

How should I live?



I should do whatever I want to do
(and that is all anyone ever does)

Today we continue our discussion of ethics. The central question of ethics, or moral philosophy, is: What ought I to do? How should I live?

Almost all standard answers to this question have one thing in common: they say that, at least sometimes, one ought to do things which are not in one's self-interest. Let's call actions of this sort **altruistic actions**.

Today our topic is a pair of related views which say that all of these standard views about ethics are mistaken.

The **psychological egoist** says that all of our actions are always self-interested — so that what conventional morality asks us to do is in fact impossible.

The **ethical egoist** says that, even if people do sometimes perform actions which are not in their self-interest, these actions are always mistakes. On this view, one ought always to do what is in one's own self-interest.

Let's discuss psychological egoism first. To understand the view we need to ask: what does it mean for an action to be **self-interested**?

One appealingly simple answer to this question is called **hedonistic psychological egoism**. This is the view stated by the 18th century philosopher Jeremy Bentham in the following passage:

NATURE has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne. They govern us in all we do, in all we say, in all we think: every effort we can make to throw off our subjection, will serve but to demonstrate and confirm it. In words a man may pretend to abjure their empire: but in reality he will remain subject to it all the while.

On this view, people only ever act in ways which they take to maximize their pleasure, and minimize their pain.

In the reading for today, Glaucon gives what can be thought of as an argument for psychological egoism.

This is based on the example of the ring of Gyges: a ring which renders its wearer invisible, and hence makes his actions free from any consequences which might result from the opinions of others.

If now there should be two such rings, and the just man should put on one and the unjust the other, no one could be found, it would seem, of such adamant temper as to persevere in justice and endure to refrain his hands from the possessions of others and not touch them, though he might with impunity take what he wished even from the market place, and enter into houses and lie with whom he pleased, and slay and loose from bonds whomsoever he would, and in all other things conduct himself among mankind as the equal of a god. And in so acting he would do no differently from the other man, but both would pursue the same course. And yet this is a great proof, one might argue, that no one is just of his own will but only from constraint, in the belief that justice is not his personal good, inasmuch as every man, when he supposes himself to have the power to do wrong, does wrong. For that there is far more profit for him personally in injustice than in justice is what every man believes, and believes truly, as the proponent of this theory will maintain. For if anyone who had got such a license within his grasp should refuse to do any wrong or lay his hands on others' possessions, he would be regarded as most pitiable and a great fool by all who took note of it, though they would praise him before one another's faces, deceiving one another because of their fear of suffering injustice. So much for this point.

The central problem for psychological egoism seems to be that there are obvious, mundane counterexamples to it. Let's start with an un-dramatic example. Have any of you ever done anything for your roommate which might at least appear to be an altruistic action?

However, a certain kind of game illustrates the fact that decisions which seem to be altruistic may be, in the end, self-interested.

You and your partner (the person sitting next to you) have been in business running drugs for the last few months. You've just been arrested by the police, who are interrogating you in separate rooms.

Here's what you know about your situation: you know that if **both** you and your partner confess, given the evidence that the police will then possess, you'll each get about 5 years for your crimes.

On the other hand, if you confess and your partner doesn't, you'll as a reward get off scot free, and your partner will be stuck serving 10 years in jail. (And the reverse if your partner confesses, and you do not.)

If you both keep quiet - and **neither** of you confesses - then the police will have only have the evidence to convict the two of you for a lesser crime - for which, you estimate, you'll have to serve 2 years in jail.

You and your partner (the person sitting next to you) have been in business running drugs for the last few months. You've just been arrested by the police, who are interrogating you in separate rooms.

Here's what you know about your situation: you know that if **both** you and your partner confess, given the evidence that the police will then possess, you'll each get about 5 years for your crimes.

On the other hand, if you confess and your partner doesn't, you'll as a reward get off scot free, and your partner will be stuck serving 10 years in jail. (And the reverse if your partner confesses, and you do not.)

If you both keep quiet - and **neither** of you confesses - then the police will have only have the evidence to convict the two of you for a lesser crime - for which, you estimate, you'll have to serve 2 years in jail.

Take a second and think about your decision. Then write down on a piece of paper either "confess" or "stay silent."

Now show the paper to your partner, and take note of how many years of each of you will have to serve in jail.

You and your partner (the person sitting next to you) have been in business running drugs for the last few months. You've just been arrested by the police, who are interrogating you in separate rooms.

Here's what you know about your situation: you know that if **both** you and your partner confess, given the evidence that the police will then possess, you'll each get about 5 years for your crimes.

On the other hand, if you confess and your partner doesn't, you'll as a reward get off scot free, and your partner will be stuck serving 10 years in jail. (And the reverse if your partner confesses, and you do not.)

If you both keep quiet - and **neither** of you confesses - then the police will have only have the evidence to convict the two of you for a lesser crime - for which, you estimate, you'll have to serve 2 years in jail.

Some years later, you and your partner are both free, and back to your old tricks. Unsurprisingly, you again get arrested, and the police again offer you the same deal.

Again take a second and think about your decision. Then again write down on a piece of paper either "confess" or "stay silent."

Now show the paper to your partner, and take note of how many years of each of you will have to serve in jail.

You and your partner (the person sitting next to you) have been in business running drugs for the last few months. You've just been arrested by the police, who are interrogating you in separate rooms.

Here's what you know about your situation: you know that if **both** you and your partner confess, given the evidence that the police will then possess, you'll each get about 5 years for your crimes.

On the other hand, if you confess and your partner doesn't, you'll as a reward get off scot free, and your partner will be stuck serving 10 years in jail. (And the reverse if your partner confesses, and you do not.)

If you both keep quiet - and **neither** of you confesses - then the police will have only have the evidence to convict the two of you for a lesser crime - for which, you estimate, you'll have to serve 2 years in jail.

Some years later, you and your partner are, once again, both free, and, once again, back to your old tricks. Unsurprisingly, you again get arrested, now for the third time, and the police again offer you the same deal.

Again take a second and think about your decision. Then again write down on a piece of paper either "confess" or "stay silent."

Now show the paper to your partner, and take note of how many years of each of you will have to serve in jail.

You and your partner (the person sitting next to you) have been in business running drugs for the last few months. You've just been arrested by the police, who are interrogating you in separate rooms.

Here's what you know about your situation: you know that if **both** you and your partner confess, given the evidence that the police will then possess, you'll each get about 5 years for your crimes.

On the other hand, if you confess and your partner doesn't, you'll as a reward get off scot free, and your partner will be stuck serving 10 years in jail. (And the reverse if your partner confesses, and you do not.)

If you both keep quiet - and **neither** of you confesses - then the police will have only have the evidence to convict the two of you for a lesser crime - for which, you estimate, you'll have to serve 2 years in jail.

Finally, you are both out of jail, and by this point are quite old. However, you decide to get together for one last big deal. But you're caught, and the police again offer you the same deal. You're getting very tired of this, but, given your age, at least **you know that this is the last time that you and your partner will ever be arrested together.**

Again write down on a piece of paper either "confess" or "stay silent" and, when you're both done, show the paper to your partner, and calculate the results.

Now add up the total number of years that each of you served in jail.

The situation in which you and your partner were placed is a [prisoner's dilemma](#).

Simple prisoner's dilemmas are games in which two agents face a decision between two courses of action, A and B, with the following properties: for each player, no matter what the other player does, B will provide a better outcome than A; but a situation in which both players do A is mutually preferable to a situation in which both do B.

The version of the prisoner's dilemma just described can be modeled by the following chart:

Courses of action	Possibility 1: Your partner confesses	Possibility 2: Your partner stays silent
Confess	5 years in jail	go free
Stay silent	10 years in jail	2 years in jail

Once you represent the choice in this way, one important fact about cases of this sort becomes clear: confessing [dominates](#) silence. No matter what your opponent does, you are better off confessing.

Courses of action	Possibility 1: Your partner confesses	Possibility 2: Your partner stays silent
Confess	5 years in jail	go free
Stay silent	10 years in jail	2 years in jail

Once you represent the choice in this way, one important fact about cases of this sort becomes clear: confessing **dominates** silence. No matter what your opponent does, you are better off confessing.

This strongly suggests that it is in your self-interest to confess — and hence that staying silent is altruism.

This may be correct for a one-off prisoner's dilemma. But we in effect just played two different types of prisoner's dilemma games: three in which you did not know how many other games would follow, and one in which you knew that no others would follow (the last one).

Is confessing always the best strategy in the first sort of case? Some data from computer simulations of prisoner's dilemmas suggests that it is not, and that the best strategies — in the sense of the strategies which are most likely to improve the status of a given agent — are not ones in which the agent always confesses.

Is confessing always the best strategy in the first sort of case? Some data from computer simulations of prisoner's dilemmas suggests that it is not, and that the best strategies — in the sense of the strategies which are most likely to improve the status of a given agent — are not ones in which the agent always confesses.

This suggests that staying silent in a multi-turn prisoner's dilemma is — at least sometimes — an action which seems to be altruistic, but is really not, since it may be (and be known to be) in the agent's best interests.

Might this model fit any of our seeming examples of altruistic actions — like the example of roommate interactions?

Can this be used to the advantage of the psychological egoist?

But now consider the last turn of our prisoner's dilemma. What was the best self-interested strategy there?

But this can be used to pose a challenge to the psychological egoism, since some of our seeming examples of self-interested actions seem more like the last turn in a prisoner's dilemma. A simple example: giving a stranger your cab in a big city.

What should the psychological egoist say about these cases?

The psychological egoist will say that this is again only an apparent case of altruism. But this time the agent is not hoping that the other person involved will reciprocate in the future; rather, they are after other things that they value, like feelings of self-satisfaction or avoiding unpleasant feelings of guilt.

No doubt this sort of account is true for some actions. But can every apparent case of altruistic action be explained in this way?

Let's work through a few such examples. Imagine a soldier who throws himself on a grenade to save his fellow troops. This certainly seems to be an unselfish act, and to be done in the knowledge that it will be much worse for that soldier's future pleasure and pain than the alternative.

The psychological egoist might say that the soldier does this action only to avoid unpleasant guilt feelings in the future. But this assumes that the soldier regards guilt feelings as so painful that death is preferable to them — which seems implausible.

Or the egoist might say that the soldier does this in order to go to heaven — and that the pleasure of this outweighs the pain of leaping on the grenade. But the soldier might be an atheist — or might be a theist who does not have the (somewhat implausible) view that jumping on the grenade is required to go to heaven.

The egoist's hypotheses might describe some acts of apparent heroism — but it seems implausible to think that they describe all of them.

Or consider the un-heroic act of holding open a door for someone in the rain.

The psychological egoist might say that one does this in the hope that that person will hold the door open for oneself in the future. But in many cases — e.g. being in a strange city — this hope will be wildly irrational.

The psychological egoist might say that is done to bring about pleasant feelings of self-satisfaction, and to avoid unpleasant guilt feelings. But many people are courteous without being especially guilt-ridden or self-satisfied.

Moreover, we can imagine a kind of Gyges-esque thought experiment. Suppose that you knew that you would be given an amnesia-inducing drug right after holding the door. Wouldn't you still, in at least some cases, hold the door open? This would appear to be problematic for the psychological egoist.

So it looks like, if we are to believe psychological egoism, we ought to have some good argument for it.

Here is one argument — **the argument from the necessity of desire**. It seems plausible that every intentional action involves some desire — one can't intentionally hold a door for someone without desiring to hold the door. But then it just follows that **we are always moved to action by our own desires**. But doesn't that make all of our actions self-interested?

We might represent this argument for psychological egoism as follows:

1. Everyone's actions are caused by their own desires.

C. Everyone's actions are done in their own self-interest.

The premise is plausible. But the argument seems to be invalid, because the fact that a desire is a desire **of mine** does not imply that it is a desire **for me**. There seems to be no impossibility in having a desire which is not a desire for my own pleasure, but is instead a desire for something quite different — like someone else's well-being.

A second argument for psychological egoism is based on the idea that humans evolved by natural selection. Oversimplifying, the theory of evolution by natural selection leads us to expect, in general, that evolution will favor those traits which improve the chances of the the bearer of the trait having a relatively large number of viable offspring. So, if this theory is correct, we should expect that altruistic tendencies will not be passed on to future generations — unless that tendency leads, in some other way, to the propagation of your own genes.

This argument might be represented as follows:

1. Human beings evolved by natural selection.
2. Altruistic actions are never favored by natural selection.
3. If an organism evolved by natural selection, each of its actions must be favored by natural selection.

C. Human beings never perform altruistic actions. (1,2,3)

1. Human beings evolved by natural selection.
 2. Altruistic actions are never favored by natural selection.
 3. If an organism evolved by natural selection, each of its actions must be favored by natural selection.
-

C. Human beings never perform altruistic actions. (1,2,3)

The second premise of this argument is likely false, for reasons (e.g., kin selection) which may or may not be relevant to our topic today. But the third premise is false. This is because the fact that a given creature evolved via natural selection does not imply that every trait of that creature is one that it has because that trait was selected for by natural selection. So showing that it is impossible for evolution to select for a trait does not show that no evolved creature could have that trait.

Psychological egoism thus seems to be open to counterexamples, and to lack (so far) a convincing argument in its favor. Let's turn instead to ethical egoism: the view that, even if people sometimes do not act in their self-interest, they **ought** to always act in their own self-interest.

The ethical egoist challenges the idea that we ought sometimes to act in the interests of others rather than in our own interests. How might one argue against the ethical egoist?

This is the question taken up by Socrates and Glaucon in the reading for today. Or, more precisely, they are considering a specific answer to this question, which might be put like this:

One should be altruistic because being altruistic has **instrumental value**: in the end, being moral is the best way to serve your interests.

Glaucon says that this is the view of “the multitude,” and describes it as the view that acting morally “belongs to the toilsome class of things that must be practiced for the sake of rewards and repute due to opinion but that in itself is to be shunned as an affliction.” (358a)

But this seems like a pretty unsatisfactory defense of altruism, for two reasons.

First, this seems (as Glaucon points out) to be less an argument that one ought to be altruistic than an argument that one ought to **seem to be altruistic**. Sometimes, the best way to do this will be to act in an altruistic manner — but certainly not always.

Second, it licenses only a very limited sort of ‘altruism’ — namely, the kind that we found in multi-turn prisoner’s dilemma games.

Here is a more serious objection to ethical egoism. It seems that the position is, in a certain way, self-refuting. For suppose that I argue in favor of ethical egoism. This would seem not very well-designed to bring about what is good for me in the future, since everyone serving their own interests is not likely to bring about what is best for me. So anyone who defends ethical egoism is either confused, or does not really believe that ethical egoism is true.

To this, the ethical egoist might respond that the argument does not show that ethical egoism is self-refuting. It shows instead that [defending](#) ethical egoism is self-refuting. There is, for all we have said, nothing incoherent in quietly believing ethical egoism to be true, and not telling anyone else about this fact.

So it seems that we have so far failed to give any argument which shows that ethical egoism is false (as opposed to not coherently defensible).

Here is one attempt to do better. It seems that one sort of irrationality results from treating like cases differently: for example, if one responds to an identical bet differently on different days, and can provide no reason for this different response, this seems to be a kind of practical irrationality. **Arbitrariness** of this sort is irrational.

But isn't the ethical egoist arbitrary in just this way? The ethical egoist should agree that he is a being of the same sort as other people, and hence that his interests are things of the same sort as the interests of others. But in deciding what to do, he takes his own interests into account, but not the interests of others. Why isn't this just the sort of arbitrariness that seems to be a mark of irrationality?

If this is right, then this is a problem for ethical egoism — and hence a problem for the view that one ought always to do what one wants to do.

Here is a last problem for ethical egoism, which is in some ways similar to Moore's reply to the skeptic.

Among the beliefs which I'm most sure of are moral judgements about certain particular cases. So, for example, I am sure that one ought not to torture an innocent person because it is very slightly in your self-interest to do so.

But ethical egoism would seem to contradict views such as this. So, since these views are true, ethical egoism is false.

The fact that ethical egoism seems to conflict with our beliefs about what one ought to do in many cases seems to show that we should only believe ethical egoism if we have a strong argument in its favor. But it is hard to see what this could be.