

Rationality

Christine M. Korsgaard

Having Reason and being Rational

According to a traditional philosophical view, dating back to Aristotle and shared by Immanuel Kant and many others, what makes human beings different from the other animals is that human beings are “rational” or have “reason.” The plausibility of this claim depends, of course, on what it is taken to mean. Both colloquially and in academia, the terms “reason” and “rational” are used in many different ways. “Reason” may be used to denote sanity, as when people say that a mentally deranged person has “lost his reason.” “Rational” may be used to mean orderly and intelligible, or in accordance with causal laws, as when people say they believe that the world is a “rational place.” “Reasoning” may be taken to describe the activity of working out what to do or believe by thinking, in which case “reason” is simply the capacity to do that. Charles Darwin uses the terms “reasoning” and “reason” in this way in *The Descent of Man* when he says:

Few persons any longer dispute that animals possess some power of reasoning. Animals may constantly be seen to pause, deliberate, and resolve. And it is a significant fact that the more the habits of any

particular animal are studied by a naturalist, the more he attributes to reason and the less to unlearnt instincts.¹

In this passage, “reason” is opposed to “instinct.” Actions from instinct are unlearned, automatic behavioral responses. When the terms are used in this way, acting from reason is acting intentionally and intelligently. Although some people have believed that all animal behavior is governed by “instinct” in Darwin’s sense, the modern study of animal minds supports Darwin’s nonchalant dismissal of that view. Corvids, our fellow primates, and cetaceans, among others, have demonstrated impressive problem-solving abilities that could not possibly be instinctive. And in many animals, behaviors that are probably instinctive are refined by learning: animals as they grow up get better at doing the things they instinctively do.

Social Scientific uses of “Reason” and “Rational”

Some uses of “reason” and “rational” are more specific than the ones described above. In the social sciences, acting “rationally” usually means acting prudently (doing what is in your own best interests)² – or acting with instrumental rationality (doing what will get you whatever ends you wish to achieve, whether they

¹ Charles Darwin in *The Descent of Man and Selection in Relation to Sex*. Princeton, New Jersey: Princeton University Press, 1981. p. 46.

² Although being “rational” in the social scientific sense is acting in your own bests interests, being “reasonable” in ordinary parlance sometimes means almost the opposite: being unwilling to impose undue burdens on others for your own sake.

are in your best interests or not). When we say that someone acts “rationally” in these senses, we usually also mean that she is subjectively guided by the standards of prudence or instrumental rationality. That is, we mean that she is motivated to do what she *believes* is in her own best interests or will get her whatever ends she wishes to achieve, and she is motivated by the fact that she believes it. This subjective dimension to the idea of being rational is worth emphasizing. If we take “rationality” to be a purely objective standard (one that does not refer at all to the subjective states of the agent), and we take survival and reproduction to be the ends of life, most animal behavior will be “rational” in the sense that it tends to achieve these ends. But, of course, this need not mean that the animals themselves grasp or employ the standards of prudence or instrumental rationality, or even that they are deliberately aiming at these ends. Evolution itself produces behavior that efficiently promotes these ends.

In general, it is important to distinguish between the evolutionary function of an animal’s behavior and what, from the animal’s own point of view, she is trying to do, if indeed there is anything that she is trying to do. In the case of automatic instinctive behavior, especially in animals not otherwise cognitively sophisticated, perhaps there is not. But when we consider more intelligent animals, there are real questions about how their own motives and intentions are related to the evolutionary function of their behavior, and it is easy to blur the line between these two things. For example, it is often suggested that the evolutionary function of altruistic behavior is to increase an animal’s own reproductive fitness in some indirect way. But this does

not mean that an animal who helps another does so with the intention of increasing her own reproductive fitness, or that there is something secretly self-interested about her conduct.

But as I have already noted, not all instrumentally effective animal behavior is instinctive and automatic. Intelligent animals do figure out how to do things that instinct does not teach them, and there is no reason to doubt that, from their own point of view, that is exactly what they are trying to do: figure out how to achieve their ends. A chimpanzee piles up boxes to reach the fruit hanging far above him; a cow learns to open the latch on her cowshed so that she can open the door and get out; a scientist puts obstacles between a raven and a food treat that the raven figures out how to overcome. But philosophers in the Kantian tradition would argue that even when an animal does engage in intelligent instrumental thinking – that is, when she figures out how to achieve a desired end – that does not necessarily show that her motivation is “rational.” These philosophers argue that rational motivation involves the awareness that the consideration on which you act *is a reason* for acting that way. When we say that an agent has a reason for what she does, we imply that there is a standard of evaluation for her action – the action is in some way “reasonable” or “rational” – and that the agent has gone some way towards meeting that standard.³ So knowing that you yourself are acting for a reason means knowing that an evaluative

³ “Gone some way towards” because you might have *a reason* for what you do while still doing something that is not what the balance of reasons favors.

standard applies to your conduct, that there is a way you should act or ought to act or that it is good or correct to act, and being motivated in part by that awareness.⁴ An animal who has figured out that taking a certain means will get him an end that he wants might be moved by that conclusion to take the means, without any thought about the normative correctness of being motivated in that way. We will come back to this conception of rationality later on.

Philosophical Conceptions of Reason and Rationality

In the philosophical tradition, “reason” is often taken to refer to the active as opposed to the passive or receptive aspects of the mind. “Reason” in this sense is contrasted with perception, sensation, and emotion, which are thought of as forms of passivity, or at least as involving passivity. The contrast is not unproblematic, for it seems clear that the kind of receptivity or responsiveness involved in sensation, perception, and emotion, cannot be understood as wholly passive. The perceived world does not simply enter the mind, as through an open door. In sensing and responding to the world our minds interact with it, and the activity of our senses themselves makes a contribution to the character of the world as we perceive it. All of this is undoubtedly true of the minds of the other animals as well. In fact, if, as

⁴ For a discussion of this view of rationality and the view of reasons that it requires, see Christine M. Korsgaard, “Acting for a Reason,” in *The Constitution of Agency* (Oxford: Oxford University Press, 2008), Chapter 7, pp. 207-229.

people usually suppose, non-human animals are in some sense “less intelligent” than human beings, it is all the more important that their perceptions should present their environment to them in ways that are already practically interpreted for them. An animal who cannot figure out what to do through reasoning must perceive the world in ways that guide his behavior without further thought – he must perceive the world around him as consisting, say, of predators who are to-be-avoided and prey who are-to-be-eaten and potential mates who are to-be-courted and offspring who are to-be-cared-for and so on. His perceptual faculties must do the work of thinking for him: in that sense, they shape the inputs of perception. But that is not something that the animal himself, considered as a conscious subject, deliberately does.

The mental activity associated with reason is activity that is attributable to the reasoner himself. Reasoning is self-conscious, self-directing activity through which we deliberately give shape to the inputs of receptivity. This happens both in the case of theoretical reasoning, when we are constructing a scientific account of the world, and in the case of practical reasoning, where its characteristic manifestation is choice based on deliberation. An animal who is rational in this sense exerts a kind of deliberate control over her own mental processes, at least those that issue in belief and action. It is the possession of reason in this sense – the exercise of deliberate control over our own mental lives – that it is most plausible to suppose is a distinctively human attribute.

Reason has also traditionally been identified with the ability to grasp, employ, or simply conform to, certain principles, which are usually conceived as *a priori*. These include the principles of logic, the principles that guide the construction of scientific theories, mathematical principles, and the principles of practical reason, including the principles of prudence and instrumental rationality mentioned in connection with the social scientific view described above. According to some philosophers, moral principles are also rational principles. In Kant's theory, for example, we evaluate our reasons by applying the categorical imperative – by asking ourselves whether the principles in which those reasons are embodied could possibly serve as universal laws. This conception of rationality is not at odds with the conception of reason as the active dimension of the mind: the principles of rationality may be taken to describe the activity in which the reasoner consciously engages when she considers whether her reasons are good ones. Even when we do not deploy these principles consciously, however, they are the principles in accordance with which we reason, the principles that describe how we think out what to do or believe. A person is called reasonable or rational when his beliefs and actions conform to the dictates of those principles, or when he is subjectively guided by them.

Reason is also identified with the capacity that enables us to identify “reasons,” the particular considerations that count in favor of belief or action. Reasons, in this sense, may be the considerations picked out by rational principles, or may be directly grasped by “reason” in the sense of the active capacity of the mind. Ordinarily, we

take reasons both to justify and to explain the actions for which they are done, or at least to make them intelligible. When we understand what an agent's reasons are, we see how the situation looked to him, and why it motivated him to act as he did. In this sense of "reason" it is plain that the other animals, like human beings, sometimes act for reasons and sometimes do not. Suppose that an animal is banging himself against the bars of his cage and we ask why he is doing that. If the answer is that he is trying to escape from the cage by breaking or bending the bars, we can see what his reason is, in the sense that we have some grasp of how the situation looks to him and what he intends to do about it. On the other hand, if the answer is that being caged has made the animal mentally ill, his behavior has a cause, but it is not something done for a reason.

Reason as a Distinctively Human Attribute

No doubt some who have believed that only human beings are "rational" or have "reason" have thought that all of animal behavior is automatic, unthinking, and instinctive. But that is not the only form the view can take. Philosophers who think that only human beings are "rational" or have "reason," may believe something along these lines: Many animals may have the capacity to think about their situations, and to be motivated to act intelligently in those situations to get what they want and avoid what they don't. But intelligence is not the same as reason. Only human beings have the capacity to think about their motivations themselves – about the potential reasons

for their actions - and to ask whether those potential reasons meet certain normative or evaluative standards. Human beings have a distinct form of self-consciousness that enables us to be aware of the motivations or potential reasons on which we are tempted to act, to evaluate those potential reasons, and to be moved to act accordingly. A parallel point may be made about the considerations on the basis of which we are tempted to believe. The standards of evaluation for our potential reasons are given by the principles of reason or rationality.

According to these views, non-human animals may have reasons for what they believe and do in the sense that they believe and act on the basis of the contents of their perceptions, desires, fears, and instincts. But they do not think about these reasons themselves, or the fact that they have them, and they do not ask themselves whether their reasons are good ones or not. Human beings, however, do, and this gives our beliefs and actions a normative character that the beliefs and actions of the other animals lack. If you are rational, when you believe something you also think that there is a sense in which you *should* believe it; when you do something, you think there is some sense (not necessarily a moral one) in which it is the appropriate or correct thing to do.

The theory that rationality in this more demanding sense marks a distinctive difference between human beings and the other animals will be most tempting if we can appeal to it to explain other things that appear to be distinctive about human life. For example, we might suppose that the fact that human beings think about the

quality of the reasons for our beliefs is part of what enabled us to develop the scientific method. Many people believe that morality itself a distinctive characteristic of human beings. If morality is a manifestation of a form of rationality that is unique to human beings, that would explain why that is so.

Some would argue that the idea that *any* capacity is unique to human beings is at odds with the theory of evolution. Darwin himself, who believed that morality is a distinctively human attribute, explained it as a result of the interaction between two capacities we share with some other animals: social instincts, and advanced cognitive faculties.⁵ Evolutionary thought need not be taken to imply that all differences between human beings and the other animals must be matters of degree. Some distinctively human attributes might emerge from unique combinations of evolved powers. The apparently immense difference in “intelligence” between human beings and the other animals might itself be better explained by the interaction of intelligence with rationality, than as a simple quantitative leap. I have already suggested that rationality is tied to a special form of self-consciousness, one that makes us aware of, and capable of evaluating, our reasons themselves. If this is so, perhaps rationality

⁵ Darwin says, “I fully subscribe to the judgment of those writers who maintain that of all the differences between man and the lower animals, the moral sense or conscience is by far the most important” (p. 70) and that “we have no reason to suppose that any of the lower animals has this capacity.” (pp. 88-90) But he proposes that “...any animal whatever, endowed with well-marked social instincts, 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.” (pp. 71-72) *The Descent of Man*, already cited.

itself can be explained as a result of the interaction between having an advanced “theory of mind” – the ability to attribute mental states to ourselves and others – and having other advanced cognitive capacities. These powers combined might enable us to think about what our own point of view and the points of view of others have in common, and to realize that those common elements set evaluative standards for our actions and beliefs.

Rationality and the Value of Human Life

Many people hold that rational beings or rational life has a greater value than non-rational beings or non-rational life. They think that this value makes rational beings either uniquely morally important, or more morally important than the other animals. This view is naturally associated with the idea that it is rationality that makes us moral beings, since that is an attribute that might appear to give rational beings a special value. But it is unclear exactly how the argument is supposed to go. Even if the capacity for moral goodness is a form of superiority, the fact that human beings alone are moral animals, if it is a fact, also makes us uniquely capable of moral evil. Perhaps the thought that rationality conveys a unique or superior value originates with the idea that rationality is one of the attributes in virtue of which human beings can be said to be made “in the image of God.”⁶ Without putting it on this theological footing,

⁶ For a protest against this line of thinking, see J.M. Coetzee, in *The Lives of Animals* (Princeton: Princeton University Press, 1999), pp. 22-23.

Immanuel Kant held that rational beings as “persons” are “ends in themselves,” while non-rational animals are “things” which may legitimately be used as mere means to the ends of rational beings.⁷ Many writers in the philosophical literature assume that what Kant meant is simply that rationality has a kind of intrinsic value which in turn confers value on those who have it. This does not make it clear why only rational beings should be the objects of moral consideration, however, and philosophers in the utilitarian tradition protest that sentience – the capacity for suffering – is sufficient to make a creature the object of moral concern.

A more promising argument for the conclusion that only rational beings are worthy of moral consideration, and one closer to the one Kant actually had in mind, holds that morality is a system of *reciprocal* rights and obligations. You cannot have the rights unless you are also capable of being meeting the obligations. Kant argues that to have a duty to another is to be constrained by that other’s will.⁸ As members of a “Kingdom of Ends,” a community of all rational beings, we all legislate moral laws for ourselves and each other and in this way exercise moral constraint over each other.⁹ But the other animals, who are incapable of willing moral laws, are no part of this system. Aristotle seems to have this kind of argument in mind as well. In the

⁷ Immanuel Kant, *Groundwork of the Metaphysics of Morals*, translated by Mary Gregor. Cambridge, UK: Cambridge University Press, 1997, p. 37, 4:428 in the Prussian Academy Edition page numbers found in the margins of nearly all translations.

⁸ Immanuel Kant, *The Metaphysics of Morals*, translated by Mary Gregor. Cambridge, UK: Cambridge University Press, p. 192, 6:442.

⁹ Kant, *Groundwork of the Metaphysics of Morals*, p. 42-43, 4:435.

Nicomachean Ethics, he argues at one point that neither justice nor friendship can exist between a (free) human being and a horse or an ox, nor a slave *qua* slave, but that a free human being can be friends or enter into relations of justice with the slave *qua* human. He explains, “for there seems to be some justice between any man and any other who can share in a system of law or be a party to an agreement.”¹⁰ Both philosophers conclude that non-human animals have only an instrumental value, but that conclusion is supposed to follow from the other animals’ inability to participate in the moral system, rather than to be the reason for it. Whether this is a good reason for excluding non-rational or non-moral creatures from moral consideration is of course a matter for further argument.

Another possible source of the idea that rational life is more morally important than non-rational life is an argument that goes roughly like this: (1) Only rational animals are self-conscious. (2) Therefore only rational animals can think about themselves and their temporally extended lives. (3) Therefore only rational animals really have genuinely temporally extended identities. (4) Not much except local pleasures and pains can be important to you if you lack a temporally extended identity. And therefore (5) not much can be very important to a non-rational animal. According to this argument, non-rational life is less morally important because the lives of non-

¹⁰ Aristotle, *The Nicomachean Ethics* Book 8, Section 11, lines 1161b1-1161b8. Translated by W. D. Ross and revised by J. O. Urmson. In Volume 2 of *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton: Princeton University Press, 1984.

rational animals are less important to those animals themselves.¹¹ All of the steps in this argument can be challenged, of course, but there is also an element of truth in some of them. Rather than taking the argument to ground a blanket distinction between rational and non-rational life, we might take it to explain why some things, such as the continuance of life, might legitimately be taken to matter more to animals who are cognitively and emotionally sophisticated in certain ways than they do to others.¹²

Normative and Descriptive Conceptions of Rationality, Rational Defects, and Marginal Cases

Those who believe that rationality is not what makes human beings worthy of moral consideration often point out that some of the human beings to whom we extend moral consideration are “not rational.” We extend moral consideration to infants, the comatose, the demented, and those with severe mental illnesses or cognitive defects. This is sometimes taken to show that it is actually “speciesism” – a prejudice in favor of our own species – and not a view about the special value of rational life, that motivates people to withhold moral consideration from non-human

¹¹ J.M. Coetzee, in *The Lives of Animals* (op. cit.), puts this view into the mouth of a fictional philosopher, Thomas O’Hearne, on p. 64, and challenges it through the reflections of his central character, Elizabeth Costello.

¹² For a more detailed treatment of this kind of argument, see Jeff McMahan, *The Ethics of Killing: Problems at the Margin of Life* (Oxford: Oxford University Press, 2002), Chapter 3.

animals. This argument, sometimes called the argument from marginal cases, raises questions about what we mean when we say of a being that he or she “is rational.”

A first point to notice is that the terms “rational” “reasonable” “reason” and their cognates may be used in either what we may call a descriptive or a normative way. When we say that someone “is rational” in the descriptive sense, we mean that there is some consideration which she believes to be relevant and on the basis of which she acts or believes, that she knows what that consideration is, and that she would offer it in answer if you asked her why she believes or does what she does. Being “rational” in this sense contrasts with being deranged, or out of control, or engaging in some form of expressive action – slapping someone in rage, screaming in terror – or acting in an automatic, unreflective way – ducking, wincing, or doing something entirely out of habit. In these kinds of cases we can explain what the person does, but the explanation does not take the form of citing a consideration on the basis of which she made a decision or formed a belief. When we say that someone “is rational” in the normative sense, we mean that she is rational in the descriptive sense *and* that the consideration of the basis of which she acted or formed a belief is a good reason, measured by the evaluative standards by which rational beings judge the quality of our reasons. When we use the terms normatively, we are prepared to call a consideration “a reason” only if it is a good reason.

Suppose, for example, that Edna tells us she plans to vote for a certain presidential candidate whom we know to be racially prejudiced, unintelligent, ignorant

of foreign policy, on the take from lobbies, etc., and we ask her why. She says he is the nominee of the XYZ party, and she always votes with the XYZ party. Edna is rational in the descriptive sense: her action is based on a consideration that she takes to be relevant, that the candidate has been nominated by her party. We may pose an objection to her plan either by saying “that’s a bad reason to vote for him” or “that’s no reason to vote for him.” If we say it’s “a bad reason,” we are using the term “reason” in the descriptive sense: a consideration on the basis of which someone acts. If we say it’s “no reason” we are using the term “reason” in a normative sense, essentially to mean a good reason. Someone can reason badly and still be “a rational being” in the descriptive sense.

There is one kind of case that cuts across the distinction in an interesting way: cases in which someone believes or acts on the basis of a consideration, but one whose efficacy for her she is not aware of and whose relevance she might consciously deny. We think of Freudian slips, prejudices, implicit biases, and various forms of self-deception as working in this way. Perhaps Edna is in fact drawn to vote for the XYZ candidate because he shares a prejudice of hers or because the opposing party’s candidate is a member of a racial minority that Edna unconsciously views as inferior. A critic of Edna might say that that is “the real reason” she is voting for the XYZ candidate. The consideration is in a sense operating as a cause, rather than as a reason, because Edna is not aware of its influence, but it makes her conduct intelligible in

something like the way a reason does: it shows us how the situation looks through her eyes, and how her action is a response to that.

We generally assume that being rational in the descriptive sense puts someone in the way of being rational in the normative sense; if you can reason, you can figure out or at least be taught how to reason well. But the categories of persons singled out in the argument from marginal cases, for various reasons, are not capable of reasoning well, and may not be capable of reasoning at all. Are they therefore “not rational beings” even in the descriptive sense?

The trouble with drawing this conclusion is that rationality is not merely a property, like having blue eyes. It is a way of functioning – a way of determining what to believe and to do. More generally, an organism is not merely a collection of properties, but a functional unity, whose various parts and capacities work together to produce a certain form of life. A rational being who lacks some of the properties that together make rational functioning possible is not non-rational, but rather defectively rational, and therefore unable to function well. For example, a small child or someone with severe cognitive disabilities may be aware of the desires that tempt her to act, and capable of refraining from acting on them if she knows that it would be wrong to do so, but unable to evaluate them correctly.¹³ An addict, on the other hand, may be able to evaluate her desires correctly, but is unable to refrain from acting on

¹³ For a more sophisticated view, along these lines, of the rationality of children see Tamar Schapiro, “What is a Child?” *Ethics* 109 (4): 715-738 (1999).

them. A comatose person does not have a different way of functioning, but rather is unable to function at all. It is not as if you could simply subtract “rationality” from a human being, and you would be left with something that functions like a non-human animal. What we owe to defectively rational persons is not to treat them however it is right to treat non-human animals, but rather duties of care that are responsive to the defects in their capacity to function rationally.

The point of advancing these considerations is not necessarily to object to the intended conclusion of the argument from marginal cases – that sentience is sufficient to make a creature the object of moral concern. Even if we agree with that conclusion, we may think that the argument from marginal cases makes morally significant errors about how the objects of moral concern should be identified. For example, categorizing “infants” as a type of creature conflates the subject of a temporal phase of a life with the subject of a life as a whole, which arguably is the proper object of moral concern.¹⁴ The subject of a life is not a mere collection of the subjects of various temporal phases of the life, but something that is connected in a certain way throughout the life. The way it is right to treat infants of every species is responsive not just to the properties they exhibit now, but to the fact that they are going to grow up in a certain way. In a similar way, whether a being is rational does not depend on the properties he exhibits now. A rational being is one “designed” to function in a

¹⁴ I borrow the phase “subject of a life” from Tom Regan, who identifies such subjects as the objects of moral concern. See *The Case for Animal Rights*. Berkeley: University of California Press, 1983.

certain way, and a creature so designed but with defects does not therefore become a different kind of creature. The proper moral response to such a creature reflects those facts.

Conclusion

“Rationality” refers to a range of properties that characterize the mental lives of the creatures deemed “rational.” It can mean simply sane and well-oriented towards the world, and it can mean capable of acting intentionally and intelligently. In these senses, non-human animals can be “rational.” Philosophers who suppose rationality is a unique human attribute or one that human beings have to a uniquely high degree usually mean something more specific. Rationality in this sense involves an awareness of the considerations that tempt us to believe and act in certain ways – of our potential reasons – and the ability to evaluate those reasons in accordance with *a priori* principles that determine whether they are good reasons or not.

Some people believe that rationality in this sense makes human beings uniquely worthy of moral consideration, or more morally important than the other animals. This may be because they believe that only human beings are moral animals and that morality itself is a system of reciprocity in which non-rational creatures cannot participate. Or it may be because they believe that the opposite of a rational mental life is a mental life without much temporal extension, and that therefore nothing but local experiences can be very important to non-human animals themselves. These

views can certainly be challenged, both empirically and philosophically. Even if rationality is a distinctively human attribute, it may be the source not of a special moral value or importance for human life, but rather of a different kind of special moral standing: human beings alone among the animals have moral obligations to the other animals with whom we share the world.¹⁵

For Further Reading

Aristotle, *The Nicomachean Ethics*. Translated by W. D. Ross and revised by J. O. Urmson. In Volume 2 of *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton: Princeton University Press, 1984.

J.M. Coetzee, J. M. *The Lives of Animals*. Princeton: Princeton University Press, 1999.

Darwin, Charles. *The Descent of Man and Selection in Relation to Sex*. Princeton, New Jersey: Princeton University Press, 1981.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*, translated by Mary Gregor. Cambridge, UK: Cambridge University Press, 1997.

¹⁵ A detailed explication of this position is forthcoming from Oxford: Oxford University Press, in Christine M. Korsgaard, *Fellow Creatures: The Moral and Legal Standing of Animals*.

Korsgaard, Christine M., "Acting for a Reason," Chapter 7 in *The Constitution of Agency*. Oxford: Oxford University Press, 2008.

Korsgaard, Christine M., *Fellow Creatures: The Moral and Legal Standing of Animals*. Oxford: Oxford University Press, forthcoming.

McMahan, Jeff. *The Ethics of Killing: Problems at the Margin of Life*. Oxford: Oxford University Press, 2002.

Schapiro, Tamar. "What is a Child?" *Ethics* 109 (4): 715-738 (1999).