



Agency and morality

Citation

Korsgaard, Christine M. "Agency and Morality." In The Routledge Handbook of the Philosophy of Agency. Edited by Luca Ferrera. New York: Routledge, 2022.

Published version

<https://doi.org/10.4324/9780429202131-49>

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37373854>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles (OAP), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

Agency and Morality

Christine M. Korsgaard

ABSTRACT: The traditional categorization of ethical theories into consequentialist, deontological, and virtue-ethical theories ignores the significance of the fact that philosophers supposedly representing these categories are working with different conceptions of what an 'action' includes, or what the object of an agent's choice is. It therefore overlooks the fact that philosophers like Kant and Aristotle believe that right actions are chosen for their goodness. These theories give rise to the question what makes an action good. Contemporary 'constitutivist' theories argue that the normative standards determining the goodness of action arise from the function of action, which is to constitute the efficacy and so the agency of those who choose them.

1. Action and the Structure of Moral Theory

According to a story widely accepted in the twentieth and twenty-first centuries, there are three main domains of value relevant to morality. Goodness in the general sense applies to the ends that are worth aiming at or realizing, the states of affairs we should bring about for their own sakes. Rightness applies in the first instance to actions, and by extension to policies when the actions which those policies dictate are right. And virtue applies to people and to their dispositions and

characteristics—specifically to the characteristics of people that incline us to do what is right.

According to a related and also widely accepted story, we can categorize different types of ethical theory according to which of these domains of value contains the considerations that determine what is right. ‘Deontologists’ believe that actions are right or wrong in themselves, intrinsically, and that our pursuit of the good must be constrained by the necessity of doing only what is right. ‘Consequentialists’ believe that actions are right when they tend to maximize the production of good consequences. ‘Virtue theorists’ believe that the right action must be identified as the one that the virtuous person would choose to do.¹

This story presupposes that all of these schools of thought mean the same thing by ‘action.’ Its simplicity falls apart when we begin to question that. Suppose that Jack drops his keys on the sidewalk, and Jill scoops *them up and runs after him, in order to ensure that Jack is in possession of his keys*. How much of that description is included in the correct description of Jill’s action?

According to a broadly empiricist account, favored by consequentialists, Jill has an end—ensuring that Jack is in possession of his keys—and in light of that end she chooses to do something, namely scoop up the keys and run after Jack. The end, on this account, is something that stands outside of the action and purports both to explain and justify it. Perhaps Jill desires the end because Jack is her friend and she cares about him, or perhaps she desires the end because she sympathizes readily

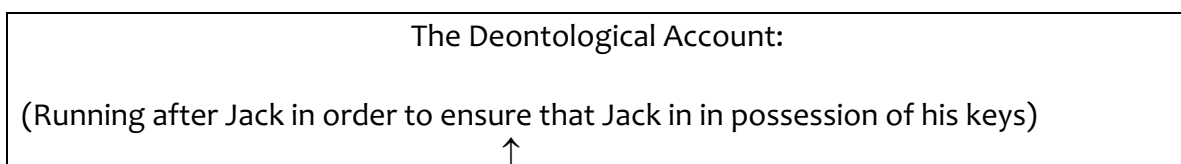
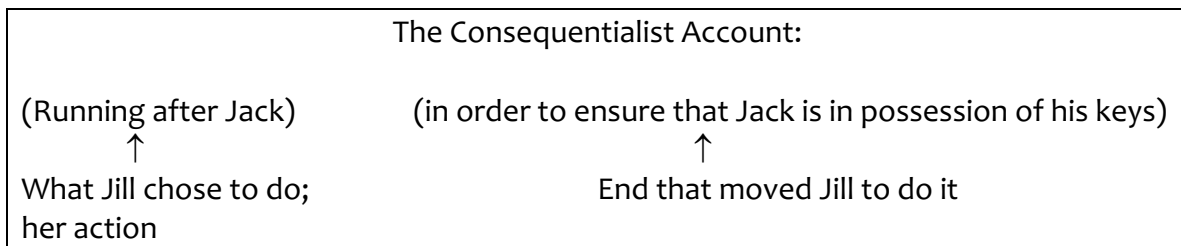
with anyone who loses their keys, or perhaps she holds the end as a matter of principle, believing that lost property should always be returned to its owner when possible. Usually, in the empiricist story, the end is not chosen, but if it is, it is chosen as a means to or an instance of some other end that in turn is not chosen. So the end is given to the agent, and she chooses an action that she thinks will bring about the end. According to consequentialists, whether the action is right depends on whether the end that it promotes is a good one, or perhaps on whether it is the best one that she could have achieved in the circumstances. As John Stuart Mill says, ‘All action is for the sake of some end, and rules of action, it seems natural to suppose, must take their whole character and color from the end to which they are subservient.’ (Mill 1979, p. 2)

What I am somewhat tendentiously going to call the ‘deontological’ account—because it is Kant’s, and Kant is usually taken to be the primary exemplar of deontology—differs from the consequentialist account in two ways. According to the deontological account, our ends are not given to us, but rather are chosen; and in particular, they are part of what we are choosing, when we choose an action. In other words, the object of Jill’s choice, and her action, is ‘to run after Jack in order to ensure that he remains in possession of his keys.’ Jill’s end may have suggested itself to her by desire or sympathy or principle, but she did not have to pursue it. No causal force such as natural desire impelled her to pursue it, nor did she choose it independently of the action, although she may have been on the lookout for actions

through which she might promote this sort of end. But the end is not in any way set for her when she decides what to do. Instead she makes a calculation according to which, given that Jack needs his keys, and given the possible moral and prudential costs of running after him, running after him in order to ensure that he has his keys is a thing worth doing, a good thing to do.

That is the second difference between the two accounts: on what I am calling the ‘deontological’ account, rightness is a form of goodness, a form of goodness that applies to actions, just as virtue is a form of goodness that applies to people. Or if it is not, our account of which actions are right is derived from an account of what makes an action morally good. I say this because we sometimes use the word ‘right’ for actions that involve the same outward act that morally good actions do, whether they are themselves morally good or not. We are using the term this way when we say that someone ‘does the right thing for the wrong reason,’ as we might if Jill, for example, ran after Jack in hopes of getting a reward for retrieving his keys.

So here are the two possibilities we have canvassed so far:



What Jill chose to do; her action What moved Jill to do it: the goodness of the action

To make it a little easier to talk about this, from now on I am going to call things like ‘running after Jack’ an ‘act’ and things like ‘running-after-Jack-in-order-to-ensure-that-Jack-is-in-possession-of-his keys’ an ‘action.’ Then we can say that according to the consequentialist account, what we choose are acts, while deontologists think what we choose are, in this technical sense, actions.

I’ve called the second possibility the ‘deontological’ account in spite of the fact that either a deontologist or a virtue ethicist might conceivably hold it. Kant and Aristotle are often supposed to represent deontology and virtue ethics, respectively, yet they are the two main exemplars of the kind of account I have just described. Both of them begin their accounts of what is right, in the sense of we ought to do, from an account of the moral goodness of actions. I’ve also called it the ‘deontological’ account in spite of the fact that it is not clear that everyone we would style a ‘deontologist’ holds it.² The fact is that in the tradition of moral philosophy, many philosophers have simply been unclear about what they mean when by an ‘action,’ and in particular about whether the end is included in the content of the action or not. When we turn to Aristotle and Kant, however, we find accounts according to which the thing that the agent chooses, and the thing that has moral value, is what I have been calling an ‘action,’ that is, an act-for-the-sake-of-an-end.

Kant believed that the sort of thing that an agent chooses, and therefore the sort of thing that is a candidate for being morally good or bad, is the sort of thing that we describe by giving what he called the agent's 'maxim.' Characteristically, the maxim, and therefore the action, includes both the act that the agent performs and the end for the sake of which he performs it. It has to include both of those, because the way that we test a maxim to determine whether it is in accordance with duty is by applying the categorical imperative, the moral principle that tells us to act only on those maxims which we can will to be universal laws (Kant, 1998, 4:421-424). The question raised by the categorical imperative is whether the agent making the choice could will that every agent who has *that end* should pursue it by *those means* at the same time as he himself wills to pursue *that end* by *those means*, without generating a contradiction in his will. If the act can be a successful method of achieving the end only if not everyone attempts to use it as a method for achieving that end, then the maxim and so the action is ruled out as wrong.³ So, for instance, Kant argued that if everyone borrowed (or rather tried to borrow) money on the strength of false promises, possible lenders would not accept promises as a basis for lending money, and the act of making a false promise would no longer work as a means of getting a loan (Kant, 1998, 4:422). An agent who willed that situation at the same time as he himself willed to get money by means of a false promise would produce a contradiction in his own will, because he would be willing a situation in which the method he proposes to use to achieve his own end would no longer work. On the

other hand, if the maxim of the action could serve as a universal law, it has what Kant calls the ‘form’ of a law—an idea I will come back to—and the action it describes is therefore permissible and to that extent morally good. When an action is rejected as impermissible, the opposite action is a duty, and is in a stronger sense morally good.

The parallel in Aristotle’s ethics to the Kantian maxim is the *logos*—a description of the action that contains or expresses the reason why it is done. Aristotle is more careful than Kant to ensure that all of the aspects of the action that might be relevant to its goodness are included in the *logos*. In a number of passages in the *Nicomachean Ethics*, Aristotle tells us that the good person acts on the right reason or right principle—the *orthos logos*—specifying that this means that the good person does the right act at the right time, in the right way, and for the right end (Aristotle, 1984. NE 2.9 1109a25–30). It is clear from these formulations that Aristotle thinks that the moral goodness of the action somehow rests in the way these factors are related to each other. Since the relation between the parts of something constitute what Aristotle calls its ‘form,’ a good action is one that is well-formed.⁴ Kant thinks that the way the parts of a maxim, namely the act and the end, have to be related to each other for an action to be well-formed is ‘so that the maxim can serve as a universal law.’ Aristotle is not specific about what exactly makes an action well-formed, but since Aristotle also describes the first principles of our actions as ‘universals’ (Aristotle 1984 NE 6.8 1141b26; NE 6.8 1142a12-21; NE 7.3 1147a105), it is tempting to think that he shares with Kant the view that ‘universality’ of some sort is

what gives an action the correct form, the form that makes it good or in Aristotle's language 'kalon' (noble).

The first difference between the consequentialist and what I am calling the deontological account, then, is that they identify different objects as the 'action' that may be right or wrong. For consequentialist, it is what I have called the act, the movement through which we promote the end, that is right or wrong. But on deontological account it is the whole action, the act-for-the-sake-of-the-end, that has moral value. The second difference concerns our motive for doing the right 'action.' On the consequentialist account, we know what moves Jill to do what that account identifies as her action, which on that account is the same as her act (running after Jack). It is her end: she is moved to run after Jack in order to ensure that Jack has his keys. On the deontological account, what is Jill's motive for the choice of what that account identifies as her action (running-after-Jack-in-order-to-ensure-that-Jack-is-in-possession-of-his-keys)? The answer is that the agent sees the action *as a whole* as something that is good to do. Because it is good in virtue of the relations between its parts, we can even say, as deontologists are supposed to say, that the action is *intrinsically* good.⁵ These contrasts explains how John Stuart Mill can say that 'all action is for the sake of some end' while Aristotle insists that 'while making has an end other than itself, action cannot; for good action is itself its end' (Aristotle 1984 NE 6.5 1140b6-7). But they leaves us with a question: what, according to Kant and

Aristotle, makes an action morally good? What makes an action well-formed, and why does an action that is well-formed count as a morally good action?

2. Constitutive Standards and the Goodness of Action

When we say that something is good, we usually mean one of three things: that it is good of its kind, that it is good for some specific purpose, or that it is good as an end. If I say without further elaboration that a certain shoe is good, you will assume that it protects your feet from the environment while being comfortable and supportive, that it does not produce blisters, that it makes you less likely to slip on tenuous surfaces, things like that. These standards of goodness are derived from the function of a shoe, and may therefore be identified as ‘internal’ or ‘constitutive’ standards for shoes: normative standards that arise from the nature of the object to which they are applied. A good X is one that serves the function of an X well, and the things that I have mentioned above, protecting your feet and your footing and so on, are the functions of shoes. If I say that a certain shoe makes a good doorstop, on the other hand, you will assume that it is massive and heavy enough to hold a door open, though massiveness and heaviness are not ordinarily especially good properties in a shoe. This is an ‘external’ standard, one we apply to the shoe only in virtue of some special purpose we have in mind for it. In both of these cases we are normally talking about things that are good as means to other ends—to the tasks of protecting one’s feet, the usual purpose of shoes, or holding a door open, a specific purpose in the

case at hand. Finally, if I say that happiness or virtue is good as an end, you will assume that these things are worth aiming at for their own sakes. Which of these things do we have in mind when we say that an action is morally good?

At first glance, it might seem to make better sense to talk about the goodness of acts than it does to talk about the goodness of actions. Acts do seem to have a general function—namely to produce ends. An act, after all, is not just a movement, but a movement undertaken by an agent, and an agent who undertakes an act usually—some would say always—has an end in view. Those who suppose that agents always act ‘under the guise of the good’ may even suppose that the agent always has some *good* end in view. If that is right, we might think that a good act is one that is effective, or even one that is effective for good.

A little thought, however, shows us why this cannot be quite right. All acts effect something or other, because all acts are movements embedded in the causal network. If nothing else, the movements of my body displace molecules in the air around me, but that doesn’t make those movements ‘effective’ in the sense we need here. When we say that an act is effective, we mean that it tends to effect the end that the agent had in mind to produce through his movement. So it is not a mere act, but only an act within the context of a whole action, that can be said to be effective. So there is at least one standard for good actions: a good action is one whose act tends to achieve the end of that action, the one the agent has in view. In other

words, the action must accord with what Kant calls the hypothetical imperative: the act is a means to the agent's end (Kant 1998 4:414-415).

You might wonder why I have said that the good action is one whose act 'tends to' achieve its end, rather than one that just *does* achieve its end. Isn't success better than a mere tendency? It is not, because the kind of efficacy that we associate with agency is not exhausted by the idea of the agent actually bringing the end about. In fact, an agent who brought about his end only accidentally—say by a deviant causal pathway—would have failed as an agent. If I fire my gun wildly astray, but the bullet ricochets off a cast iron fence and happens to hit the target in exactly the spot that I intended, I have not made a good shot. Indeed, once I understand the situation, I would feel that I have been the beneficiary of a lucky accident, and that I didn't really *do* anything at all. To be successful in action is not merely to do something that brings about your end. To be successful in action you have to use a means that makes you into *the kind of thing* that *reliably* achieves that sort of end. The efficacy that makes an action good is not just the efficacy of the act. It is the efficacy of the agent.

This in turn has a further implication. In order for an action to be efficacious in the right way, the efficacy that it exhibits has to be the efficacy of the agent *considered as an agent*. This implies not only that the means must tend to bring about the end, but that the action as a whole has to be attributable to the agent as such. The source of the action has to be the agent herself. It is obvious that it cannot just

be a force at work on the agent from the outside. If I knock Jack off a cliff by shoving Jill into him when she is standing behind him and he is on the edge, I make Jill's body into the sort of thing that reliably knocks people off cliffs, but she has not made herself into that, and her agency is not involved. It is perhaps less obvious but nevertheless also true that it also cannot be just a force at work on the agent from the inside, unless we have a quite particular conception of what counts as 'inside.' If Jill shoves Jack off a cliff under the influence of a twitch or a spasm or while sleepwalking or under the influence of a posthypnotic suggestion, there is a sense in which she has done the shoving herself, but it is not the right sense. The action must originate in a choice that somehow represents the activity of the agent herself.

Before we consider how that condition is met, let's take stock. According to the deontological account, actions, acts-for-the-sake-of-ends, may be morally good or bad. This must mean they are good of their kind, good for some special purpose, or good in themselves. In order for actions to be good of their kind, they must have a function, which good actions succeed in serving or serve well. Surprising as it may seem, we have now identified a function for actions. An agent who undertakes an action undertakes to achieve some end in the world through her own efficacy. The function of an action, therefore, is to make its agent efficacious for a certain end. In order to do that, the action has to meet two standards. First, the act must be related to the end in such a way that it makes the agent efficacious, that is, it must make the agent into the kind of thing that tends to achieve that end. And, second, the choice

must be related to the whole action in a way that ensures that it is the agent herself, and not some alien force working in her or on her or through her, that is in this way efficacious. Putting these two conditions together, she must *make herself* into the kind of thing that is efficacious for that end. An action that meets these two standards succeeds in making its agent efficacious, and therefore is good of its kind. It is good considered just as an action.

We have already seen how to meet the first of these two standards: an act is related to its end in a way that makes the agent efficacious if the action as a whole conforms to the what Kant called hypothetical imperative, the principle of taking the means to one's ends. How do we guarantee that it meets the second standard, that it is the agent's efficacy *as an agent* that is in this way established?

Aristotle argued that the action must have its source in the agent's character, his virtues and vices, in order for the action to be an expression of his own choice, and therefore of his own efficacy (Aristotle 1984 NE 3.2). That secures us part of the conclusion we want here: chosen actions are *by their very nature* morally good or bad. Moral goodness or badness is not some external standard that happens to apply to human actions, the way that the standards for serving as a good doorstop might happen to apply to a shoe. To put it another way, morality is not a special purpose we happen to have for actions. Actions have moral properties considered simply as actions.

But Kant's theory goes even further. According to Kant, an action has its source in the agent herself only if it has its source in the agent's autonomy, that is, only when she acts on principles which she herself recognizes, and so legislates to herself, as universal laws. But if the agent is thinking clearly, the only principles that she can recognize as universal laws are the principles that really can be universal laws. That means that only the morally good agent is genuinely efficacious, that is, only she makes herself genuinely efficacious through her actions. Her actions make her efficacious because they achieve the end she has in mind, and because achieving *that end* by using *that means* is a principle of action that she herself endorses. To put it more intuitively, she does what she herself thinks that she and other agents should do. It follows that only the morally good action is good as an action, good of its kind. Autonomy, which for Kant involves acting on moral principles, is a constitutive standard for actions, one that they must meet in order to be good of their kind.

It is a difficult question whether Aristotle also holds a view of this kind. It depends on whether his view allows us to see moral badness as a form of what Kant would have called 'heteronomy,' the enslavement of the agent to forces outside of her own will (Kant 1998 4:432-433). Aristotle's identification of *akrasia*, or moral weakness, as a condition that is different from moral badness complicates this question (Aristotle 1984 NE VII). But Aristotle's view raises another question we must deal with here. As I noted earlier, while Mill suggests that actions (by which he means

what I have called ‘acts’) are good or bad as means to good ends, Aristotle says that good actions are ends in themselves. What are we to make of Aristotle’s claim?

In Kant’s theory, a good action is one whose maxim can be a law, and a required action is one whose maxim must be a law, because the opposite maxim is impermissible. If the maxim of an action must be a law, then the action is one that it is necessary for the agent to do. In that sense, it is the nature of this action that it is something that is to be brought about for its own sake, and the things that are to be brought about for their own sakes are good. But I think a comparison that Aristotle often makes gives us a better way to put the point: the comparison between health and virtue. Healthy activity is a good thing for its own sake because there is a way in which our lives consist in physical activity, and healthy activity is the excellence of physical activity. In the same way, morally good action is a good thing for its own sake there is a way because in which our lives consist of rational, chosen, action, and morally good action is the excellence of rational chosen action. Doing what is right is good simply because it is a way of living well.

Related Chapters:

I.3 Agency, Functions, and Teleology

II.4 Rational Agency

VI.1 Agency, Desire, and Motivation

VI.2 Agency, Will, and Freedom

VI.3 Agency and Responsibility

VII. 2 Agency and Autonomy

IX. 1 Agency, Reasons, and Rationality

IX. 2 Agency and Practical Reasoning

IX. 3 Agency and Normativity

References:

Aristotle. (1984) *The Nicomachean Ethics* (translated by W. D. Ross, revised by J. O.

Urmson) in *The Complete Works of Aristotle: The Revised Oxford Translation*.

Edited by Jonathan Barnes. Princeton: Princeton University Press. References to Aristotle's works are given by the standard Bekker page, column, and line numbers.

Herman, Barbara. (1993) "Leaving Deontology Behind" in *The Practice of Moral*

Judgment. Cambridge, MA: Harvard University Press, pp. 208-240.

Kant, Immanuel. (1998) *Groundwork of the Metaphysics of Morals*. Translated by Mary

Gregor. Cambridge: Cambridge University Press. References to Kant's works are given in the standard way by the page numbers in the relevant volume of Kants gesammelte Schriften, which appear in the margins of most translation.

Korsgaard, Christine M. (1996) "Kant's Formula of Universal Law" in Korsgaard,

Creating the Kingdom of Ends. Cambridge: Cambridge University Press.

Mill, John Stuart. (1979) *Utilitarianism*. Edited by George Sher. Indianapolis: Hackett Publishing Company.

Tenenbaum, Sergio (ed.) (2010). *Desire, Practical Reason, and the Good*. Oxford: Oxford University Press.

For Further Reading:

1. The volumes below contain both classical accounts and contemporary essays on the three types of ethical theory.

Darwall, Stephen, editor. *Consequentialism*. Blackwell Publishers Ltd: 2003.

Darwall, Stephen, editor. *Deontology*. Blackwell Publishers Ltd: 2003.

Darwall, Stephen, editor. *Virtue Ethics*. Blackwell Publishers Ltd: 2003.

2. The book below defends the Kantian version of a constitutivist account of the moral goodness of action.

Korsgaard, Christine M. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press, 2009.

3. The books below defend another version of constitutivism, one in which the nature of action provides the standards for the rationality of action.

Velleman, J. David. *Practical Reflection*. Princeton: Princeton University Press, 1989.

Reissued by the Center for the Study of Language and Information, 2007.

Velleman, J. David. *The Possibility of Practical Reason*. 1st edition. Oxford: Oxford University Press, 2000. 2nd Edition, Michigan Publishing Services, 2015

Velleman, J. David. *How We Get Along*. Cambridge: Cambridge University Press, 2009.

¹ 'Must be identified' deliberately leaves it open whether the identity in question is metaphysical or epistemological. Later I will suggest that, at least in its Aristotelian versions, it can only be epistemological. That means that Aristotelian virtue theory is not actually an alternative to consequentialism and deontology, which are views about the metaphysics of the right.

² An alternative strategy is to deny that Kant is a deontologist. See Herman 1993.

³ For a defense of this interpretation of categorical imperative, see Korsgaard 1996.

⁴ If this is Aristotle's view, then he does not hold that what makes an action right is that the virtuous person would choose it. What makes it right is that it is well-formed. But Aristotle does hold that the virtuous person is the only reliable judge of whether an action is well-formed. See note 1.

⁵ A caveat is in order. What makes the action good on these accounts is the relation between its parts, and in that sense its goodness is intrinsic, but we cannot ascertain whether that relation holds simply by examining the action itself, as we might expect to do if the goodness is intrinsic. For instance, Kant's argument against false promising depends on the idea that if everyone made false promises, potential lenders would come to realize that fact. That obviously depends on causal relations that are not intrinsic to the action.

AUTHOR BIO:

Christine M. Korsgaard is the Arthur Kingsley Porter Professor of Philosophy at Harvard. She is the author of *The Sources of Normativity* (Cambridge 1996), *Creating the Kingdom of Ends* (Cambridge 1996), *The Constitution of Agency* (Oxford 2008), *Self-Constitution: Agency, Identity, and Integrity* (Oxford 2009), and *Fellow Creatures: Our Obligations to the Other Animals* (Oxford 2018).