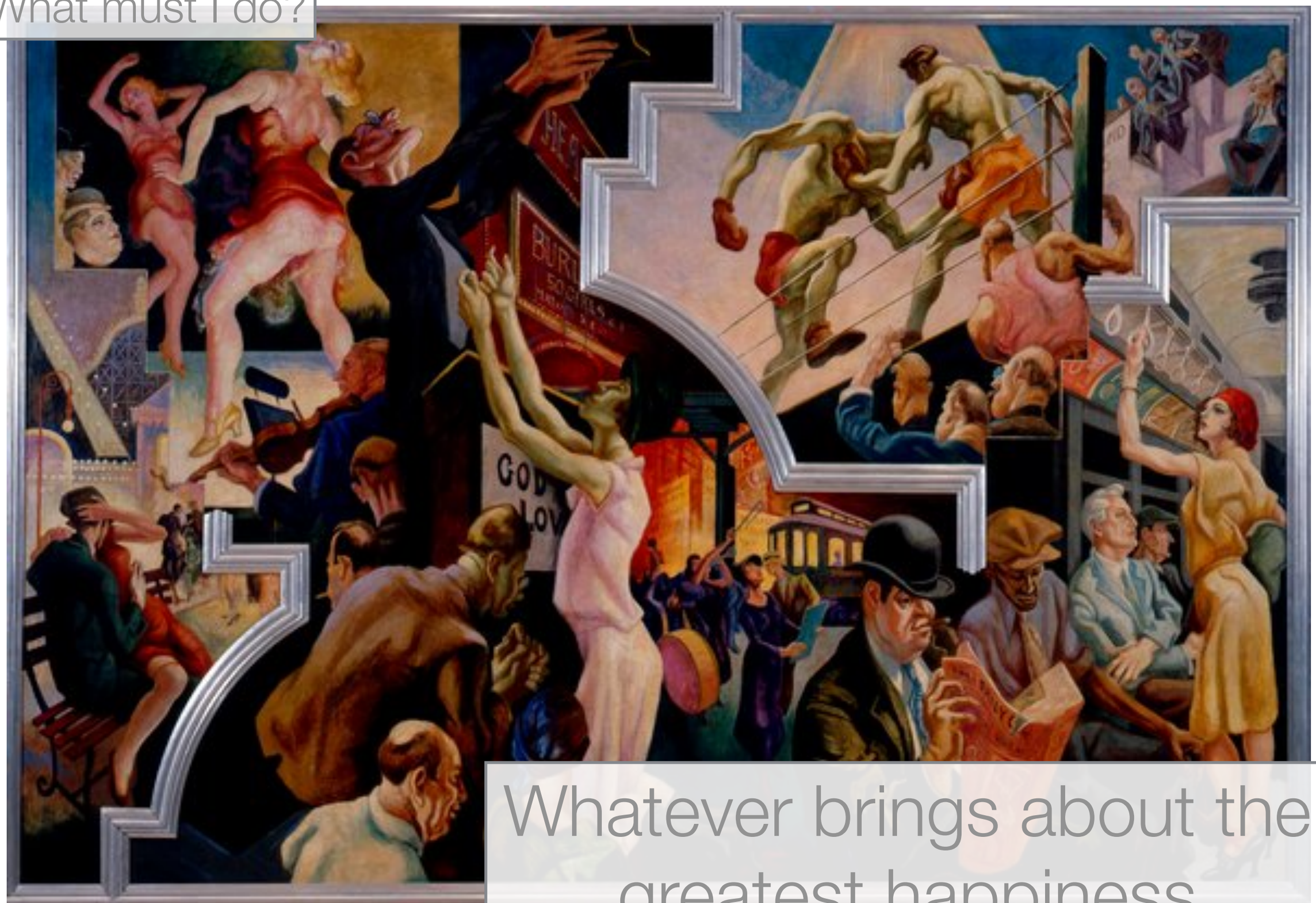
What must I do?



Whatever brings about the greatest happiness

Suppose that some actions are right, and some are wrong. What's the difference between them? What makes some actions right, and others wrong?

Here is one very simple, but also very plausible, answer to this question:

Consequentialism
An action is the right thing to do in certain circumstances if, of all the actions available in those circumstances, it would produce the best outcome.

Consequentialism says, simply, that we should judge actions by their consequences. Whatever will lead to the best overall outcome is what one ought to do.

A slightly different way to get the general idea is this: if I am deciding between doing action A and action B, I should try to figure out what the world would be like if I did A, and what the world would be like if I did B; and I should do whichever action would lead to the better world.

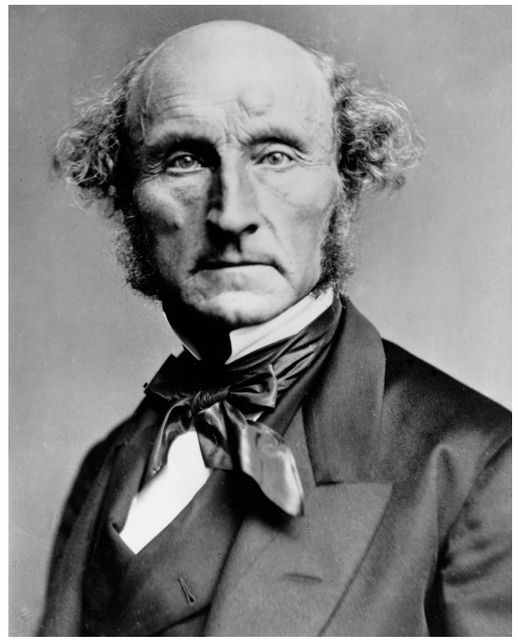
Consequentialism

An action is the right thing to do in certain circumstances if, of all the actions available in those circumstances, it would produce the best outcome.

This view raises a question: what makes one outcome, or state of affairs, better than another?

To answer this question is to give a theory of value: a theory about what makes one state of the world better, or worse than, another.

In the reading for today, John Stuart Mill gives the following statement of his theory of value.



> The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure. To give a clear view of the moral standard set up by the theory, much more requires to be said; in particular, what things it includes in the ideas of pain and pleasure; and to what extent this is left an open question. But these supplementary explanations do not affect the theory of life on which this theory of morality is grounded – namely, that pleasure, and freedom from pain, are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain.

This view is sometimes called hedonism:

Hedonism
One state of affairs is better than another if and only if it involves the best overall distribution of pleasure and pain.

Consequentialism
An action is the right thing to do in certain circumstances if, of all the actions available in those circumstances, it would produce the best outcome.

+

Hedonism
One state of affairs is better than another if and only if it involves the best overall distribution of pleasure and pain.

=

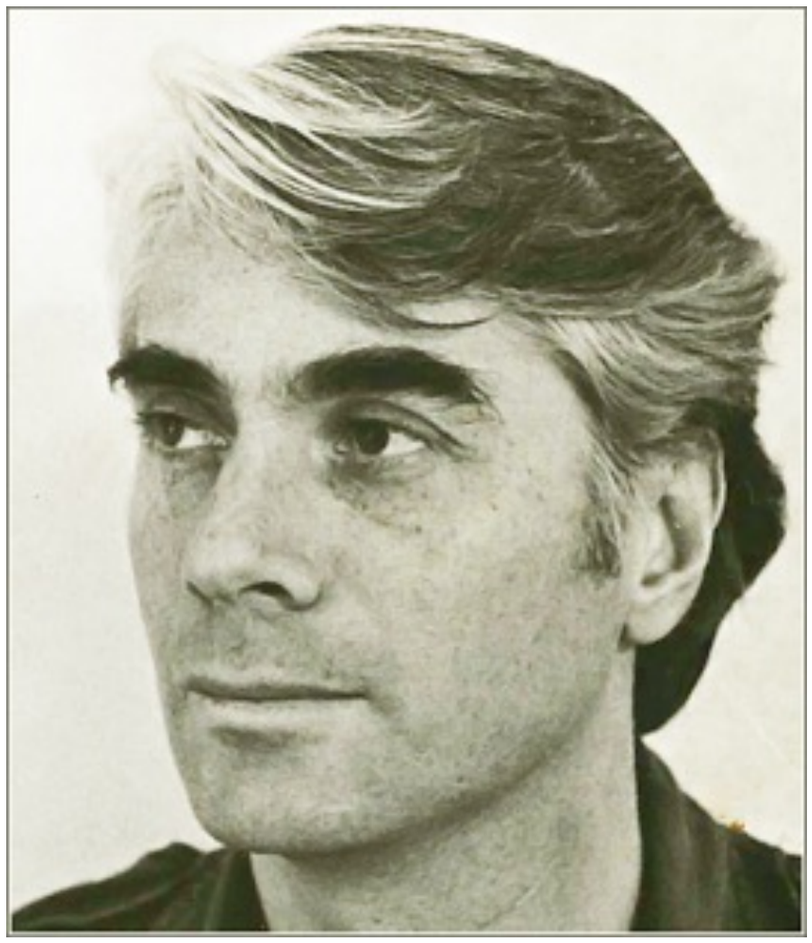
Utilitarianism
An action is the right thing to do in certain circumstances if, of all the actions available in those circumstances, it would produce the best overall distribution of pleasure and pain.

Utilitarianism
An action is the right thing to do in certain circumstances if, of all the actions available in those circumstances, it would produce the best overall distribution of pleasure and pain.

Utilitarianism is a very clear and plausible-sounding view about ethics. This is the view which is often summed up with the slogan that one ought always to act to cause the greatest happiness for the greatest number. It is a paradigmatically unselfish theory: no one's pleasures and pains are more important than anyone else's.

Today we are going to discuss two sorts of cases which can be used to pose problems for utilitarianism. The first is a challenge to utilitarianism in particular; the second challenge consequentialism in general.

The first challenge can be brought out by Robert Nozick's example of the experience machine.



Suppose there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences, we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences, selecting your life's experiences for, say, the next two years. After two years have passed, you will have ten minutes or ten hours out of the tank, to select the experiences of your *next* two years. Of course, while in the tank you won't know that you're there; you'll think it's all actually happening. Others can also plug in to have the experiences they want, so there's no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everyone plugs in.) Would you plug in? *What else can matter to us, other than how our lives feel from the inside?* Nor should you refrain because of the few moments of distress between the moment you've decided and the moment you're plugged. What's a few moments of distress compared to a lifetime of bliss (if that's what you choose), and why feel any distress at all if your decision *is* the best one?

What must the utilitarian say about the relative goodness of the state of affairs in which everyone (or almost everyone) plugs in and the state of affairs in which no one does?

Suppose you face the decision whether to get into the experience machine. What would a utilitarian say about what you ought to do?

Suppose now that you face the decision of whether you should put everyone into the experience machine. (The machines are maintained by extremely reliable robots.) What would a utilitarian say about what you ought to do?

Does it matter if people ask you (or beg you) not to plug them in?

Is Nozick right that these consequences of utilitarianism, and hedonistic consequentialism, are incorrect?

It is important to see that, even if you agree with Nozick, his example does not show that Consequentialism is false, but only that a particular version of that view is false. One might agree with Nozick about the experience machine, and still be a Consequentialist, if one holds that what makes one state of affairs better than another can sometimes depend on facts other than sensations of pleasure and pain.

What might make one state of affairs better than another, if not the overall distribution of pleasure and pain? This is a difficult question, to which many different answers have been given. Some answers are:

The extent to which the desires of agents are satisfied.

The extent to which the states of affairs contain beauty, or love, or friendship, or something else taken to be of objective value.

The extent to which the states of affairs maximize the well-being, or welfare, of agents.

Corresponding to each of these views about what makes one outcome better than another is a different version of consequentialism. For example, “preference-satisfaction consequentialism” is the view that one should always act in such a way that maximizes the extent to which the desires of people are satisfied.

Let's turn from our evaluation of Utilitarianism in particular to a challenge for Consequentialism in general.

One general feature of consequentialism is its indifference to how consequences are brought about. What matters when deciding what to do is what one's various actions will bring about, not what those actions are.

One consequence of this general feature might be stated like this:

> **Act/omission indifference**
> Whether I bring about some state of affairs by doing something or failing to do it is morally irrelevant.

Some aspects of this principle are quite appealing. For example, the principle refuses to let people stand idly by as others suffer, on the grounds that one is not the cause of that suffering. One whose failure to act leads to suffering is, according to consequentialism, just as responsible for it as one whose action leads to that suffering.

But some troubling consequences of this principle are brought out by the following example, due to Judith Jarvis Thomson.

David is a great transplant surgeon. Five of his patients need new parts—one needs a heart, the others need, respectively, liver, stomach, spleen, and spinal cord—but all are of the same, relatively rare, blood-type. By chance, David learns of a healthy specimen with that very blood-type. David can take the healthy specimen's parts, killing him, and install them in his patients, saving them. Or he can refrain from taking the healthy specimen's parts, letting his patients die.

What does the consequentialist say that David ought to do in this case? What ought he to do?

This sort of case might lead you to think something like this: killing someone in order to save the lives of others is never morally permissible.

If this were true, this looks like it would be trouble for the Consequentialist, since it is hard to argue that killing someone, especially when it saves the lives of others, can never lead to an outcome which is, overall, the best of the available options.

But, as some of Thomson's other examples show, matters are not quite this simple.

Consider one of her examples involving a trolley car:

> Edward is the driver of a trolley, whose brakes have just failed. On the track ahead of him are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Edward can turn the trolley onto it. Unfortunately there is one person on the right-hand track. Edward can turn the trolley, killing the one; or he can refrain from turning the trolley, killing the five.

Is it permissible for Edward to turn the trolley? If so, wouldn't this be a case in which it is permissible — perhaps even obligatory — to kill one person in order to save five lives?

But then why might it be OK for Edward to turn the trolley, but clearly not permissible for the doctor to cut up his healthy specimen?

One might try to explain the difference here like this: Edward is choosing between killing one and killing five; either way, he is killing someone. David is choosing between killing one and letting five die, and this is something quite different. We have a stronger duty to avoid killing than to prevent people from dying.

But it is not clear that this is the right explanation of the difference between Edward and David, as is brought out by the example of Frank.

> Frank is a passenger on a trolley whose driver has just shouted that the trolley's brakes have failed, and who then died of the shock. On the track ahead are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Frank can turn the trolley onto it. Unfortunately there is one person on the right-hand track. Frank can turn the trolley, killing the one; or he can refrain from turning the trolley, letting the five die.

Here it seems as though Frank is faced with a choice between letting five die, and killing one — so his choice seems, in this respect, just like David's (the surgeon's). But it seems as though it is morally permissible for Frank to turn the trolley, even though it is not morally permissible for David to cut up the healthy specimen.

More complications are introduced by yet a third trolley example:

George is on a footbridge over the trolley tracks. He knows trolleys, and can see that the one approaching the bridge is out of control. On the track back of the bridge there are five people; the banks are so steep that they will not be able to get off the track in time. George knows that the only way to stop an out-of-control trolley is to drop a very heavy weight into its path. But the only available, sufficiently heavy weight is a fat man, also watching the trolley from the footbridge. George can shove the fat man onto the track in the path of the trolley, killing the fat man; or he can refrain from doing this, letting the five die.

Many people think that it is not permissible for George to push the fat man. But why is this any different from turning the trolley to kill the one on the right hand section of the trolley tracks? After all, in both cases, you are killing one rather than letting 5 die.

One thought is this: the fat man has a right not to be pushed onto the tracks in a way that people standing on trolley tracks don't have a right not to be run over by trolleys.

This sort of thought also promises to make sense of the example of David the surgeon; perhaps healthy specimens have a right not to be cut up, but that dying patients in need of transplants have no right to be saved.

This way of thinking about these cases is very different than the way of approaching them suggested by Consequentialism. According to this view, we should think about what we ought to do by first thinking about the rights and obligations of the people involved and not, at least in the first instance, about which action would bring about the best outcome.

Beginning next time, we will begin discussing this other, non-consequentialist approach to ethical questions.