Ape Autonomy? Studies in Natural Moral Psychology

Kristin Andrews
York University

In *Philosophical Perspectives on Animals: Mind, Ethics, Morals*, K. Petrus & M. Wild, eds. Transcript. 2013.

Introduction

Much of what science has been uncovering about great ape cognition, social organization, and behavior suggests that apes should be granted moral standing. In addition, there have been claims that the growing understanding we have about great apes and some other species might be sufficient for concluding that some animals have morality, or have the ability to engage in moral and immoral behavior (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 laboratory research we know that great apes can recognize intentions (Uller 2004; Call et al. 2004; Warneken & Tomasello 2006) and engage in cooperative behavior (Hirata 2003; Melis et al. 2006). We also have evidence 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).

This sort of evidence has led some to conclude that many species have morality to some degree, since there is a smooth continuity between moral and amoral agents, just as there is a smooth continuity between human and nonhuman minds. On such continuity views, the psychological capacities required for moral agency are not all-or-nothing. Frans de Waal takes this position when he argues that morality requires both empathy and

reciprocity, and that many species demonstrate primitive versions of these necessary requirements. He takes these behaviors as evidence that animals have a moral sense, though one that is not as developed as the human moral sense. On the other hand, one might 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 & 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 the differences, the important similarities between species include empathy, altruism, cooperation and perhaps a sense of fairness.

What remains unclear in such claims is what it means to "have morality". Initially, one might think that asking whether great apes have the qualities of a moral agent is a more specific question. A moral agent is someone who can be held responsible for her behavior, whose behavior can appropriately be judged morally acceptable or not. Nonhuman animals are typically not seen as moral agents, even though they are thought to have agency, and may be considered moral patients. Something more is needed, as reflected in the following passage from the *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 again we run into problems, because it is unclear what it means to be a demand of morality. Answers vary, because determining a demand of morality can only be answered from within the framework of a moral theory, and the different moral theories will provide different answers. Subsequently, the 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 a Humean sense, but could still be a moral agent under 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 saliencies 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.

This lack of clarity 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. But the additional burden is augmented by 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 the animal are irrelevant or

not sufficient for being deemed a moral agent. 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 the 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 other species has a theory of mind. While there is no evidence that great apes have anything like the philosopher's representational concept of belief, there is evidence that great apes

have other cognitive capacities that can fulfill the same functions that are sometimes seen as requirements for moral agency. As well, I will argue that recognizing social norms drive the development of metacognitive social abilities such as theory of mind, and so we should not expect that theory of mind is 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 action. For each condition, having a theory of mind seems to be implicated. In

response, one might argue that great apes have a theory of mind, and hence this requirement does not interfere with their possible moral agency. I will not take this path, because I think that we don't currently have conclusive evidence about whether or not great apes have a theory of mind. This is both due 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). 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. Though it is emphasized differently within the different moral theories, this demand seems to cut across theories as a basic requirement for moral agency. An autonomous agent knows that her actions have consequences, and she thinks she can affect change in the world through her actions. As in moral agency, there are various accounts of autonomy that reflect the demands of a particular moral theory. However, we can point to the following as a baseline for 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). Put this way, an autonomous agent is contrasted with an individual whose every act is controlled by external forces, so an animal whose behavior was completely controlled by, e.g., fixed action patterns (such that all behavioral sequences are inflexible and determined by environmental stimuli) could not be an autonomous agent. However, an animal might be an autonomous agent if his behavior is flexible and the result of internal

cognitive processes rather than reflex or association with environmental stimuli. For such animals, the question would then be whether the internal processes are of the right sort.

However, it is argued that all nonhuman animals lack the rational processes necessary for autonomy (Kant 1798; Korsgaard 2006). Korsgaard, for example argues that animals lack autonomy in the sense of the normative self-government that allows one 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 the object of his fear or desire, and 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).

This passage from Korsgaard reflects the view that having autonomy requires having a theory of mind—the ability to attribute propositional attitudes to oneself and others. To realize *that* one fears or desires some state of affairs, one needs to have the attitude concepts and have the ability to accurately represent the contents of the propositional attitudes.

The argument could go like this: in order to have moral agency, one needs also to have autonomy. An autonomous agent is able to act for reasons. And since acting for reasons requires the ability to recognize that one has reasons for actions, and reasons for actions can be seen as sets of beliefs and desires that motivate behavior, it follows that

acting for reasons requires having a theory of mind. Thus, the worry arises that the moral agent must have a theory of mind, and if nonhuman animals lack a theory of mind they cannot be a moral agent.

This argument raises a number of concerns. An obvious first issue is that the argument presupposes a Kantian moral framework according to which the ability to act from one's own reasons is essential for moral agency. Contrast this to a consequentialist account in which a moral actor may fail to consider abstract reasons for action, but instead weigh two projected outcomes in order to determine which is best by some criterion.

Setting that concern aside, we can still take seriously the objection, but question the claim that the ability to consider one's own beliefs and desires is the only internal process that can generate normative self-government. I think we ought to reject that assumption, in part given the high standards set out in Korsgaard's requirements for autonomy. On this view, a moral agent has 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. This flies in the face of common sense according to which children have a very early entry into the domain of morality, even though the range of their moral behaviors is quite limited. By examining moral development starting with ten-year old boys, Kohlberg presumed that there was moral reasoning of some sort at this age, and that assumption mirrors the commonsense conception that children are moral agents, even though they cannot be held responsible

for all their actions. That we recognize children as not *fully* responsibility for their actions is reflected in their special legal status. Since children are still developing their cognitive capacities and their ability to control their impulses and emotions, children are limited in what they can do. Given the acceptance of the "ought implies can" principle, children enjoy this special status. 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 are limited in their 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 so utterly dependent on a caregiver.

So, we see that there are degrees of these cognitive capacities, and hence in 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 morally significant domain, whatever processes drive the attempt for self-betterment.

There is another concern about the abstract reasons view that Korsgaard promotes, which requires that people have explicit knowledge of their reasons and are able to analyze those reasons when engaging in moral behavior. Research on adult moral reasoning suggests that adults do not generally consider their reasons when making moral judgments. Jonathan Haidt's studies on moral reasoning suggest that adult humans do not engage in a consideration of reasons when making moral judgments (Haidt 2001). While such findings do raise some worry for the strong view of autonomy, it should be pointed out that on this view autonomy only requires that one *can* consider one's grounds for action, and this is consistent with the findings that humans very often do not consider their reasons for their moral judgments. However, to take this position as a response to

Haidt's findings would be to accept that humans are not living up to their moral obligations most of the time.

These considerations about actual human moral cognition and behavior raise questions about whether considering reasons for action is necessary for being an autonomous agent. In order to examine the possibility that there are other ways of having autonomy, let's return to Christman's general 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, and whether apes have them.

Let's focus on two aspects of this definition in turn; first we'll look at being directed by internal considerations, and then we'll turn to the issue of being one's own person. What does it mean to be directed by internal considerations rather than those that are externally imposed? Let me suggest a sufficient condition: the ability to distinguish intentional from nonintentional action. An intentional action is done purposefully, and would often be 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.

There is evidence that apes have the cognitive capacities necessary for some degree of self-improvement and for identifying intentional behavior. The current view about chimpanzee social cognition, for example, is that chimpanzees understand that

others have mental states such as seeing, and that seeing motivates individuals to act (Hare et al. 2000; Hare, Call, and Tomasello 2001; Plooij 1978; Goodall 1986; de Waal 1996). It is also thought that chimpanzees understand goals and intentionality (Uller 2004; Tomasello & Carpenter 2005; Warneken & Tomasello 2006). For example, Claudia 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. While Gergerly concluded that the infants' surprise response to "irrational" behavior suggests that they attribute goals and rationality, Uller's conclusion was more circumspect, even though her infant chimpanzees subjects responded in the same way as the human infants. She concludes that chimpanzees, at least, have an understanding of agency.

This finding concerning infant chimpanzees is consistent with other experimental studies of intentionality attribution. 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). Warneken & Tomasello found that chimpanzees respond appropriately to the communicative gestures of human caregivers (Warneken & Tomasello 2006). While engaged in what appeared to be casual social interaction with the experimenter, the chimpanzee would be tested on their ability to respond to a nonverbal request for help. In one condition, the experimenter dropped an object and requested that a chimpanzee pick it up.

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 who 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 form 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, in preparation). 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 a justified position to take.

Now, on to the second aspect of Christman's definition. What does it mean to be one's own person, and act from one's authentic self? One plausible interpretation is that an authentic self is one that is self-created, rather than given to the agent fully formed. To create one's self requires the ability to deliberately change oneself. There are various ways in which one can act to change or control her behavior. One way of changing oneself is to work things out in the space of reasons. But other ways would be to develop habits of behavior and to learn how to control emotional responses. It would require learning.

The ability to self-create by purposefully changing oneself has not been given direct attention by great ape researchers. If researchers are going to directly examine the question of whether great apes work to improve themselves, the importance of this ability needs first to be articulated. However, there are some findings that suggest great apes do act to improve themselves in various ways. For example, the social learning literature indicates that great ape species do learn from observing others' behavior (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 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 great ape self-improvement. The field of cross-species personality research has identified six personality factors in chimpanzees (King & Figueredo 1997) and five factors for orangutans (Weiss, King, & Perkins 2006). While the initial drive of this research was to determine whether there are personality differences in other species, and what the personality factors consist of, future work could be on changes in personality and the events and behaviors that drive personality changes. Such investigation may also further our understanding of whether apes or other animals take steps toward something like 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 is 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, autonomy seems to be a concept on a gradation that includes a number of different cognitive capacities. One of these capacities is surely a theory of mind, but other capacities include the ability to recognize intentional agency, the ability to learn from others and from experience, the ability to refrain from action and to self-modulate one's emotional state. There is evidence that great apes have many of these capacities to some extent, and that insofar as one wants to cast a wide moral net, great apes may be caught up alongside young humans and unreflective adults.

Consequences of one's actions

The first argument about the relation between theory of mind and moral agency had a decidedly Kantian flavor. This second argument is more aligned with consequentialist thinking, and is based on the role of moral knowledge for moral agency. Plausibly, a moral agent needs to have some degree of knowledge about how one's 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 seen as 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, and it takes sophisticated cognition to develop a theory of mind, because it 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.

I have challenged the premise that a theory of mind is necessary for making accurate predictions of behavior in such environments (Andrews 2009). 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. Instead of relying of attribution of propositional attitudes, adult humans predict from their

understanding of the person *qua* person, not from their understanding of the person as a container for belies 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 default to 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 folk prediction, I deny the claim that appeal to propositional attitudes is necessary for moral cognition. Further, the Machiavellian version of the Social Intelligence Hypothesis cannot be correct. This is because in a world without theory of mind, 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 is based on quotidian situations. It would only be behaviors that are outside the norm that one would have difficultly 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 is correct, 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 society's 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. In other words, the explainer must recognize that an actor is behaving outside the range of normal behavior. This requires a construal of the situation as anomalous. And this 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 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, are authority dependendent, and don't involve moral concepts. 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 morality and convention, 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).

There is some reason to reject such a distinction. Norms of convention may be based on the harm concept, and what appears to be an exclusive distinction may simply be two ends of a continuum. For example, the conventional norm against chewing gum in class might be a norm about avoiding harm, albeit 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, dried 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.

Nonetheless, there may still be some important theoretical distinction between moral and conventional rules, and there may be border 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 the premises only entail that one would have an *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. This is because we want to count as having at least some degree of autonomy those individuals who do not 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 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 justifications for their response to a trolley problem (Hauser et al. 2007). This result suggests that adult humans don't have privileged access to all the norms they use when making moral judgments.

Perhaps the best response to these concerns is to look at the evidence for social norms among great ape communities. I have discussed some of this evidence elsewhere (Andrews 2009). But here let me point to some of the findings form the ape culture literature. The 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 local 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 is potential prey in the environment. 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 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 prey. 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.

In responding to these objections 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, and knowledge of norms together with a desire to follow those norms results in what is, plausibly, normative behavior. And there is good reason to accept such as moral behavior, because for many humans, morality is similarly limited to knowledge of social norms and the desire to conform to them.

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 have limits on 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 cognitive capacity, is not warranted. However, this discussion has suggested a kind of upper limit to the sort of moral agency that apes could enjoy (or suffer). 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 also describes something that can rightly be 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 at that stage, it is only from within the perspective of a particular moral theory that we can conclude that that does not count as moral agency.

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.

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., & 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.

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.

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, J. (2001). The emotional dog and its rational tail. Psychological Review.

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.

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.

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.

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.

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.

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.

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.