

Class #27 - Kantian Deontology
Kant, from *Groundwork of the Metaphysic of Morals*

I. Good Will, Duty, and Inclination

The core claim of utilitarianism is that the morality of an act depends on its consequences. Consider an act such as going out of your way to help a struggling person cross the street. This is a good act.

It remains a good act even if it ends with an unforeseeable bad consequence, as in the case in which a previously-unseen truck runs over and kills both of you.

The utilitarian, given the bad outcome, must describe your act as bad.

But your intention is the same in both cases.

How could the same act be good in one case and bad in the other?

Kant's moral theory accounts for our intuition that the two acts are equally morally worthy. In both cases, our will, our desire to do our moral duty, is good.

It is impossible to conceive anything at all in the world, or even out of it, which can be taken as good without qualification, except a good will (Kant 536a).

Since the will in each case is the same, the moral worth of each action is the same. For Kant, morality is always autonomous, or independent of consequences.

A good will is not good because of what it effects or accomplishes - because of its fitness for attaining some proposed end: it is good through its willing alone - that is, good in itself (Kant 537a).

Utilitarianism makes morality heteronomous, involving factors outside of us and over which we have little or no control.

Kant claims that since morality is unaffected by consequences, it must be purely formal, determined by the agreement of one's maxims with one's duties.

Our first step toward understanding Kant's moral theory is to clarify his notion of the will, which he calls reason in its practical employment, and how it relates to our inclinations.

We have inclinations, or desires, as do all animals.

Our inclinations will naturally conflict in some cases with the inclinations of others.

We might want something that belongs to someone else.

We might be inclined to take it.

Kant will argue that it is our duty never to steal.

Let's grant that claim for the moment, assuming that stealing is wrong.

Still, Kant argues, refraining from stealing is not in itself praiseworthy.

We can not be said to have a good will merely by refraining from committing a bad act.

We might be acting in accordance with our duty for merely selfish reasons.

Consider saving some one's life with the expectation that you will receive a substantial a reward.

Perhaps you save Lady Gaga from drowning.

The moral worth of the action itself is unclear.

If you are doing it for reward, you may be pursuing your self-interest.

To have moral worth, Kant claims, an action must be done from the motive of duty.
To determine if an action is morally worthy, then, we have to determine the content of our will.
An action can only be seen as morally worthy if we are acting against our inclinations.
Otherwise, we might be acting for the wrong reasons.

Kant discusses four ways in which duty and inclination may meet.

- DI1. Acting contrary to duty
- DI2. Acting consistently with duty and with inclination
- DI3. Acting consistently with duty, but not with immediate inclination, though we might have some inclination
- DI4. Acting consistently with duty but contrary to inclination

Examples of DI1 include robbing, murdering, and lying.
We might be inclined to perform such acts, for personal gain, say.
Or, we might be disinclined.
In either case, such acts are not morally valuable.

To illustrate DI2, Kant discusses a shopkeeper charging a fair price.
The shopkeeper does no wrong.
But since the act is in his self-interest, we can not see the moral value in it.

Thus people are served *honestly*; but this is not nearly enough to justify us in believing that the shopkeeper has acted in this way from duty or from principles of fair dealing; his interests required him to do so. We cannot assume him to have in addition an immediate inclination towards his customers, leading him, as it were out of love, to give no man preference over another in the matter of price. Thus the action was done neither from duty nor from immediate inclination, but solely from purposes of self-interest (Kant 537b).

To illustrate DI3, consider sitting next to someone with a twenty-dollar bill falling out of his pocket.
Imagine that taking the bill would be easy, and you need the money.
You may be inclined to steal it, but refrain out of fear of being caught.
In this case, you again act in the right way but for the wrong reasons.

Thus it is only in the case of DI4 that our acts are clearly morally worthy.
Acts such as returning lost money or volunteering one's time to help others, if done for the right reasons, may qualify as morally good.
To qualify, though, they must not be done out of selfish motives.
Among selfish motives are the desire for the happiness that someone with a generous spirit might get from doing good work.

To help others where one can is a duty, and besides this there are many spirits of so sympathetic a temper that, without any further motive of vanity or self-interest, they find an inner pleasure in spreading happiness around them and can take delight in the contentment of others as their own work. Yet I maintain that in such a case an action of this kind, however right and however amiable it may be, has still no genuinely moral worth. It stands on the same footing as other inclinations... (Kant 537b-538a).

One might try to interpret Kant as holding merely the weak claim that we can see only see moral value in acts done contrary to inclination, though other acts of types D2 or D3 may also have moral worth.

We might see the weak interpretation as charitable to Kant.

But Kant is committed to the stronger claim.

Only acting contrary to inclination creates moral worth.

The weak claim is less controversial, but the strong claim is really Kant's position.

In fact, Kant claims that there may never be actions of moral worth.

Even when we think that we are acting purely out of duty, we may be misleading ourselves.

Kant has provided a strict standard for morality.

We will put aside for now questions about whether it is too strict.

The basic intuition to which Kant appeals, that the only good thing is a good will, seems fair enough, and worthy of pursuit.

A good will is the reasoned desire to do one's moral duty.

II. Categorical Imperatives

We assumed, for the sake of discussing the relation between duty and inclination, that we actually have moral duties and that we know what these duties are.

But all we have seen so far is that if there are any moral duties or acts, they would have to be independent of our interests or inclinations.

Thus, commands of morality would have to be categorical, independent of particular facts of a situation.

There is an imperative which, without being based on, and conditioned by, any further purpose to be attained by a certain line of conduct, enjoins this conduct immediately. This imperative is categorical. It is concerned, not with the matter of the action and its presumed results, but with its form and with the principle from which it follows; and what is essentially good in the action consists in the mental disposition, let the consequences be what they may. This imperative may be called the imperative of morality (Kant 540b).

Our sole duty is to obey the categorical imperative.

An imperative is a command.

A categorical imperative admits of no exceptions, a command which is independent of the consequences.

Kant describes three versions of the single moral law called the categorical imperative (CI), one rule in three supposedly-equivalent forms.

CI1. The formula of universal law

CI2. The formula of the end in itself

CI3. The formula of the kingdom of ends

Kant discusses the three versions of the CI without presuming the existence of any such imperatives.

We must never forget here that it is impossible to settle *by an example*, and so empirically, whether there is any imperative of this kind at all: we must rather suspect that all imperatives which seem to be categorical may none the less be covertly hypothetical (Kant 542a).

Kant shows that if there are moral acts, then they would have to conform to a categorical imperative.

Imperatives are categorical if they command independent of any consideration of consequences.

III. The Formula of Universal Law

CI1, the formula of universal law, like all categorical imperatives, refers to maxims for action rather than the content of an act itself.

Act only on that maxim through which you can at the same time will that it should become a universal law (Kant 542b).

A maxim is a general rule of which an action is an instance.

Maxims are generalized versions of one's intentions.

The maxims of my actions are of the form: Whenever I am in situation x, I shall do action y.

To fully generalize my maxim, to universalize it, I apply it to everyone: Whenever anyone is in situation x, he/she shall do action y.

Kant's claim in the formula of universal law is that if an action is to be moral, it must be possible to will the universalization of the maxim which guides the act.

Kant is thus abstracting not only from the particular act (which might be done from self-interested motives) but also from the particular agent of the act.

There are thus two steps of abstraction which enable any imperative to command

The categorical imperative can be implemented as a test on an action we are considering whether to perform.

Maxims fail the categorical imperative test (in the formula of universal law) if they lead to contradictions.

As an example of how to use the formula of universal law, consider: Should I plagiarize my philosophy paper?

Step 1. Determine your maxim, the rule that guides your action

Step 2. Consider the situation if everyone did the same, if the maxim were to be universalized.

Step 3. Observe that it would be impossible to will this situation because it would eradicate the notion of paper writing. No one would assign papers since they would not be the kinds of exercises intended.

Result. Plagiarizing my philosophy paper would be contrary to duty.

Note that this result does not depend on our not liking a world of plagiarized papers.

It depends on its impossibility.

Plagiarized papers are no papers at all, in the relevant sense.

For another example, consider lying.

The essence of morality, according to the formula of universal law, is universalizability.

When we lie, we are actually willing that others lie to us.

But we do not want others to lie to us.

So we both want and do not want others to lie to us.

That makes it impossible to will a lie.

We can be inclined to lie, due to our base nature.

In such cases, we do not will a universal lie.

We only will that we may be an exception to a universal law of truth-telling.

IV. Contradictions

For it to be possible to universalize a maxim, it does not suffice to be willing to accept that everyone act as I do.

We have to imagine a world in which everyone acts as I do.

An action will be immoral if such a world is impossible.

Further, even if such a world were possible, an action will be immoral if we can not consistently will such a world.

Consequentialist considerations of whether we would like to live in such a world are irrelevant.

A maxim fails the first version of the categorical imperative if it leads to a contradiction.

There are two types of contradictions which can cause a maxim to fail.

C1. Contradiction in the world

C2. Contradiction in the will

For C1, a maxim can fail because it is not possible to have a world in which a maxim is universalized.

Consider jumping up and down while remaining motionless.

For C2, a maxim can fail because, though such a world is possible, it is not possible to will this world without contradiction.

If a maxim creates a contradiction in the world, it also creates a contradiction in the will, since we can not consistently will a contradiction.

To illustrate the use of the categorical imperative test, Kant provides four examples.

The four examples are categorized according to a division of perfect duties and imperfect duties, and duties to one's self and duties to others.

A perfect duty is one whose violation leads to a contradiction in the world.

An imperfect duty is one whose violation leads to a contradiction in the will.

Thus, there are four different illustrations, each corresponding to a kind of duty: perfect duties to one's self, perfect duties to others, imperfect duties to one's self, and imperfect duties to others.

The first two illustrations concern perfect duties, ones which lead to contradictions in the world.

The first illustration concerns suicide.

The generalized maxim is: Whenever I am in despair, I may kill myself.

The universalized maxim is: Whenever any one is in despair, he/she may kill him/herself.

But it is natural to want to maintain our lives.

It is then seen at once that a system of nature by whose law the very same feeling whose function...is to stimulate the furtherance of life should actually destroy life would contradict itself and consequently could not subsist as a system of nature (Kant 543a).

The second illustration concerns false promising in order to borrow money.

If everyone were to promise falsely to borrow money, then no one would believe such promises.

The institution of promise-making would disappear.

No promises would be possible.

The utilitarian considers factors irrelevant for Kant when evaluating the morality of false promising.

For example, the utilitarian worries about one's reputation, about whether I would like to live in a world in which promises are made falsely, and about whether one could actually get the money.

For Kant, all that is relevant is the impossibility of a world in which the maxim were universalized.

The third and fourth illustrations concern imperfect duties, ones which lead to contradictions in the will. In these illustrations, the maxims can be universalized.

It is possible to have a world in which every one follows the maxims.

But there is a contradiction in willing the maxims to be universal.

The third example concerns laziness, and letting our talents rust.

Consider Whitney Houston or [Dave Navarro](#).

Still, while we think that wasting good talent is lamentable, it may not be morally wrong.

Is it possible for a rational person to be lazy?

Kant defines rationality in such a way that laziness, in this sense, is incompatible with rationality.

Kant's claim captures our sense that the decision to spend one's life in a drugged-out haze rather than honing one's skills is a bad decision.

The fourth illustration concerns imperfect duties to others.

Consider willing to neglect others in need.

A world in which no one helps anyone else may be possible.

But if that maxim were universalized, we would also be willing to neglect ourselves.

We both want help from others, but do not want to help others.

Thus there is a contradiction in the will.

Another way to see that we can not universalize the maxim of neglecting others in need is to consider that the maxim 'never help anyone' must fail.

So, it is not the case that we may never help anyone.

That is, we must help someone sometime(s).

We can distinguish between duties of justice and duties of beneficence.

We must always comply with our duties of justice, never violating the categorical imperative.

But, we must select situations in which to be beneficent.

V. Not the Golden Rule

The formula of universal law sounds a bit like the golden rule.

It differs in several important respects.

The golden rule says that the actions we perform which affect others are only permissible if we are willing to have others affect us in the same way.

The golden rule is thus silent on actions which affect only one's self.

Kant, for example, believes that suicide is (at least sometimes) morally impermissible, as is neglecting one's natural talents.

The first and third illustrations involve duties to oneself.

More importantly, the golden rule allows me to treat others badly as long as I am willing to be treated badly myself.

If I am willing to be punched, I can punch some one.

There are masochistic people in the world.

But it seems wrong to think that they are permitted to harm others.

According to the categorical imperative, it is irrational to want to be mistreated.

In contrast to the golden rule, Kant argues that we may never mistreat others, no matter our inclination.

VI. The Formula of the End in Itself

The second version of the categorical imperative counsels us never to use persons as mere means.

Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end (Kant 545b).

Using someone as a mere means ordinarily involves deceit or coercion.

For example, we might involve someone in a plan to which they would not consent.

The formula of the end in itself is based on Kant's claim that all persons are due respect as rational beings.

All rational beings are equally able to make and break the moral law.

Notice that Kant is providing us with a criterion for personhood.

The ability to legislate morality for oneself is the mark of being a person, a member of the moral community.

The will is therefore not merely subject to the law, but is so subject that it must be considered as also *making the law* for itself and precisely on this account as first of all subject to the law (of which it can regard itself as the author (Kant 546b).

Kant expresses the distinction between persons and other things in terms of price and dignity.

Things with price have value only conditionally, or hypothetically.

Persons, in contrast, have dignity, which is unconditional, or categorical, value.

VII. The Kingdom of Ends

The kingdom of ends is the third version of the categorical imperative and a more positive way of viewing the moral law.

The first two versions provide a test for actions that rules decisively against infractions.

The formula of the kingdom of ends shows the ends of morality insofar as ends are compatible with the universality and categoricity of moral law.

We can see the kingdom of ends in Kant's distinction between price and dignity.

In the kingdom of ends everything has either a *price* or a *dignity*. If it has a price, something else can be put in its place as an *equivalent*: if it is exalted above all price and so admits of no equivalent, then it has a dignity (Kant 548a).

The formula of the kingdom of ends recognizes that we, as the makers of ends and sources of value, have goals and desires.

These ends should mesh with the ends of all other rational beings.

I understand by a "*kingdom*" a systematic union of different rational beings under common laws (Kant 547b).

Persons have value exactly because they can create value.

VIII. Hypothetical Imperatives and Autonomy

As we saw, a fundamental claim of Kant's moral theory is that the value of our actions must be independent of the consequences of our actions.

We call Kant's theory deontological, or duty-based, precisely because it says that morality is a system of categorical, abstract, perfectly general rules.

Utilitarianism, in contrast, understands morality as a system of hypothetical imperatives.

Moral commands cannot be hypothetical imperatives, according to Kant, because then the value of an action would depend in its consequences, which we can not predict or control.

- NC NC1. Consequences are out of our control.
- NC2. Morality is within our control.
- NCC. So, morality can not be based on consequences.

If morality were based on consequences, it would not be about duty, but self-interest.

Further, our ability to reason does not seem to lead us to happiness.

Indeed, we have no real control over our happiness.

We don't even really know what will make us happy.

Now it is impossible for the most intelligent, and at the same time most powerful, but nevertheless finite, being to form here a determinate concept of what he really wills. Is it riches that he wants? How much anxiety, envy, and pestering might he not bring in this way on his own head! Is it knowledge and insight? This might perhaps merely give him an eye so sharp that it would make evils at present hidden from him and yet unavoidable seem all the more frightful, or would add a load of still further needs to the desires which already give him trouble enough. Is it long life? Who will guarantee that it would not be a long misery? Is it at least health? How often has infirmity of body kept a man from excesses into which perfect health would have let him fall!... (Kant 541a-b).

If our ability to reason has any purpose, it must be independent of goals, it must be good in itself.

We should seek our happiness since it will help avoid temptation away from the moral law.

But our own happiness is not a moral matter.

Even if we are inclined to violate the moral law, we can act morally.

The moral quality of our action will be determined only by the content of our will.

Since the commands of morality do not come from the consequences, they must come from ourselves.

Kant calls the fact that we give the moral law to ourselves autonomy.

He implicitly assumes two axioms: that we are free to act and that morality is possible.

Then, he argues for autonomy.

- A A1. A moral action must be done, independently of your desires.
- A2. So there must be some reason to do it.
- A3. The reasons do not come from outside of us.
- AC. Therefore, we give the moral law to ourselves, i.e. we are autonomous.

A1 and A2 are supposed to be obvious.

A3 relies on the insight that external motivation is a consequentialist notion.

If we act for external reasons, then we would undermine the universal character of morality.

External conditions are always different, and not controllable by the individual.
 Consider the role of examples in our discussion of utilitarianism.
 Kant denies that these examples can tell us anything important about morality.
 They involve how we feel about something, our intuitions and inclinations.
 Systems of morality based on hypothetical imperatives are heteronomous.
 But discussions of morality should proceed exclusively by pure reason.
 The rational subject gives the moral law to himself and the system may be pure and autonomous,
 independent of capricious desires or circumstances.
 Kant's concept of autonomy is closely linked to our ordinary concept of freedom.
 For Kant, our moral freedom consists in the irrelevance of external factors to our morality.
 Freedom, then, is our ability to make and obey the categorical imperative.
 Note that one is most free when one is following the objective moral law, which constrains you from
 acting otherwise!

IX. Comparing Utilitarianism and Kantian Deontology

Here is a chart to help you compare utilitarianism and Kantian ethics.

	Utilitarianism	Kantian Ethics
	Action-Guiding, moral theory	Action-Guiding, moral theory
Why should one be moral?	Pain/ Pleasure Favor/ Disapproval	Duty, not inclination
What, generally, determines if an action is good or bad?	Consequences in the world	One's own Good Will
What tool do we use to evaluate actions?	Greatest Happiness Principle	Categorical Imperative Test
What is the minimum we have to do to be moral persons?	Create the greatest happiness for the greatest number	Never break the moral law, the CI
How can we exceed the moral minimum?	There is no supererogation. One must consider one's own interests impartially.	Sometimes aiding others in meeting their ends.
Why do persons have value?	They can be happy.	They are the bearers of rational life.
Scope	Broad	Narrow
Precision of the theory	Imprecise	Precise

X. Criticisms of Kant's Moral Theory

There are times when we think that lying and killing are morally acceptable.

Consider the example of Danish fishing boats ferrying Jews away from Nazi-controlled regions.

It is hard to see how we could formulate maxims governing such acts without violating the categorical imperative.

In such cases, are we deciding to break the moral law or do we want our morality to permit these acts?

It is clear that Kant dismisses the objection, and maintains his exceptionlessness.

Kant decries such examples as irrelevant.

He considers a case of an inquiring murderer asking you the whereabouts of a loved one.

Kant explicitly denies that lying to the murderer is morally acceptable.

Truthfulness in statements that cannot be avoided is the formal duty of man to everyone, however great the disadvantage that may arise therefrom for him or for any others.. [By telling a lie] I do wrong to duty in general in a most essential point. That is, as far as in me lies, I bring it about that statements (declarations) in general find no credence, and hence also that all rights based on contracts become void and lose their force, and this is a wrong done to mankind in general. (Kant, "On a Supposed Right to Lie Because of Philanthropic Concerns," Academy edition 426).

Kant allows that we can refuse to answer the murderer's question.

But to lie in order to achieve a particular outcome would make morality heteronomous.

Kant is holding tight to the claim that our intentions exclusively determine morality.

But even the best intentions can lead to bad consequences.

We can, as Mill says, often reasonably predict the consequences of our actions.

Falling back on good intentions seems morally irresponsible.

In the case of the inquiring murderer, my duty to tell the truth seems overwhelmed by my duties to family and friends.

I seem to have some control over some consequences, and it is naively idealistic to think otherwise.

For Kant, we are never responsible for the bad consequences of our truths, even though we are always responsible for the consequences of a lie.

If something bad happens, then someone else is responsible.

"Get your own moral house in order," is a tough position, especially when we have reasonable expectations of being able to influence others.

While Kant claims that appeals to intuition are illicit in determining the morality of an action, it seems that we must rely on our intuitions when choosing a moral theory.

A more abstract, and perhaps more worrisome, criticism of Kant's theory is that different descriptions of the same acts may result in different outcomes of the categorical imperative test.

Consider [an example, from Bernard Williams](#), of Jim who must choose between shooting one man or letting twenty men die.

Jim finds himself in the central square of a small South American town. Tied up against the wall are a row of twenty Indians, most terrified, a few defiant, in front of them several armed men in uniform. A heavy man in a sweat-stained khaki shirt turns out to be the captain in charge and, after a good deal of questioning of Jim which establishes that he got there by accident while on a botanical expedition, explains that the Indians are a random group of the inhabitants who, after recent acts of protest against the government, are just about to be killed to remind other possible

protestors of the advantages of not protesting. However, since Jim is an honoured visitor from another land, the captain is happy to offer him a guest's privilege of killing one of the Indians himself. If Jim accepts, then as a special mark of the occasion, the other Indians will be let off. Of course, if Jim refuses, then there is no special occasion, and Pedro here will do what he was about to do when Jim arrived, and kill them all. Jim, with some desperate recollection of schoolboy fiction, wonders whether if he got hold of a gun, he could hold the captain, Pedro and the rest of the soldiers to threat, but it is quite clear from the set-up that nothing of the sort is going to work: any attempt at that sort of thing will mean that all the Indians will be killed, and himself. The men against the wall, and the other villagers understand the situation, and are obviously begging him to accept. What should he do? (Bernard Williams, *A Critique of Utilitarianism*)

We can describe Williams' example in two different ways.

Description 1:

Choose between
a. shooting a man
and
b. not shooting a man.

Description 2:

Choose between
a'. saving 19 lives
and
b'. aiding a corrupt military.

If we describe the act in the first way, then the categorical imperative test says that we can not shoot the man.

If we describe it in the second way, it seems that the categorical imperative test can allow us to shoot the man.

Kant would respond that Description 2 is incorrect.

In formulating one's maxims, we must focus on our own moral life.

The correct description of an act should involve only our actions and not their effects on others.

Kant's response presupposes that there is one and only one correct description of the act.

A description should be objective, non-controversial, and morally neutral.

Here, Kant relies on Hume's fact/value distinction.

Hume says that there are facts, in the world, and values, which we impose on it.

Facts are objective.

We project our values on to the world of facts.

For Hume, these values are subjective (though he claims universality).

(Hume uses the fact/value distinction for another purpose, to establish subjectivism, but that does not matter here.)

For Kant, values are universal, derived from the rationality that we all share.

Describing the proposed act as 'saving 19 lives' involves desires, consequences, and other people.

But so does 'shooting a man'.

How about 'pulling a trigger'?

But that is no good for similar reasons.

There is nothing wrong with pulling a trigger on a paint gun, or a water gun.

How about 'moving my finger while...'

By insisting that our descriptions of our actions be strictly objective and morally neutral, we force them to abandon all sense of the action itself, and why it might be wrong.

The big question here is whether there is a fact/value distinction.

If not, then Kant will have real trouble describing acts in any morally neutral way.

Different descriptions of the same acts will result in different outcomes of the categorical imperative test.

But whether there is a fact/value distinction or not is not a settled matter.