What must I do?

Whatever I want to do

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

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 rational 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 — and that it is impossible for them to do otherwise.

In the reading for today, though, 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 adamantine 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 is that there seem to be obvious counterexamples to it — cases in which people behave in ways quite unlike the way that Glaucon imagines that people will behave when equipped with the ring of Gyges. Some are dramatic, like a fireman running into a burning building to save someone. Others are very un-dramatic, like holding the door for someone to enter a building on a rainy day.

The defender of psychological egoism can come up with hypotheses about cases like these. She might say that the fireman and the door-holder, for example, are motivated by the fact that they find pleasure in thinking of themselves as, respectively, heroic and courteous.

But these hypotheses in many cases seem quite implausible. Why believe that in every case the motivation is really about one's own future pleasure or pain?

Let’s work through a few such examples. Imagine a soldier who throws himself on a grenade to save his fellow troops. This certain 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-glorious 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 holding 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 controversial. 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 rational egoism: the view that, even if people sometimes do not act in their self-interest, they ought to act in their own self-interest.

The rational 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 rational 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)

Here is one way to develop the view that Glaucon attributes to "the multitude." Consider the following sort of situation:

The prisoner's dilemma

You are one of two prisoners arrested for a crime. You, and the other prisoner, are each rational, and you each know that if you both stay silent, and don't confess, you will each be convicted of a fairly minor crime, and get 2 years in jail each. If you turn State's evidence and the other prisoner stays silent, then you will get off with nothing, and the other prisoner will get 10 years; exactly the opposite will happen if the other prisoner turns State's evidence, and you stay silent. If you both confess, you both get 5 years. Is it rational for you to confess, or stay silent?

Suppose that we think that it is rational to not confess — to stay silent. How might this be used to support the view that it is rational to be moral?

One might, however, question this defense of staying silent. For one thing, one might doubt whether it rational to stay silent; the choice, after all, can be represented in 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

It looks like whatever your partner does, you are better off confessing. So how could it be rational to stay silent?

One might reply that confessing is in your collective self-interest; but this raises a close relative of the question we were trying to answer in the first place: why would it ever be rational to act against my self-interest, in favor of the interests of some group of which I am a part?

We might reply by pointing out that, even if confessing is rational in a one-turn prisoner's dilemma, in a multi-turn prisoner's dilemma - in which your partner's decision might be based on your past decisions - staying silent might well be rational. Perhaps decisions about whether to be moral are, in important respects, like multi-turn prisoner's dilemma games.

But, first, many moral decisions are not equivalent to multi-turn prisoner's dilemma games of the relevant sort. Remember the seeming counterexamples to psychological egoism discussed above: the examples of self-sacrifice, and of holding the door for someone in a strange city. These are equivalent to the last turn of a multi-player prisoner's dilemma game - and the self-interested rational choice in that case seems to be, pretty clearly, that one should confess.

Second, 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.

Here is a more serious objection to rational egoism. It seems that the position is, in a certain way, self-refuting. For suppose that I argue in favor of rational 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 rational egoism is either confused, or does not really believe that rational egoism is true.

To this, the rational egoist might respond that the argument does not show that rational egoism is self-refuting. It shows instead that defending rational egoism is self-refuting. There is, for all we have said, nothing incoherent in quietly believing rational 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 rational egoism is false (as opposed to not coherently defendable).

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 rational egoist arbitrary in just this way? The rational 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 rational egoism — and hence a problem for the view that one ought always to do what one wants to do.