

**The Humane Society Institute for Science and Policy**  
**Animal Studies Repository**

2013

# Ape Autonomy? Social Norms and Moral Agency in Other Species

Kristin Andrews  
*York University*

Follow this and additional works at: <http://animalstudiesrepository.org/autono>

 Part of the [Animal Studies Commons](#), [Comparative Psychology Commons](#), and the [Other Animal Sciences Commons](#)

---

## Recommended Citation

Andrews, K. (2013). Ape autonomy? Social norms and moral agency in other species. *Animal minds & animal ethics: connecting two separate fields*, 173-196.

This Book Chapter is brought to you for free and open access by the Humane Society Institute for Science and Policy. It has been accepted for inclusion by an authorized administrator of the Animal Studies Repository. For more information, please contact [eyahner@humanesociety.org](mailto:eyahner@humanesociety.org).

# Ape Autonomy? Social Norms and Moral Agency in Other Species

Kristin Andrews  
York University

## Introduction

Once upon a time, not too long ago, the question about apes and ethics had to do with moral standing—do apes have interests or rights that humans ought to respect? Given the fifty years of research on great ape cognition, life history, social organization, and behavior, the answer to that question seems obvious. Apes have emotions and projects, they can be harmed, and they have important social relationships.

Today the question is more likely to be about whether apes have morality in some sense—whether they are autonomous agents, have social norms, moral emotions, or even abstract concepts of right and wrong. Some argue that given what we now know, the great apes do fall into the sphere of what we might call moral agents (de Waal 2006; Bekoff & Pierce 2009). From field research, we know that within species, great ape communities differ from one another in their methods of food processing and social interaction to an extent that is recognized as cultural (McGrew 1992; Whiten et al. 1999; van Schaik 2003). And from captive research we know that great apes can recognize intentions and goals (Uller 2004; Call et al. 2004; Warneken & Tomasello 2006) as well as others' emotional states (Parr and Hopkins 2000; Parr 2001). There is evidence of prosocial behavior in chimpanzees including consolation behavior (comforting individuals who are subject to aggressive attacks) (Fraser et al 2008; Fraser and Aureli 2008) and helping behavior (Hirata 2003; Melis et al. 2006). Indeed, de Waal argues that great apes are capable of empathy, act to help others in flexible and creative ways, and engage in reciprocal exchanges of goods and services (de Waal 2006; 2009).

There have been two approaches to endorsing moral agency in nonhuman animals. One approach, taken by de Waal, is to see moral agency as on a continuum, so that he sees a smooth continuity between moral and amoral agents, just as he sees a smooth continuity between human and nonhuman minds. This allows him to conclude that many species have morality to some degree, since he takes empathy and reciprocity to be jointly sufficient for entry into the domain of moral agent, and he thinks many species demonstrate primitive versions of these requirements. De Waal concludes that many animals have a moral sense, though one that is not as developed as the human moral sense.

The other approach is to be a pluralist about kinds of moral agency, and argue that an animal who lacks many of the cognitive capacities of adult humans can still be a moral agent because there are different kinds of moral agents, and animal species can have their own form of morality. This view is defended by Bekoff and Pierce (2009) who argue that some species have a distinct form of morality that is not a precursor to human morality. Because they take *morality* to mean "a suite of other-regarding behaviors that cultivate and regulate complex interactions within social groups" (Bekoff & Pierce 2009, 82), they take the complexity of animal behavior, social organization, and cognitive flexibility to demonstrate that other species have morality in this sense. Central to the view is that different species have different norms, and that this makes animal morality species-relative. Despite their differences, important similarities between species include the capacities for empathy, altruism, cooperation and perhaps a sense of fairness. Unlike de Waal's focus on the mechanisms of morality, Bekoff and Pierce are more

interested in the existence of those beneficial behavioral regularities, which they identify functionally in terms of the benefits offered to the group.

However, there are certain theoretical barriers toward considering either de Waal's empathy and reciprocity or Bekoff and Pierce's other-regarding behaviors as sufficient criteria for moral agency. A moral agent is someone who can be held responsible for her behavior, whose behavior can appropriately be judged morally acceptable or not. This suggests that there is something cognitive about morality that isn't reflected in Bekoff and Pierce's approach. But the mechanisms that de Waal suggest may not be sufficient. Consider the following passage from the entry on moral agents in *The Routledge Encyclopedia of Philosophy*:

*Moral agents are those agents expected to meet the demands of morality. Not all agents are moral agents. Young children and animals, being capable of performing actions, may be agents in the way that stones, plants and cars are not. But though they are agents they are not automatically considered moral agents. For a moral agent must also be capable of conforming to at least some of the demands of morality (Haksar 1998).*

Here the question shifts from the nature of moral agency to the question of what counts as a demand for morality. Of course answers to that question will vary, because determining a demand of morality can only be answered from within the framework of a moral theory, and different moral theories will provide different answers. Subsequently, different moral theories will draw different conclusions about which individuals are included as moral agents. For example, it is plausible that a psychopath that is impaired in empathic understanding may not be considered a moral agent in the Humean sense, but could still be a moral agent in a Kantian framework, so long as his rational abilities are not impaired. Principlist views might require knowledge of moral rules, characterological views might require knowledge of moral valiances associated with personality traits, and particularist theories will stress the importance of procedural knowledge and skill in recognizing moral saliences of situations. Depending on one's moral theory, a demand of morality might include the ability to follow the Categorical Imperative, the ability to compute utility, the ability to emotionally relate to others, or the ability to understand and develop the virtues. And while it's true that all these theories share the view that a moral agent must be able to evaluate the actions of others, they differ when it comes to what counts as a moral evaluation, and thus the cognitive capacities required for meeting the demands of morality would vary.

This disagreement accentuates the burden for those who want to argue that animals do have morality. Since the null hypothesis is that animals are not moral agents, it will require more work to defend the view that nonhumans have morality in some sense. The additional burden is due to the fact that there is no consensus on the properties required to be a moral agent. As a result, any argument for moral agency can be deflected by claiming that the properties identified in a species are irrelevant or not sufficient for moral agency. The lack of consensus on moral theory also makes it easier to find direct arguments against the claim that animals have morality. That's because one can find one property associated with some account of moral agency and argue that animals lack that property, and hence cannot be a moral agent.

For example, the recent discussions of animal morality that have focused on the ability to understand the moral sentiments have been criticized not on empirical grounds, but on philosophical ones (e.g. Wright 2006, Korsgaard 2006, Kitcher 2006, Singer 2006). Frans de Waal argues that great apes, as well as dolphins and elephants, demonstrate empathy in behaviors such as targeted helping, consolation, cooperation, and sensitivity to fairness (de Waal 2006). However, the critics argue that these behaviors do not count as moral behaviors, because humans have some capacity that the animals lack. That is, while animals might do the right sort of thing, they may not do it for the right reasons (if they are even

acting for reasons at all). It isn't the behavior that matters, but the mechanisms that drive the behavior. The issue then becomes the sort of capacity required to make the moral-looking behavior into truly moral behavior.

I will look at one capacity that is sometimes thought to be a necessary condition for moral agency—a theory of mind. A theory of mind is the ability to understand that self and others act for reasons, and that propositional attitudes drive behavior. As we will see, there are several reasons for thinking that this capacity is required for moral agency. However, I will argue that this concern should not lead one to conclude that great apes cannot be moral agents, even for those who think it unlikely that any non-human species has a theory of mind. While there is no current evidence that great apes have anything like the philosopher's representational concept of belief (Call and Tomasello 2008), this is no reason to reject the possibility that they are moral agents. My aim is to undermine in-principle arguments against animal moral agency by showing that moral agency is possible without a theory of mind. I will argue that great apes have other cognitive capacities that can fulfill the same functions that are sometimes seen as requirements for moral agency, such as knowing the likely consequences of an action by predicting the participants' behaviors and emotional responses. This argument is based on my work on the evolution of mindreading (understood narrowly as the ability to attribute representational mental states such as belief), according to which the drive to explain behavior is fundamental to the development of theory of mind (Andrews 2009, 2012). The need to explain behavior is, in turn, driven by a recognition of the behavior as norm-violating. If recognizing social norms drives the development of metacognitive social abilities such as theory of mind, and the existence of social norms implies the existence of moral agency, then theory of mind is not a necessary condition for moral agency.

### **Moral behavior and theory of mind**

With a theory of mind comes the concepts of belief and desire, and with those concepts come the understanding that people act for reasons that consist of their beliefs and desires. An impoverished mother chooses to forgo eating nutritious food herself because she believes she cannot afford nutritious food for both her and her children, she wants her children to be healthy, and she believes they need nutritious food for healthy development. These days, some people are quick to make judgments about people who subsist on fast food, and might call into question the woman's fast-food-eating behavior. But those same people, once they realize the woman's reason for her food choice, would likely change their mind about her behavior as morally suspect, and perhaps even see her as a tragic hero. Knowing another's reason for action allows us to make a better-informed judgment about the behavior in question.

A full-fledged moral agent will be someone whose behavior can be seen as moral or immoral, and someone who can judge her own action and others' actions through a normative lens. At minimum, she must be an intentional agent who purposefully affects change in her environment. On this minimal conception, two things must hold: (1) the individual must be autonomous; (2) the individual must understand the consequences of her actions. For each condition, having a theory of mind seems to be implicated: insofar as autonomy requires knowing someone's reasons for action, it would require a theory of mind; since knowing the consequences of an action requires being able to predict behavior, and having a theory of mind has been touted as extremely beneficial in predicting behavior, theory of mind seems to be involved here as well.

One response to such worries is to argue that great apes do have a theory of mind, and hence these requirements do not interfere with their possible moral agency. I will not take this path, because I think that we don't currently have conclusive evidence on this point. This is due both to problems with the methodologies used to study theory of mind in other species, and with our understanding of theory of

mind in humans (Andrews 2005, 2012). I will argue that theory of mind is not needed for either condition. Let us now examine each of the conditions in turn.

## **Autonomy**

Having reasons for actions, or having autonomy, or being an intentional agent are different ways of stating one demand of morality that cuts across moral theories. An autonomous agent will understand that her actions have consequences, and she can evaluate those consequences. I will assume the generic account of autonomy given by Christman: "to be autonomous is to be one's own person, to be directed by considerations, desires, conditions, and characteristics that are not simply imposed externally upon one, but are part of what can somehow be considered one's authentic self" (Christman 2009). Here an autonomous agent is contrasted with an individual whose every act is controlled by external forces such as fixed action patterns. An animal could be an autonomous agent if his behavior is flexible and the result of internal cognitive processes rather than a reflex or some species-typical behavior. For animals with flexible behavior, the question is then whether the internal processes are of the right sort.

I will defend an umbrella view of autonomy, which includes both diagnostic features and constitutive cognitive capacities. The diagnostic element takes seriously our acceptance of children as moral agents. The cognitive capacities involved will be related both to this developmental approach and the evolutionary one I sketched above. My view is at odds with one view that has been widely discussed in the animal morality context, namely Korsgaard's Kantian approach. Let me introduce her view in order to use it as a foil for the umbrella account.

Following Kant, Korsgaard has argued that all nonhuman animals lack the rational processes necessary for autonomy (Kant 1798; Korsgaard 2006). On Korsgaard's view animals lack autonomy because they lack what she calls *normative self-government* which is the ability to decide whether an act is justified and then act from that judgment rather than from one's desire. She writes:

*What it [normative self-government] requires is a certain form of self-consciousness: namely, consciousness of the grounds on which you propose to act as grounds. What I mean is this: a nonhuman agent may be conscious of it as fearful or desirable, and so as something to be avoided or to be sought. This is the ground of his action. But a rational animal is, in addition, conscious that she fears or desires the object, and that she is inclined to act in a certain way as a result. That's what I mean by being conscious of the ground as a ground. She does not just think about the object that she fears or even about its fearfulness but about her fears and desires themselves (Korsgaard 2006, 113).*

Korsgaard understands self-consciousness as consciousness about one's beliefs and desires—one's reasons—which is another way of describing having a theory of mind. A straightforward argument that moral agency requires having a theory of mind begins with the idea that moral agency requires autonomy, and autonomy requires acting for reasons, which in turn requires realizing that you have reasons for performing an action. Since reasons for action are sets of beliefs and desires that motivate behavior, and having a theory of mind is the ability to think about beliefs and desires, it follows that acting for reasons requires having a theory of mind. Or so the argument goes.

But we shouldn't let it go too far, at least in terms of the question of ape moral agency. Besides being an argument from within a Kantian framework, and hence rather irrelevant for other moral theories, the consequences of this view lead to counterintuitive conclusions about human children, who we should want to retain in the sphere of moral agents. While doing so, however, it is important to see what we can retain as well as what we should jettison from Korsgaard's account. I will question the notion that the

ability to consider one's own beliefs and desires is the only internal process that can generate something like normative self-government. I also want to make clear that this discussion isn't intended to criticize Korsgaard's project, only its applicability to the question at issue, which is whether we can make sense of children as moral agents without a theory of mind, and hence potentially apply that status to other apes as well.

So, on to the diagnostic aspect. If a moral agent needs explicit knowledge of her reasons for action and the ability to analyze them, it would follow that children, who do not even begin to understand belief until around 4 years old (Wellman et al. 2001), and who don't recognize some of the implications of belief attributions until mid-childhood (Apperly & Robinson 2001; 2002; 2003) are not autonomous agents. Even worse, it would follow that teenagers, who appear to be impaired at considering their own reasons for action (Pillow 1999; Morris 2000; Moshman 2004), are also outside the sphere of moral agency. At nine years-old, a child who burns a cat would not have acted immorally, on this view, because he isn't yet endowed with the cognitive capacities for evaluating his reasons for action. However, such a child will likely grow up to engage in anti-social behavior toward humans; we know what sort of a moral agent he will turn into. Rather than considering the nine year-old anti-social child as amoral, our common response is to retain moral language and judgments in describing the child and his actions. Psychologists who study moral cognition also base their work on the intuition that young children are within the sphere of the moral, as when Kohlberg asked what sort of moral reasoning ten year-old boys engaged in (Kohlberg 1971), and when Turiel and colleagues examined whether toddlers have a distinction between moral and conventional rules (Turiel and Nucci 1978).

While we take children to be responsible and in the sphere of the moral, we don't hold them fully responsible for their actions, as is reflected in the leniency given to children who commit crimes. Because children and adolescents are still developing their cognitive capacities and their ability to control their impulses and emotions, they are limited in what they *can* do, and hence are limited in what they *should* do. But this doesn't mean that children are not moral agents, and that their behavior cannot be categorized as good or bad. It just means that they have some degree of moral agency. And while children are limited to varying degrees during development, we start to treat them as moral agents when they become mobile and not entirely dependent on a caregiver.

Human children are seen as moral agents because they have degrees of the cognitive capacities required for normative self-government, and they live in a world of social norms that they can grow into and that their behavior is evaluated and shaped in terms of. Insofar as there are degrees of these cognitive capacities, there are also degrees of normative self-government. But as soon as an individual begins to attempt to shape her own behavior, we take her to have entered into the domain of moral agency, whatever mechanisms drive the attempt for self-betterment.

The case of children raises questions about whether a consideration of reasons for action, and hence a theory of mind, is the mechanism necessary for being an autonomous agent. To determine what those mechanisms might be, we can turn back to Christman's definition of autonomy: "to be autonomous is to be one's own person, to be directed by considerations, desires, conditions, and characteristics that are not simply imposed externally upon one, but are part of what can somehow be considered one's authentic self" (Christman 2009). We can appeal to this definition in order to ask about the cognitive capacities required in the case of children, and then turn to the question of whether apes have them.

Christman points to two requirements for autonomy: being directed by internal considerations, and being one's own person. I suggest that we can identify an individual who is directed by internal considerations by examining whether she can distinguish intentional from nonintentional action. An intentional action is done purposefully, and is often described as being done for reasons. If someone can sort intentional

actions from other kinds of actions, then there is at least an implicit recognition that these two kinds of actions are different, and that some actions are the responsibility of the agent, and others are not. While it would be experimentally difficult to determine whether nonhuman apes understand some of their own actions as intentional and others as unintentional, there is ample evidence that they understand this distinction in others. And, though it does not directly follow, we can infer that the individual who understands this about others likely also understand this about herself, given the dearth of human subjects who can understand others' mental states but not their own. (This view is even consistent with Carruthers's (2009) argument that mindreading others evolved before the ability to know one's own mind, because he is not committed to there being any developmental lag between using one's mindreading capacity on a third party and one's self; what is unique about his view is that mindreading supersedes any introspective access to one's own mind).

There is ample evidence that apes and human children understand intentional action in terms of individuals having goals and intentions. In one study, Uller found that chimpanzees, like human children (Gergeley et al. 1995), seem to perceive the behavior of geometric shapes moving in the right way as intentional (Uller 2004). For both humans and chimpanzees, a violation of expectation paradigm was used to measure the subjects' responses. In the child study, Gergeley concluded that the infants' surprise response to "irrational" behavior suggests that they attribute goals and rationality to the geometric shapes. In another study, Call and colleagues found that chimpanzees are more impatient with humans who are unwilling to give them food compared with humans who are unable to give them food; they beg more from the capable person who is unwilling than they beg from the person who is unable to access the visible food (Call et al. 2004). In addition, Jensen found that in an experimental setting chimpanzees will punish a conspecific who steals their food, by pulling a rope that causes the food to fall out of reach of both individuals. However, when the food is taken away by a human and given to the conspecific, the chimpanzee will not punish by pulling the rope (Jensen et al. 2007).

There is one study that suggests chimpanzees do not understand capability, and since understanding capability is necessary for distinguishing between intentional and unintentional actions, this study challenges the dominant view (Vonk & Subiaul 2009). In the study, chimpanzees are given the opportunity to beg for food from one of two humans, one of whom is unable to supply the food. Across a number of experimental conditions, the researchers found that the chimpanzees did not beg from the capable human significantly more often than from the incapable human, thus suggesting that the chimpanzees are unable to make this distinction. However, I have argued that this experiment fails to pick out the distinction (Andrews, 2012). In the experimental set-up, the humans were not actually rendered incapable; instead, different apparatuses were used to suggest that movement was restricted, though in fact the human could have maneuvered around the barriers. Given the concerns with this study and the findings of other studies, I think that the received view, that chimpanzees are sensitive to intentional agency, and can distinguish between intentional action and nonintentional action, is justified.

The second aspect of Christman's definition is being one's own person, and acting from one's authentic self. I interpret this as the ability to self-create rather than having one's self as given, since self-creation requires deliberately changing oneself.

The ability to self-create by purposefully changing oneself has not been given direct attention by great ape researchers or developmental psychologists. However, the observation that children spontaneously begin practicing adult actions as part of play at an early age influenced Maria Montessori's development of her early childhood education methods, many of which have been positively evaluated by developmental psychologists (Lillard 2005). Like human children, great apes engage in observational learning (Whiten 2000, Tomasello et al. 1987, Call & Tomasello 1994). Orangutans will position themselves so that they are only a few inches away from the behavior that they are observing, and will

subsequently attempt the behavior themselves (Call & Tomasello 1994). Some scientists think that great apes practice behaviors in order to develop competences (Anne Russon, pers. communication). More research on the issue of practice within the social learning project can help to determine whether and to what extent the great apes act to purposefully change themselves.

There are other areas of research that are ripe for exploring the question of self-improvement. Both human infants (Gartstein and Rothbart 2003) and many nonhuman animal species, like adult humans, have identifiable personality factors. Six personality factors have been identified in chimpanzees<sup>1</sup> (King & Figueredo 1997) and five in orangutans (Weiss, King, & Perkins 2006)<sup>2</sup>. The personality research program could take on the question of whether individuals understand their own personality traits, and whether they act to modify them. Such self-help activity would be compelling evidence for self-improvement.

While there may be prima facie reason for thinking that a moral actor needs to have a theory of mind because of considerations of autonomy, I think we should reject this connection. Autonomy, in the present context, is better understood as an umbrella concept, and while it can be defined in such a way so as to require theory of mind, defining it that way excludes not only nonhuman animals, but many humans as well. Rather, following de Waal if we understand autonomy as a concept on a gradation that includes a number of different cognitive capacities, we can make sense of our inclusion of children into the sphere of the moral, and we can better understand the variety of cognitive capacities that may be involved in moral activity. One of these capacities is surely a theory of mind, but other capacities include the ability to recognize intentional agency and the ability to learn from others and from experience. There is evidence that great apes have these capacities to some extent, and that insofar as one wants to cast a wide moral net, great apes may be included alongside young humans and unreflective adults.

### **Consequences of one's actions**

While the first requirement I gave for being a moral agent has historically been associated with Kant's moral theory, the second requirement is more aligned with consequentialist thinking, and is based on the role moral knowledge plays in moral agency. Plausibly, a moral agent needs to have some degree of knowledge about how her actions will affect another. But this requirement entails that we must predict how our actions impact others; we must be generally accurate predictors of behavior. And, one might argue, a theory of mind is necessary for accurately predicting intentional behavior.

The ability to see self and others as acting from beliefs and desires is often thought to be a necessary condition for the human ability to copiously and successfully predict behavior, and an ability that evolved to allow our ancestors to succeed in a complex Machiavellian social environment. One familiar account of the evolution of belief/desire attribution comes from the Social Intelligence Hypothesis, according to which human cognitive ability evolved as a result of our ancestors' complex social environment, rather than pressures of the physical environment. The pressures that come from living among others created a need to become better psychologists who are able to make better predictions of behavior, and these pressures led to the development of mental state concepts and a corresponding logic (Humphrey 1978, Jolly 1966). Behavioral and neurological evidence supports the claim that there is a correlation between sociality and cognitive abilities in a number of different taxa, and one explanation for the extraordinary social and cognitive abilities in humans stems from our success in confronting the task of better predicting and manipulating others' behavior (Byrne & Whiten 1988, Dunbar 1998). The story goes like this: In large social groups it is necessary to accurately predict the behavior of conspecifics. Having a theory of mind allows people to make more accurate predictions of behavior, because it lets a predictor understand someone's reasons for action (or the causes of action). Since theory of mind requires the postulation of theoretical entities such as belief and desire, and it requires the development of some mechanism for

using these theoretical entities to make predictions of behavior, only a sophisticated cognitive agent should be able to develop it.

I have challenged the premise that a theory of mind is necessary for making accurate predictions of behavior in such environments (Andrews 2009, 2012). I argue that a theory of mind is not necessary to predict the behavior of individuals who lack a theory of mind, and is not even the predominant method used to predict adult human behavior. Rather, I argue that adult humans predict from their understanding of the person *qua* person, not from their understanding of the person as a container for beliefs and desires. We develop our models of people by using a host of heuristics, and develop our predictions from our understanding of an individual's personality traits, stereotypic qualities, past behavior, emotional state, similarity to self, and so forth. Such models might include some propositional attitudes, but they need not. Thus, in a world without a theory of mind one could predict behavior using the same general process, by appeal to the model of the target; it's just that the model would lack any reference to belief. In our world, which is rich with belief attribution, we only use predictions based on propositional attitudes when the model at hand fails. And the model only fails in unusual or unfamiliar situations.

If all this is right, then reason attribution is not an automatic process for humans, but rather a deliberate and controlled process that is appealed to only when the agent deems it necessary. Given my account of predicting behavior, I challenge the claim that an appeal to propositional attitudes is necessary for moral cognition, and I claim that the version of the Social Intelligence Hypothesis that uses theory of mind to explain the correlation between sophisticated cognitive abilities and complex social groups cannot be correct. This is because in a world without mindreading, the kind of predictions one couldn't make would be predictions in anomalous or bizarre situations. Quotidian predictions would follow directly from the model, since the model was developed on the basis of quotidian situations. It would only be behaviors that are outside the norm that one would have difficulty in predicting. It might be tempting to claim that a theory of mind is needed to predict behavior in such situations, and I think that temptation leads to truth, with one caveat. In order to predict behavior in an anomalous situation, we must first seek to understand the situation. The odd situation and the odd behavior are things that need to be construed, and the act of construal consists of formulating an explanation for the behavior. A successful explanation will resolve the affective tension that drove the need to explain the behavior in the first place, and it will include additional information that can be added to the model. In some cases, the only information that will resolve that tension will be a belief attribution, and in those cases, having a theory of mind will be needed to make a prediction of behavior. However, note that the prediction is derivative of the explanation. That is, one must first be able to explain behavior in terms of belief before being able to predict in terms of belief. Thus, predicting behavior based on the attribution of beliefs and desires relies on a prior ability to construe behavior as being caused by beliefs and desires.

These considerations drive the argument that an understanding of a society's standards or normative rules is necessary for developing a theory of mind. This is because the ability to explain behavior develops before or alongside the development of a theory of mind, and in order to explain behavior in terms of beliefs, the explainer must have recognized the situation as anomalous. That is the explainer must have recognized that the actor was behaving outside the range of normal behavior, and the construal of the situation as anomalous implies at least an implicit understanding of society's norms.

An immediate objection is to deny that the norms that are identified are moral norms, and to claim that they are instead merely conventional norms. While moral norms may be taken to be universal, authority independent, and involve concepts such as help, harm, justice, rights, and cooperation, conventional norms are not universal, authority depending, and don't involve the moral concepts. Rather, conventional norms are seen as "behavioral uniformities which coordinate interactions of individuals within social systems" (Nucci & Turiel 1978). Turiel's research on children's early recognition of conventional vs. moral

violations was based on assumptions about the nature of the distinction between moral and conventional, something that moral philosophers have, despite their best efforts, not managed to come to consensus on (Kelly et al. 2007). Recent work on the moral/conventional distinction suggests that children and adults may not see a clear distinction between the two in some cases; for example, Nichols found that children, and to a lesser extent college students, responded to etiquette violations that involved disgusting behaviors in the same way they respond to moral transgressions (Nichols 2002, 2004). Given their analysis, Kelly et al. write that there is "a growing body of evidence justifying substantial skepticism about *all* the major conclusions that have been drawn from studies using the moral/conventional task" (Kelly et al. 2007).

I agree that there is reason to reject the distinction. Norms of convention may be based on the harm concept, and what appears to be a distinction may simply be two ends of a continuum. For example, the conventional norm against chewing gum might be a norm about avoiding harm, though low level or distant harm. One can tell a story about why children ought not chew gum in class: it gets all over the place, dries up balls under desks and gooey pieces stuck on the bottom of shoes. This causes harm in the sense of annoyance or disgust to the person who encounters the used pieces of gum, and might cause harm in the long run if resources that could be used to teach students are funneled into an expensive janitorial service. What causes these sorts of harms may depend on the other conventions of the society. Littering and leaving dog droppings are two behaviors that became seen as harmful in the west within recent memory, and social norms changed due to shifts in how people feel about such behaviors.

Nonetheless, there may still be some important theoretical distinction between moral and conventional rules, and there may be borderline cases that are less clear. What such a distinction may consist of takes us back to the concern I raised at the beginning of the paper; different moral theories will identify different necessary features of morality. Cleanly dividing norms into moral and conventional cannot be done outside any particular moral framework. But if the lay folk don't see a distinction, and the ethicists cannot agree on what a distinction would amount to, perhaps there is reason to think that it is nothing more than a theoretical construct.

A second concern is that my argument only requires having implicit knowledge of the social norms, and implicit knowledge is not open to the self-reflective analysis that is required on some accounts of autonomy. On my account, it is sufficient to have procedural knowledge how to act in the face of norm violations; I want to count individuals as having a degree of autonomy even if they fail to have explicit knowledge of the moral norms. As in my response to the last objection, here too I want to appeal to what people actually do. There is reason to suspect that most humans don't have explicit moral knowledge of at least some moral norms. One reason to suspect this is so comes from a series of studies on adult human moral reasoning. In these studies, Hauser and colleagues found that adult humans confabulate when providing moral justification for their response on how to respond to a trolley problem (Hauser et al. 2007). And in other studies, Jonathan Haidt found that insofar as adults make reference to moral principles, it is only post hoc and ad hoc attempts to rationalize their emotional reactions to moral situations (Haidt 2001). This result suggests that either adult humans don't have privileged access to all the norms they use when making moral judgments, or they don't use principles when making moral judgments.

To illustrate, consider a literary example. Huck Finn is a character who implicitly tracked the moral norms, even though he also knew that the legal principles of his society were in conflict with his intuitions. After lying to a man about having seen Jim, the runaway slave that Huck was floating down the Mississippi with, we hear Huck's inner soliloquy:

*"They went off and I got aboard the raft, feeling bad and low, because I knowed very well I had done wrong, and I see it warn't no use for me to try to learn to do right; a body that don't get STARTED right when he's little ain't got no show—when the pinch comes there ain't nothing to back him up and keep him to his work, and so he gets beat. Then I thought a minute, and says to myself, hold on; s'pose you'd a done right and give Jim up, would you felt better than what you do now? No, says I, I'd feel bad—I'd feel just the same way I do now. Well, then, says I, what's the use you learning to do right when it's troublesome to do right and ain't no trouble to do wrong, and the wages is just the same? I was stuck. I couldn't answer that. So I reckoned I wouldn't bother no more about it, but after this always do whichever come handiest at the time (Twain 1885, 118).*

This example illustrates that a person can live according to norms without being able to offer reasons for action in terms of those norms or even without the ability to recognize them as norms.

Just as Huck Finn was able to track norms that he wasn't explicitly aware of, the great apes may also be able to follow norms (and norm violations) without requiring any metacognitive awareness of such norms. I have discussed some of the evidence about norms in ape societies elsewhere (Andrews 2009). But here let me point to some of the findings from the ape culture literature. Culture research involves collaborations between researchers who study wild apes across different sites. The typical ape behaviors are listed, and researchers indicate whether or not that behavior is seen among their community. Using this method, we know that there is great behavioral variation within species of great apes, and that this variation does not reflect only behavior on the environment, but social behavior as well.

A good example of a sophisticated social behavior that involves something that looks like group norms is the group hunting behavior of the Tai forest chimpanzees in the Ivory Coast (Boesch 2005). There are cultural differences in hunting; some chimpanzee communities do not hunt at all, even though there are monkeys in the surroundings. In other communities chimpanzees hunt alone. But in the Tai forest, chimpanzees only hunt in groups, and each group member fulfills a particular role. In a four-member hunting party, the roles are Driver (who initiates a hunt by forcing the prey to move through the trees in a single direction), Blocker (who climbs trees to herd the prey toward the Driver), Chaser (who climbs under the prey to capture it), and Ambusher (who quietly climbs in front of the prey to block escape). Roles are determined functionally and flexibly, and may change during the course of the hunt depending on the positioning of the individuals and the monkey. The chimpanzees take on the role that they ought to take on, given the situation. After the hunt, the meat is shared between the hunters depending on their role in the hunt. Drivers and Ambusers rarely capture the monkey, and they receive about three times less meat than do the captors (Boesch 2002; Boesch & Boesch-Achermann 2000). However, if an Ambusher accurately predicts the prey's behavior, and the behavior of the other hunters, then he is given just as much meat as the captors. The hunting and meat-sharing behavior looks like a group norm because the group expects that meat be shared in this way, and each individual acts to enforce the expectation.

Other ape behaviors, such as the special treatment of infants in chimpanzee societies, have been argued by primatologists to serve as what they call proto social norms (von Rohr et al. 2011). Chimpanzee infants are given deferential treatment by other community members. Newborn chimpanzees are extremely interesting to other community members, and adults will watch intently but not try to approach the new member of their group. Juveniles and older infants, however, act on their interest by trying to approach or touch the new infant, which leads the mother to act hostilely or defensively; these young chimpanzees quickly learn not to approach newborn infants. Once infants are old enough to venture away from their mother, adults are extremely tolerant of infants climbing over them and even stealing their food or tools. Adults are also reported to self-handicap when playing with infants. A violation the norms regarding infants, such as infanticide or other aggressive acts towards infants, often results in an uproar of vocal

protest from adult females, and can also cause third party intervention. The authors take such behaviors to be a potential proto social norm because they involve behavioral regularities that, when violated, elicit bystander reactions, including punishment (in the sense of indignation toward the transgressor) and the expression of moral emotions.

I hope to have made the prima facie case that there can be sensitivity to the normative without the advanced cognitive capacities required to consider reasons for action. Without a theory of mind, one can have moral knowledge of the society's norms and standards in the sense of tracking such norms, and that coupled with a motivation to follow the society's norms results in a moral status that is not unlike that of many humans.

### **Ape moral agency**

I've challenged the two central arguments connecting theory of mind to moral agency. If theory of mind develops from an understanding of moral norms, then some moral knowledge is antecedent to having a theory of mind. In addition, I have raised some worries about the consistency of holding both that most humans are moral agents and that most humans are limited in their ability to engage in the kind of normative self-government required by Korsgaard. I suggest that there are other cognitive mechanisms that could be used to engage in self-governing and self-creation.

But let me clearly state the limitations of my argument. I have not shown that great apes should be considered moral agents, but only that the prima facie case against ape moral agency, based on the idea that they lack a particular cognition capacity, is not warranted. However, this discussion has suggested a kind of upper limit to the sort of moral agency that apes might have. I have argued that one can have knowledge of the norms of society without a theory of mind, and one can have the cognitive capacities needed to change oneself without a theory of mind. Taken together, it follows that without a theory of mind one could still purposefully act to become a better group citizen, one could know that there are such moral norms, want to conform to them, and control one's behavior so as to better conform to them.

While this describes the upper level of the kind of moral agency available to a creature without a theory of mind, it still describes something that is identified as a level of morality. Without a theory of mind, one can still participate in Kohlberg's conventional stage of moral development, in which one recognizes and acts so as to follow the group's rules. While such an individual may remain there, it is only from within the perspective of a particular moral theory that we can conclude that he does not count as a moral agent.

At any rate, once at the conventional level of morality, one has some concept of acceptable and unacceptable, good and bad, right and wrong. And at this point the individual has entered the domain of the normative.

<sup>1</sup> The chimpanzee personality factors include correlates for all the human ones (Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness) plus a factor for Dominance.

<sup>2</sup> Orangutans showed only the correlates to Neuroticism, Extraversion, Agreeableness, and Dominance, but also showed a factor that is a combination of Conscientiousness and Openness that was called Intellect

## Bibliography

- Andrews, K. (2005). Chimpanzee theory of mind: Looking in all the wrong places? *Mind and Language*, 20(5), 521-536.
- Andrews, K. (2009). Understanding norms without a theory of mind. *Inquiry*, 52(5), 433-448.
- Andrews, K. (in preparation). *Persons as We Know Them: A Pluralistic Folk Psychology*.
- Apperly, I. A., & Robinson, E. J. (2001). Children's difficulties handling dual identity. *Journal of Experimental Child Psychology*, 78(4), 374-397.
- Apperly, I. A., & Robinson, E. J. (2002). Five-year-olds' handling of reference and description in the domains of language and mental representation. *Journal of Experimental Child Psychology*, 83(1), 53-75.
- Apperly, I. A., & Robinson, E. J. (2003). When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children's reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, 85(4), 297-311.
- Bekoff, M., & Pierce, J. (2009). *Wild Justice: The Moral Lives of Animals*. Chicago: University of Chicago Press.
- Boesch, C. (2005). Joint cooperative hunting among wild chimpanzees: Taking natural observations seriously. *Behavioral and Brain Sciences*, 28(5), 692.
- Boesch, C. (2002). Cooperative hunting roles among Tai chimpanzees. *Human Nature*, 13(1), 27-46.
- Boesch, C., & Boesch-Achermann, H. (2000). *The Chimpanzees of the Tai Forest: Behavioural Ecology and Evolution*. Oxford: Oxford University Press.
- Byrne, R., & Whiten, A. (Eds.). (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. New York: Oxford University Press.
- Call, J. & M. Tomasello. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Science*, 12(5): 187-192.
- Call, J., & Tomasello, M. (1994). The social learning of tool use by orangutans (*Pongo pygmaeus*). *Human Evolution*, 297-313.
- Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). 'Unwilling' versus 'unable': Chimpanzees' understanding of human intentional action. *Developmental Science*, 7(4), 488-498.
- Carruthers, Peter. 2009. "How we know our own minds: The relationship between mindreading and metacognition". *Behavioral and Brain Sciences*, 32: 121-182.
- Christman, J. (2009). Autonomy in Moral and Political Philosophy. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Fall 2009 ed.).
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6(5), 178-190.
- Fraser, O.N., D. Stahl & F. Aureli. 2008. "Stress reduction through consolation in chimpanzees". *Proceeding of the National Academy of the Sciences*, 105(25): 8557-8562.

Fraser, O.N. & F. Aureli. 2008. "Reconciliation, consolation and postconflict behavior specificity in chimpanzees". *American Journal of Primatology*, 70: 1114-1123.

Gartstein, Maria A. & Mary K. Rothbard. 2003. "Studying infant temperament via the Revised Infant Behavior Questionnaire". *Infant Behavior and Development*, 26(1): 64-86.

Gergely, G., Nadasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56(2), 165-193.

Goodall, J. (1986). *The Chimpanzees of Gombe : Patterns of Behavior*. Cambridge, MA: Harvard University Press.

Haidt, Jonathan. 2001. "The emotional dog and its rational tail: A social intuitionist approach to moral judgment". *Psychological Review*, 108: 814-834.

Haksar, V. (1998). Moral Agents. In E. Craig (ed.), *Routledge Encyclopedia of Philosophy*. (London: Routledge, 1998). Retrieved January 14, 2004, from <http://www.rep.routledge.com/article/L049>

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59(4), 771-785.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61(1), 139-151.

Hauser, M., Cushman, F., Young, L., Jin, R. K.-X., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind and Language*, 22(1), 1-21.

Hirata, S. (2003). Cooperation in chimpanzees. *Hattatsu*, 95, 103-111.

Humphrey, N. (1978). Nature's psychologists. *New Scientist*(29), 900-904.

Jensen, Keith, Josep Call & Michael Tomasello. 2007. "Chimpanzees are vengeful but not spiteful". *Proceeding of the National Academy of the Sciences*, 104(32): 13046-13050.

Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science*, 153, 501-506.

Kant, I. (1798/1977). *Anthropologie in Pragmatischer Hinsicht* (translated as *Anthropology from a pragmatic point of view*) Carbondale: Southern Illinois Press.

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. T. (2007). Harm, affect, and the moral/conventional distinction. *Mind and Language*, 22(2), 117-131.

King, J. E., & Figueredo, A. J. (1997). The five-factor model plus dominance in chimpanzee personality. *Journal of Research in Personality*, 31(2), 257-271.

Kitcher, P. (2006). Ethics and Evolution: How to Get Here from There. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 120-139). Princeton: Princeton University Press.

Kohlberg, L. (1971). *From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development*. New York: Academic Press.

Korsgaard, C. (2006). Morality and the distinctiveness of human action. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 98-119). Princeton: Princeton University Press.

- Lillard, Angeline S. 2005. *Montessori: The Science Behind the Genius*. New York: Oxford University Press.
- McGrew, W. C. (1992). *Chimpanzee Material Culture: Implications for Human Evolution*. Cambridge: Cambridge University Press.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees Recruit The Best Collaborators. *Science*, 311, 1297-1300.
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, 84(2), 221-236.
- Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development*, 49(2), 400-407.
- Nichols, S. (2004). *Sentimental Rules: On the Natural Foundation of Moral Judgment*. Oxford: Oxford University Press.
- Parr, Lisa A. 2001. "Cognitive and physiological markers of emotional awareness in chimpanzees (*Pan troglodytes*)". *Animal Cognition*, 4(3-4): 223-229.
- Parr, L. A. & W. D. Hopkins. 2000. "Brain temperature asymmetries and emotional perception in chimpanzees, *Pan troglodytes*". *Physiology and Behavior*, 71: 363-371.
- Plooij, F. X. (1978). Some basic traits of language in wild chimpanzees. In A. Lock (Ed.), *Action, gesture, and symbol: The emergence of language*. New York: Academic Press.
- Singer, P. (2006). Morality, Reason, and the Rights of Animals. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 140-160). Princeton: Princeton University Press.
- Smetana, J. 1981: Preschool children's conceptions of moral and social rules. *Child Development*, 52, 1333-1336.
- Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, 70(1), 1-131.
- Tomasello, M., Davis-Dasilva, M., Camak, L., & Bard, K. (1987). Observational learning of tool-use by young chimpanzees. *Human Evolution*, 175-183.
- Turiel, E. and Nucci, L. 1978: Social interactions and the development of social concepts in preschool children. *Child Development*, 49, 400-407.
- Twain, Mark. (1884/1885). *The Adventures of Huckleberry Finn*. New York: Charles L. Webster and Company.
- Uller, C. (2004). Disposition to recognize goals in infant chimpanzees. *Animal Cognition*, 7(3), 154-161.
- van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., et al. (2003). Orangutan cultures and the evolution of material culture. *Science*, 299(5603), 102-105.
- von Rohr, Claudia Rudolf, Judith M. Burkart & Carel P. van Schaik. 2011. "Evolutionary precursors of social norms in chimpanzees: A new approach". *Biology and Philosophy*, 26: 1-30.

Vonk, J., & Subiaul, F. (2009). Do chimpanzees know what others can and cannot do? Reasoning about 'capability'. *Animal Cognition*, 12(2), 267-286.

de Waal, F. (2009). *The Age of Empathy: Harmony*.

de Waal, F. (2006). *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press.

de Waal, F. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in infants and young chimpanzees. *Science*, 311(5765), 1301-1303.

Weiss, A., King, J. E., & Perkins, L. (2006). Personality and subjective well-being in orangutans (*Pongo pygmaeus* and *Pongo abelii*). *Journal of Personality and Social Psychology*, 90(3), 501-511.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655-684.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. *Nature*, 399(6737), 682-685.

Whiten, A. (2000). Primate culture and social learning. *Cognitive Science: A Multidisciplinary Journal*, 477-508.

Wright, R. (2006). The uses of anthropomorphism. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 83-97). Princeton: Princeton University Press.

Wright, R. (2006). The uses of anthropomorphism. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 83-97). Princeton: Princeton University Press.