on harm and fairness violations. Such a biased focus makes some sense, given that the same bias appears to have played out in research on human moral cognition, where most work has also focused on harm and fairness domains (see Haidt & Joseph, 2007). However, in the case of primates, the bias toward studying only the harm and fairness domains winds up being especially unfortunate, because primates' natural behavior suggests they may have well-formulated norms in the domains of ingroup and authority. Future work, therefore, should focus far more on these domains, exploring whether primates do, in fact, make moral evaluations when they see violations in these areas.

Endnotes

1. Incidentally, work on *human* moral cognition has also disproportionately focused on the harm and fairness domains (see Haidt & Joseph, 2007).
2. The Dictator game can also be thought of as a measure of fairness considerations, specifically as a test of how and when participants decide to give others a fair split. However, given that this test has mostly been discussed in terms of prosocial preferences in the primate field, we've presented this work as part of the harm domain (i.e., as a measure of how and when primates are willing to increase others' welfare).

References

- Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science: A Journal of the American Psychological Society / APS*, *17*, 159–163.
- Barnes, J. L., Hill, T., Langer, M., Martinez, M., & Santos, L. R. (2008). Helping behaviour and regard for others in capuchin monkeys (*cebus apella*). *Biology Letters*, *4*, 638–640.
- Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*, 912–915.
- Bräuer, J., Call, J., & Tomasello, M. (2006). Are apes really inequity averse? *Proceedings of the Royal Society of London, Biology*, *273*, 3123–3128.
- Brosnan, S. F. (2006). Nonhuman species' reactions to inequity and their implications for fairness. *Social Justice Research*, *19*, 153–185.

- Brosnan, S. F., & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, *425*, 297–299.
- Brosnan, S. F., Schiff, H. C., & de Waal, F. B. (2005). Tolerance for inequity may increase with social closeness in chimpanzees. *Proceedings of the Royal Society of London, Biology*, *272*, 253–258.
- Burkart, J. M., Fehr, E., Efferson, C., & van Schaik, C. P. (2007). Other-regarding preferences in a non-human primate: Common marmosets provision food altruistically. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 19762–19766.
- Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). ‘Unwilling’ versus ‘unable’: Chimpanzees’ understanding of human intentional action. *Developmental Science*, *7*, 488–498.
- Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*, 1222–1226.
- Chater, N., Vlaev, I., & Grinberg, M. (2008). A new consequence of Simpson’s paradox: Stable cooperation in one-shot prisoner’s dilemma from populations of individualistic learners. *Journal of Experimental Psychology: General*, *137*, 403–421.
- Chen, M. K., & Santos, L. R. (2006). Some thoughts on the adaptive function of inequity aversion: An alternative to Brosnan’s social hypothesis. *Social Justice Research*, *19*, 201–207.
- Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The responses of female baboons (*papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning? *Journal of Comparative Psychology*, *109*, 134–141.
- Cronin, K. A., Schroeder, K. K., Rothwell, E. S., Silk, J. B., & Snowdon, C. T. (2009). Cooperatively breeding cottontop tamarins (*saguinus oedipus*) do not donate rewards to their long-term mates. *Journal of Comparative Psychology*, *123*, 231–241.
- Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. *Current Biology*, *15*, 543–548.
- de Waal, F. B., Leimgruber, K., & Greenberg, A. R. (2008). Giving is self-rewarding for monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 13685–13689.
- de Waal, F. B. M. (2006). Joint ventures require joint payoffs: Fairness among primates. *Social Research: An International Quarterly*, *73*, 349–364.
- Dubreuil, D., Gentile, M. S., & Visalberghi, E. (2006). Are capuchin monkeys (*cebus apella*) inequity averse? *Proceedings of the Royal Society of London, Biology*, *273*, 1223–1228.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology*, Vol. 2 (4th ed., pp. 357–411). New York: McGraw-Hill.
- Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, *15*, 447–452.
- Garcia, J., Hankins, W. G., & Rusiniak, K. W. (1974). Behavioral regulation of the milieu interne in man and rat. *Science*, *185*, 824–831.
- Graham, J., & Haidt, J. (2010). *Moral foundations sacredness scale*. Retrieved July 2, 2010, from http://faculty.virginia.edu/haidt/mft/MFT_sacredness_scale.doc.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*, 1029–1046.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.
- Haidt, J. (2008). Morality. *Perspectives on Psychological Science*, *3*, 65–72.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*, 98–116.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*, 55–66.
- Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The Innate Mind*, *3*, 367–391.
- Haidt, J., & Kesebir, S. (2009). Morality. In S. T. Fiske & D. Gilbert (Eds.), *The handbook of social psychology* (5th ed.). New Jersey: John Wiley & Sons.
- Hamilton, W. D. (1963). The evolution of altruistic behavior. *The American Naturalist*, *97*, 354–356.
- Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*, 557–559.
- Hare, B., Addessi, E., Call, J., Tomasello, M., & Visalberghi, E. (2003). Do capuchin monkeys, *cebus apella*, know what conspecifics do and do not see? *Animal Behaviour*, *65*, 131–142.
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*, 771–785.
- Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*, 139–151.
- Hauser, M. D. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Harper Collins.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2005). “Economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *The Behavioral and Brain Sciences*, *28*, 795–815, discussion 815–855.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*, 1480–1484.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al. (2006). Costly punishment across human societies. *Science*, *312*, 1767–1770.
- Horner, V., Proctor, D., Bonnie, K. E., Whiten, A., & de Waal, F. B. (2010). Prestige affects cultural learning in chimpanzees. *PLoS ONE*, *5*, e10625.
- Jensen, K., Call, J., & Tomasello, M. (2007a). Chimpanzees are rational maximizers in an ultimatum game. *Science*, *318*, 107–109.
- Jensen, K., Call, J., & Tomasello, M. (2007b). Chimpanzees are vengeful but not spiteful. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 13046–13050.

- Jensen, K., Hare, B., Call, J., & Tomasello, M. (2006). What's in it for me? Self-Regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society of London, Biology* 273, 1013–1021.
- Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 76, 728–741.
- Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109, 224–234.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., Ge, L., & Pascalis, O. (2005). Three-month-olds, but not newborns, prefer own-race faces. *Developmental Science*, 8, F31–F36.
- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, 104, 12577.
- Kitchen, D. M., Cheney, D. L., & Seyfarth, R. M. (2005). Male chacma baboons (*papio hamadryas ursinus*) discriminate loud call contests between rivals of different relative ranks. *Animal Cognition*, 8, 1–6.
- Lakshminarayanan, V. R., & Santos, L. R. (2008). Capuchin monkeys are sensitive to others' welfare. *Current Biology*, 18, R999–1000.
- Mahajan, N., Martinez, M. A., Gutierrez, N. L., Diesendruck, G., Banaji, M. R., & Santos, L. R. (2011). The evolution of intergroup bias: Perceptions and attitudes in rhesus macaques. *Journal of Personality and Social Psychology*, 100, 387–405.
- Masserman, J. H., Wechkin, S., & Terris, W. (1964). "Altruistic" behavior in rhesus monkeys. *The American Journal of Psychiatry*, 121, 584–585.
- Moretti, L., & di Pellegrino, G. (2010). Disgust selectively modulates reciprocal fairness in economic interactions. *Emotion*, 10, 169–180.
- Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 362, 731–744.
- Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). 'Unwilling' versus 'unable': Capuchin monkeys' (*cebus apella*) understanding of human intentional action. *Developmental Science*, 12, 938–945.
- Pokorny, J. J., & de Waal, F. (2009). Monkeys recognize the faces of group mates in photographs. *Proceedings of the National Academy of Sciences*, 106, 21539.
- Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, 7, 157–160.
- Pusey, A. E., Packer, C., Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W., & Struhsaker, T. T. (1987). *Primate societies*. Chicago: University of Chicago Press.
- Rosati, A., Santos, L. R., & Hare, B. A. (2009). Primate social cognition: Thirty years after Premack and Woodruff. In M. Platt & A. A. Ghazanfar (Eds.), *Primate neuroethology*. Cambridge, MA: MIT Press.
- Rozin, P. (1999). The process of moralization. *Psychological Science*, 10, 218–221.
- Rozin, P., Haidt, J., & Fincher, K. (2009). From oral to moral. *Science*, 323, 1179.
- Russell, Y. I., Call, J., & Dunbar, R. I. (2008). Image scoring in great apes. *Behavioural Processes*, 78, 108–111.
- Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, macaca mulatta, know what others can and cannot hear. *Animal Behaviour*, 71, 1175–1181.
- Saunders, C. D., & Hausfater, G. (1988). The functional significance of baboon grooming behavior. In Colbern D. L., & Gispén W. H. (Eds.), *Neural mechanisms and biological significance of grooming behavior* (pp. 430–432). New York: Academic Science.
- Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health*. New York: Routledge.
- Sidanius, J. (1993). The psychology of group conflict and the dynamics of oppression: A social dominance perspective. In S. Iyengar & W. McGuire (Eds.), *Explorations in political psychology* (pp.183–219). Durham, NC: Duke University Press.
- Silberberg, A., Crescimbeni, L., Addessi, E., Anderson, J. R., & Visalberghi, E. (2009). Does inequity aversion depend on a frustration effect? A test with capuchin monkeys (*cebus apella*). *Animal Cognition*, 12, 505–509.
- Silk, J. B. (1999). Male bonnet macaques use information about third-party rank relationships to recruit allies. *Animal Behaviour*, 58, 45–51.
- Silk, J. B., Brosnan, S. F., Vonk, J., Henrich, J., Povinelli, D. J., Richardson, A. S., Lambeth, S. P., Mascaró, J., & Schapiro, S. J. (2005). Chimpanzees are indifferent to the welfare of unrelated group members. *Nature*, 437, 1357–1359.
- Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, 104, 17228.
- Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. *Language in mind: Advances in the study of language and thought*, 277–311.
- Stevens, J. R. (2010). Donor payoffs and other-regarding preferences in cotton-top tamarins (*saguinus oedipus*). *Animal Cognition*, 13, 663–670.
- Subiaul, F., Vonk, J., Okamoto-Barth, S., & Barth, J. (2008). Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Animal Cognition*, 11, 611–623.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 102.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worche & W. G. Austins (Eds.), *The social psychology of intergroup relations*. (pp. 33–47). Monterey, CA: Brooks/Cole Publishing Co.
- Takimoto, A., Kuroshima, H., & Fujita, K. (2009). Capuchin monkeys (*cebus apella*) are sensitive to others' reward: An experimental analysis of food-choice for conspecifics. *Animal Cognition*, 13, 249–261.
- Tomasello, M., Call, J., & Hare, B. (2003a). Chimpanzees understand psychological states: The question is which ones and to what extent. *Trends in Cognitive Sciences*, 7, 153–156.
- Tomasello, M., Call, J., & Hare, B. (2003b). Chimpanzees versus humans: It's not that simple. *Trends in Cognitive Science*, 7, 239–240.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences*, 28, 675–691.

- Tse, P. U. (2008). Symbolic thought and the evolution of human morality. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 1*. Cambridge, MA: MIT Press.
- Vonk, J., Brosnan, S. F., Silk, J. B., Henrich, J., Richardson, A. S., Lambeth, S. P., Schapiro, S. J., & Povinelli, D. J. (2008). Chimpanzees do not take advantage of very low cost opportunities to deliver food to unrelated group members. *Animal Behaviour*, *75*, 1757–1770.
- Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*, 1301–1303.
- Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *Plos Biology*, *5*, 1414–1420.
- Yamamoto, S., & Tanaka, M. (2010). The influence of kin relationship and reciprocal context on chimpanzees' other-regarding preferences. *Animal Behaviour*, *79*, 595–602.