

## **Primates, philosophers and the biological basis of morality: a review of primates and philosophers by Frans De Waal, Princeton University Press, 2006, 200 pp**

**Massimo Pigliucci**

Received: 22 February 2007 / Accepted: 22 February 2007 / Published online: 31 May 2007  
© Springer Science+Business Media B.V. 2007

### **Introduction**

Philosophical inquiries into morality are as old as philosophy, but it may turn out that morality itself is much, much older than that. At least, that is the main thesis of primatologist Frans De Waal, who in this short book based on his Tanner Lectures at Princeton, elaborates on what biologists have been hinting at since Darwin's (1871) book *The Descent of Man* and Hamilton's (1963) studies on the evolution of altruism: morality is yet another allegedly human characteristic that turns out to be built over evolutionary time by natural selection.

This sort of intellectual project has historically generated tension between “the two cultures” (Snow 1959), and just a few years ago E. O. Wilson (1998) was predicting an eventual reduction of all humanities to biology by a continuing expansion of sociobiology and evolutionary psychology. It is interesting, therefore, to find in De Waal's book a clear willingness toward dialog, both on the part of the scientist and by the four philosophers (Robert Wright, Christine Korsgaard, Philip Kitcher and Peter Singer) called upon to comment on the main essay. Indeed, the idea that science can and should inform philosophical debates is becoming rather popular among both scientists and philosophers (e.g., Gibbard 1997; Sober and Wilson 1998; Skyrms 2000; Green et al. 2001; Pigliucci 2003; Clayton and Schloss 2004). That said, there is plenty of room for disagreement, which is of course where things get interesting.

De Waal's objective is two-fold: to make a case for the idea that humans are by nature moral, and to sketch out how such morality could have evolved from simpler building blocks present in our ancestors and our current close relatives. I will first examine De Waal's case for his two conclusions, then move on to some of the four philosophers' objections to parts of his thesis, and finally consider the merits and limits of De Waal's own response to his critics.

---

M. Pigliucci (✉)  
Departments of Ecology and Evolution and of Philosophy, Stony Brook University,  
New York, NY 11794, USA  
e-mail: pigliucci@platofootnote.org

### Humans are good-natured: the demise of veneer theory?

De Waal's first objective is to demolish what he calls "veneer theory," the idea that humans are, by nature, selfish and immoral, and that morality is a thin veneer precariously maintained by elaborate rituals and institutions that define civilization. Interestingly, while De Waal appropriately cites Hobbes' (1651) central tenet that social life is "by covenant only, which is artificial," he blames none other than biologist Thomas Huxley for the most powerful articulation of veneer theory (p. 7). In his famous *Evolution and Ethics*, Huxley (1894) argues that while nature is "red in tooth and claw," humanity has managed to overcome its natural condition. This requires a significant and continued effort, according to Huxley, just as a gardener has constantly to battle the continuous invasions of his plot by weeds that would quickly return it to its chaotic primordial state. In De Waal's view, Huxley's position was particularly striking considering that Darwin (1871) himself had clearly taken the opposite stance, including morality within the purview of human nature in his *Descent of Man*. How could Darwin's bulldog so thoroughly deny the power of natural selection?

De Waal points out that Huxley's position is by no means unusual (p. 8). Freud (1930) saw modern humans as continuously engaged in a struggle to overcome their (mostly sexual) animal instincts; Williams (1988) is quoted by De Waal as writing "I account for morality as an accidental capability produced, in its boundless stupidity, by a biological process that is normally opposed to the expression of such capability"; and Dawkins (1976) tells us that "We, alone on earth, can rebel against the tyranny of the selfish replicators," although he awaits until the very end of *The Selfish Gene* to give us this bit of consolation.

The most convincing evidence that there may indeed be a "veneer theory," even if perhaps nobody explicitly defends it as such, is that biologists have devised all sorts of explanations for altruistic behavior that implicitly assume the veneer view of morality. The classic accounts of altruism fall into one of the following categories: it is the result of simple self-deception, to make the actor's selfish actions more easy to disguise; or it is the product of kin selection (i.e., it is still about passing copies of one's own genes); or, finally, it is based on reciprocity (i.e., it still benefits the "altruist," though sometimes only statistically and in the long run, as in iterated versions of game theoretical situations such as the prisoners' dilemma). That is, there really is no such thing as altruism if one understands the word to mean a sacrifice of one's resources (or life) to benefit a stranger with no expectation of returns. From that perspective, then, Huxley & co.'s view of human (true) altruism as something highly unstable and somehow outside of our biology makes sense, though De Waal rightly points out that no sympathizer of veneer theory actually provides an account of how on earth *Homo sapiens* has somehow managed to transcend its biological nature (p. 10), albeit precariously and with much struggle.

The major exception to this denial of true altruism is one that, interestingly, De Waal is not sympathetic with: a variety of thinkers have invoked group selection to explain the evolution of true (ingroup) altruism, from Darwin himself to the half-baked ideas of Kropotkin (1902), to the more recent attempt by Sober and Wilson (1998). De Waal writes (p. 16) "I remain unconvinced that we need group selection to explain the origin of these [moral] tendencies—we seem to get quite far with theories of kin selection and reciprocal altruism." This implies that De Waal may in fact agree that there is no such thing as "true" altruism, but it provides him the latitude to use potent concepts of modern evolutionary theory, such as kin selection and reciprocal altruism, to investigate how the building blocks of morality may have been put in place by natural selection before the

evolution of humans – as long as one keeps distinct the idea of altruism (in the biological sense) from that of morality. But where to start looking for pertinent evidence? Where Darwin himself looked: data on emotions and sympathy in non-human animals, particularly primates.

#### The building blocks and evolution of human morality

Darwin thought that the step from social instincts to full-fledged morality was perfectly compatible with his general theory, as this passage of *The Descent of Man* quoted by De Waal (p. 14) clearly summarizes: “Any animal whatever, endowed with well-marked social instincts, the parental and filial affections being here included, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers had become as well developed, or nearly as well developed, as in man.” And Darwin saw plenty of evidence for those “social instincts,” beginning with the parent-offspring relationship that he explicitly mentions, and continuing with manifestations of sympathy at the distress of other animals. So, according to De Waal, Darwin may have gotten the specific mechanism (group selection) wrong, but his intuition and observational powers as a naturalist were right on target: full fledged morality can evolve by building on blocks that include behavior toward kins and sympathy for other members of the group, related or not.

De Waal summarizes his view of the evolution of morality in a convenient table (p. 22) where the transition from social to moral animal is seen as based on a continuity between other primates and humans. The mechanisms, as referred to above, are kin selection and reciprocal altruism, and the evidence falls into three categories: from neuroscience, where studies show that moral dilemmas (such as the “trolley problems” so popular among philosophers) activate areas of the brain involved with emotions; from psychology, which supports the idea that morality has a strong emotional component in humans; and from comparative biology, with its mounting evidence concerning the appearance of components of moral behavior in non-human primates. De Waal devotes most of the rest of his essay to the latter source of evidence, which is his own field of technical expertise.

He begins with four building blocks towards human morality: emotional contagion, personal distress, empathy and sympathy. Emotional contagion is a very basic phenomenon, widespread among social primates, where the emotional state of an individual induces a similar state in another individual. This can cause personal distress, which is largely self-directed in that it generates behaviors aimed at ameliorating one’s own (as opposed to others’) emotional situation. Empathy, however, is defined by De Waal as “the original, pre-linguistic form of inter-individual linkage that only secondarily has come under the influence of language and culture” (p. 24). De Waal assumes that empathy evolved within the context of parental care, and then extended to other members of the group, regardless of genetic relatedness. He refers to a significant amount of evidence that primates help each other in case of need (e.g., in a fight) and, more importantly, engage in behaviors that is hard not to classify as empathic, such as putting an arm around someone who has been attacked. Finally, we come to sympathy, an emotional response in which the animal feels concern for another one, thereby making the crucial move from self- to other-directed emotions. Empathy appears in humans at the age of two, but it is difficult to explain the behavior of some rhesus monkeys as anything other than true sympathy: for instance, individuals have been observed to resist pulling a chain that delivered food to them once they realized that another monkey was being electrically shocked as a result. It takes a high degree of social instinct, as Darwin would have put it, to almost starve oneself so to avoid

hurting another member of the group. (Indeed, by that standard, rhesus monkeys may be more moral than human beings, since the latter have been shown to willingly pull an electroshock lever when asked to do so by an authority figure, as in the classic experiments by Milgram 1974.)

All of this brings De Waal to present his “Russian Doll Model” of the layers necessary to build a fully human sense of morality (p. 39). At the very core of any primate moral sense is Emotional Contagion, which ensures an automatic emotional impact of the state of an individual over that of another. Building on this is the second layer, Cognitive Empathy, whereby the animal actually assesses another’s emotions. Finally, we have full fledged Attribution, where one is capable of adopting another’s perspective, as it is possible for normal human beings (but, interestingly, not for individuals with specific brain damage, who display essentially sociopathic behavior).

The rest of De Waal’s essay details tantalizing, if of course individually debatable, examples of proto-moral behavior in our close primate relatives. We learn that chimpanzees and capuchin monkeys share food outside of the mother–offspring relationship, but that they do so conditional to a system of reciprocity: observational studies show that food sharing is more likely with individuals who had earlier been grooming the one giving his food away. De Waal suggests that this biological phenomenon eventually gave rise to the psychological feeling that we call “gratitude.”

There is also some intriguing evidence of “inequity aversion” among capuchin monkeys. In experimental situations where the monkeys were offered rewards after a similar offer had been made—in full view of the subject—to another monkey, individuals were less likely to accept their own reward in situations where the other monkey had received a better deal from the experimenter. While one ought to be aware of the perils of projecting human interpretations on another primate’s behavior (indeed, De Waal devotes an entire appendix of his essay to the twin perils of anthropomorphism and “anthropodenial”), it is hard to resist the idea that the capuchins are in fact capable of the rudiments of what we would call a sense of fairness.

Finally, De Waal also discusses the other side of the moral coin: if it is true that other primates only show elements of moral sense, it is also demonstrable that human morality has precise limits that root it squarely into its evolutionary past. In particular, it is clear that human beings—despite formal protestations to the contrary—have a precise sense of a hierarchy of moral duties, where the self and its kins come first, followed by the in-group. The idea of being moral toward individuals from other groups is very recent and, as one can plainly see from a cursory reading of daily headlines, very fragile indeed. This progression is, of course, precisely what an evolutionary biologist would expect if morality were anchored in the biological phenomena of kin selection and reciprocal altruism.

#### The philosophers’ responses

The responses to De Waal’s opening piece are interesting in that they reflect the four philosophers’ individual and varied preoccupations with specific aspects or implications of De Waal’s central argument. So, Wright focuses on the use (and abuse) of anthropomorphism, Korsgaard on the distinctiveness of human action, Kitcher on the possibility that “vener theory” is actually a straw man, and Singer, of course, on the implications for animal rights.

Robert Wright distinguishes between two types of anthropomorphic language: emotional (i.e., the attribution of human-like emotions, such as compassion or outrage, to other

primates), and cognitive (when one attributes human-like mental states, such as consciousness, to our evolutionary cousins). He very reasonably suggests that, from an evolutionary perspective, surely human-like emotions appeared earlier in the great ape clade than human-like cognition. That being the case, De Waal may be going beyond what is parsimoniously needed when he uses both emotional and cognitive anthropomorphic languages, as the first one will probably suffice. Perhaps more interestingly, Wright rejects the label of “vener theorist” imposed on him by De Waal, but then proposes a twist on veneer theory itself: perhaps the veneer of morality is not just culturally grounded, but is genetically ingrained (p. 96). After all, De Waal himself admits that the building blocks of human morality are to be found in the widespread (and genetically based) processes of kin selection and reciprocal altruism. Of course, whether or not this “third way” between cultural and genetic determinism can truly be thought of as “vener” (something I suspect would find De Waal in deep disagreement) is another matter, but the possibility is certainly worth entertaining.

I find Christine Korsgaard’s attempt to defend the idea that there is a significant leap between chimp and human morality both unnecessary and rather out of step with the available science. It is unnecessary because De Waal has most certainly not denied that there are large differences between humans and our closest relatives, both emotionally and, more important, cognitively. Indeed, I do not know of any biologist that has even come close to that position. However, just because the distance is great, this does not invalidate De Waal’s main point that the building blocks for human morality have been put into place during a process of biological evolution. It seems that Korsgaard makes the same mistake for which Darwin was famously chastised by Huxley: evolution, as a natural process, has to proceed in steps, but there is no logical necessity for these steps to be small in size and very large in number. Korsgaard’s defense of the uniqueness of humans is also increasingly out of sync with the science of human evolution, arching back to a Kantian idea of morality that assumes that “the capacity for normative self-government and the deeper level of intentional control” (p. 116) is unique to human beings. Except that there is mounting evidence from the biological literature, from game theory (Nowak et al. 2000) to neurobiology (de Quervain et al. 2004) that our much vaunted self-consciousness (and intentional control) may be akin to a neural rationalizer of decisions already taken at the subconscious (and closer to the emotional) level. There may be as little room in moral theory for a self-determining Kantian agent as there seems to be in economics for the mythical ultra-rational *Homo economicus*.

Philip Kitcher, like Wright above, seems to feel that De Waal’s demolition of veneer theory is a bit too easy, suspecting that it is in fact a straw man. Indeed, in parallel with Wright, Kitcher also suggests a polar opposite to veneer theory, which he calls “solid-to-the-core” theory, which claims that morality was present in its essential components in our ancestors. He then, again, like Wright, is more attracted by some sort of third way, suggesting that “all the interesting positions lie somewhere in between” (p. 124). Kitcher acknowledges the importance of De Waal’s distinction between biological altruism (which is an oxymoron, according to evolutionary theory) and psychological altruism, which is really where the debate about morality begins (although, to be fair, Sober and Wilson 1998, have already expounded at a great length on that very distinction). He then goes on, however, to state (p. 129) that “our evolutionary relatives belong somewhere in altruism space away from the point of complete selfish indifference. Until we have a clearer view of the specific kinds of psychological altruism chimpanzees ... display, and until we know what kinds are relevant to morality, it’s premature to claim (as De Waal does) that human morality is a ‘direct outgrowth’ of tendencies these animals share.” Well, perhaps we can

then talk of human morality as an indirect outgrowth of other primates' tendencies. Unfortunately, however, Kitcher, does not attempt to sketch what other venues of inquiry are possible for the origin of human morality, certainly in part because of limitations of space. But he would probably agree with me (and De Waal) that whatever shape such an attempt may take, the only game in town is some kind of evolutionary scenario.

Peter Singer is the commentator that comes closest to accepting De Waal's arguments as presented, not the least because Singer himself made the point more than two decades ago that the origins of human morality are to be sought into our evolutionary past (Singer 1981). Where Singer distances himself from De Waal is on the claim (which Singer attributes to De Waal, but that I have not actually explicitly encountered in the latter's writings) that "all of human ethics derives from our evolved nature as social mammals" (p. 142). Of course Singer is right, and one of his analogies drives the point home very nicely: "An ability to count can be useful, but it leads by a logical process to the abstractions of higher mathematics that have no direct payoff in evolutionary terms" (p. 146). Indeed, and I would hope that the reasonable goal here is to create a consilience among disciplines, including biology, philosophy and psychology, as to what morality is and where it comes from—not a simplistic E.O. Wilson-style biologization of the subject matter. As I mentioned earlier, Singer then proceeds to discuss the implications of De Waal's arguments for animal rights, eventually agreeing (p. 154) that one could meaningfully adopt De Waal's suggestion of talking about "obligations" to other animals, rather than using the much more loaded term "rights." Indeed, De Waal's terminology in this respect is a direct consequence of his evolutionary view of inter-specific relationships.

#### Back to primates

De Waal, in his response at the end of the book, rather correctly points out that "whereas my respected colleagues focus on what seems missing rather than present in other primates, my own emphasis has rather been on shared characteristics" (p. 161). A more stereotypical example of the difference between philosophers and scientists is hard to find, and perhaps De Waal is being a little uncharitable here. Moreover, he immediately moves to make the exact same mistake mentioned above of assuming that evolution requires gradual change. It does not, and this should by now be abundantly clear from the biological literature (e.g., Schlichting and Pigliucci 1998).

Be that as it may, De Waal proposes an interesting model that bears some resemblance to Singer's "expanding circle" (p. 164), but is rather analogous to the famous floating pyramid of needs in economics. The basic idea is that our first moral duty (sanctioned by our biological urges) is toward ourselves, followed by the family and clan (kin selection), the community and tribe or nation (ingroup), humanity at large (outgroup) and finally all life forms (extended outgroup). Which levels of the pyramid are visible (i.e., which moral duties are actually carried out) depends on the conditions, so that it may require a society with ample resources and sophisticated moral reasoning to really care for all life on earth. Obviously, even twenty-first century humanity hasn't gotten to that level yet.

De Waal here blurs, possibly purposely, the boundaries between biological moral instincts and reasoned moral discourse, stating rather bluntly that "it is not just that we are biased in favor of the innermost circles (ourselves, our family, our community, our species), we *ought* to be. Loyalty is a moral duty" (p. 165, emphasis in the original). He goes on a few sentences later to add: "measured along the dimension of kinship, bonding, and group membership, an intellectually disabled human does indeed possess greater moral

value than any animal.’’ I am not certain what Singer’s reaction to this sort of reasoning would be, but most philosophers will likely regard it as a crude attempt at, again, over-biologizing morality.

Then again, once I overcame my own initially skeptical reaction to De Waal’s suggestion, I began to realize that it is hard to mount an effective reply to it. He is not saying that all morality reduces to biological instincts, but he is saying that ‘‘a viable moral system rarely lets its rules get out of touch with the biological imperatives of survival and reproduction’’ (p. 163). How un-Kantian of him. If one wishes to disagree, then one should put forth a coherent view of where any additional ‘‘ought’’ comes from and why it is justified.

This begins to sound suspiciously like what is often referred to as the naturalistic fallacy. As Hume (1739) famously put it in the *Treatise of Human Nature* (III(i)1): ‘‘In every system of morality, which I have hitherto met with, I have always remark’d, that the author proceeds for some time in the ordinary way of reasoning ... when of a sudden I am surpris’d to find, that instead of the usual copulations of proposition, *is*, and *is not*, I meet with no proposition that is not connected with an *ought* or an *ought not*. This change is imperceptible, but is, however, of the last consequence. For as this *ought* or *ought not*, expresses some new relation or affirmation, ‘tis necessary that it shou’d be observ’d and explain’d; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it.’’ Given the typically understated ironic tone, it has been notoriously open to interpretation whether Hume meant that the transition from *is* to *ought* simply cannot and should not be made, or that it can be made, but it requires some work. I do not know if De Waal read Hume, but the primatologist’s argument may be taken as the beginning of an answer to the philosopher’s demand. Whether it is a good answer remains to be seen.

## References

- Boyd R (2006) The puzzle of human sociality. *Science* 314:1555–1556
- Clayton P, Schloss J (2004) Evolution and ethics: human morality in biological and religious perspective. William B. Eerdmans Publishing Co., Grand Rapids, pp 339
- Darwin C (1871/1982) *The descent of man and selection in relation to sex*. Princeton University Press, Princeton
- Dawkins R (1976) *The selfish gene*. Oxford University Press, New York
- de Quervain DJ-F, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, Buck A, Fehr E (2004) The neural basis of altruistic punishment. *Science* 305:1254–1258
- Freud S (1930/1961) *Civilization and its discontents*. Norton, New York
- Gibbard A (1997) Wise choices, apt feelings. In Darwall, S, Gibbard, A, Railton P (eds) *Moral discourse and practice: some philosophical approaches*. Oxford University Press, New York, pp 179–198
- Greene JD, Somerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293:2105–2108
- Hamilton WD (1963) The evolution of altruistic behavior. *Am Nat* 97:354–356
- Hobbes T (1651) *Hobbes’s leviathan*. Clarendon Press, Oxford (preprinted from the ed. of 1651)
- Hume D (1739) *A Treatise of Human Nature*. Clarendon Press, Newyork, 40/1978
- Huxley TH (1894/1989) *Evolution and ethics*. Princeton University Press, Princeton
- Kropotkin (1902/1972) *Mutual aid: a factor of evolution*. New York University Press, New York
- Milgram S (1974). *Obedience to authority: an experimental view*. Harper & Row, New York
- Nowak MA, Page KM, Sigmund K (2000) Fairness versus reason in the ultimatum game. *Science* 289:1773–1775
- Pigliucci M (2003) On the relationship between science and ethics. *Zygon* 38:871–894
- Schlichting CD, Pigliucci M (1998) Phenotypic evolution, a reaction norm perspective. Sinauer, Sunderland
- Singer P (1981) *The expanding circle: ethics and sociobiology*. Farrar, Straus & Giroux, New York
- Skyrms B (2000) Game theory, rationality and evolution of the social contract. *J Consciousness* 7:269–284

- 
- Snow CP (1959/1998) *The two cultures and the scientific revolution*. Oxford University Press, Cambridge
- Sober E, Wilson DS (1998) *Unto Others: the Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Cambridge
- Williams GC (1988) Reply to comments on ‘‘Huxley’s evolution and ethics in sociobiological perspective’’. *Zygon* 23:437–438
- Wilson EO (1998) *Consilience: the unity of knowledge*. Knopf: distributed by Random House, New York