Philosophy 1100 – Introduction to Ethics

Lecture 3 – Three Different Types of Ethical Theories

The ethical theories that philosophers have advanced fall into three main groups. To understand these different types, one needs to think about the different types of moral or ethical judgments, broadly conceived, that people make.

1. Everyday Moral/Ethical Judgments

Question: What are some of the **different types** of ethical judgments, broadly conceived, that people make?

The expression "different types" is not the most precise expression, since we can divide things up into different types either by means of very broad categories, or by very fine grained ones. What I have in mind here is a division based upon categories that are as broad as possible.

One thing that people certainly do is to make judgments about the actions of people, and that's certainly a broad category. So we characterize actions as right and wrong, we say that someone should or shouldn't have done something, or that something was a good thing to do, or a bad thing to do.

Do we make moral judgments about things other than actions?

Well, we also make moral judgments about people. We say that people like Hitler and Stalin were very bad people – or that they were evil people – and that other people are very good people.

So we make moral judgments about actions, and about people. Are there any other broad categories of things that we make ethical judgments about?

Another category – a slightly less obvious one, I think – is that of traits of character. Consider traits of character like being an honest person, or being a loving person, or a person who keeps his or her promises. We think of these as good traits to have, while we think of traits such as that of being a cruel person, or a dishonest person, or an unfaithful person as bad traits to have. Or we speak about virtues and vices.

We have, then, at least these three broad categories of moral judgments, based on the things that we make moral judgments about:

(1) Moral judgments about actions being right or wrong;

(2) Moral judgments about people being good or bad;

(3) Moral judgments about traits of character being good or bad, being virtues or vices.

Is there any other type of ethical judgment, broadly conceived? The answer is that there is, and it is an extremely important one. But before going on to consider what that might be, let us focus on the three categories that we have noticed so far, and think about the following questions:

(a) Are these three types of judgments all independent of one another, or are they somehow logically related?

(b) If they are logically related, how are they related? Is one of these three categories more fundamental than the other two, so ethical judgments of those other two types can be analyzed in terms of ethical judgments of the more fundamental type?

Here, for example, is a possible view, and one that some philosophers appear to accept:

1. Of the three types of ethical judgments we have considered so far, the fundamental ones are those that are about traits of character, about virtues and vices.

2. Talk about the good or badness of people can then be analyzed in terms of judgments about the goodness or badness of traits of character. At least as a starting point, then, one might say:

A person is a good person to the extent that they possess good traits of character rather than bad ones, and a bad person to the extent that they possess bad traits of character rather than good ones.

3. Similarly, talk about the rightness and wrongness of actions can be analyzed in terms of judgments about the goodness or badness of traits of character. In particular, the following sort of analysis might be suggested:

An action of a certain type T is a **morally right action** if and only if there is some **virtuous trait of character** C such that people with that virtuous character trait C are **disposed to perform** actions of type T.

An action of a certain type T is a **morally wrong action** if and only if there is some **bad trait of character** C such that people with that bad character trait C are **disposed to perform** actions of type T.

Class Discussion: What is one to say about this theory? Does it seem right that of the three types of ethical judgments we have considered so far, it is judgments about the goodness or badness of traits of character that are logically the most basic? What **alternative** view (or views) might be proposed here? Which seems to you the **most plausible**?

2. Another, Very Important Type of Evaluative Judgment

There is, however, another type of ethical judgment, broadly conceived – one might speak of an **evaluative** judgment – that one tends not to think of when one talks about morality or ethics, but that is very important. To see what it is, ask yourself whether there are things other than actions, people, and traits of character that one refers to as good or bad.

Consider the following:

"How was the party? Was it a good one?"

"Today was a bad day for Sue. Everything went wrong that could have gone wrong."

"Bruce was a nice person, but he had a rather bad life, with lots of unhappiness and suffering, and some tragic events."

"That plane crash was really bad."

In these sentences, the terms "good" and "bad" are being applied not to actions, nor to people, nor to traits of character. They are being applied to things like occurrences, or events, or parts of a person's life, or a person's life as a whole.

Moreover, it would seem that such judgments can be applied to situations that do not involve any people at all. Compare, for example, the following two possible **worlds**:

World 1: A world that contains no human beings or other intelligent beings, but that does contain plants and animals, all of which are herbivores.

World 2: A world that is just the same as World 1 with respect to the types of plants and herbivores that it contains, but that rather that containing only herbivores, contains a large number of carnivores as well, as well as many more natural disasters, such as forest fires.

So World 2 will contain much more pain than World 1, with various animals being hunted down and killed by other animals, and animals dying painfully in things like forest fires. Doesn't one want to say, then, that World 1 is a world that is **better than** World 2?

The basic idea, then, is that in addition to ethical or evaluative judgments concerning the goodness or badness of people, the rightness and wrongness of actions, the goodness or badness of traits of character, one also has:

> Judgments or propositions about the **goodness** or **badness** of **events and states of affairs**, about the **desirability** or **undesirability** of such things.

3. One of the Most Fundamental Questions in Meta-Ethics

We are now in a position to consider one of the most important, and one of the most fundamental questions in meta-ethics. It concerns the relation between evaluative judgments or propositions of the final sort just mentioned, and evaluative judgments or propositions of the other three types considered earlier.

To simplify things, suppose that one decided that as regards ethical judgments or propositions about actions, people, and traits of character, the basic judgments or propositions were those concerning actions, asserting that certain actions were morally right, or morally wrong, or morally permissible, or morally obligatory, that certain actions should or should not be done, or may be done. The questions that are crucial are then as follows:

(1) Are judgments or propositions about the rightness and wrongness of actions logically more basic than judgments about the good and badness, the desirability or undesirability, of events and states of affairs, or is it the other way around, with the latter being more fundamental? Or are both types of judgments fundamental?

(2) If one of these two types is more fundamental, which one is it, and how is the other type of judgment related to that more fundamental type?

How might one of these two types of judgments be related to the other? One possibility is that judgments or propositions about the rightness or wrongness of actions might be fundamental, and judgments or propositions about the goodness or badness of states of affairs might be analyzed in one of the following ways:

"A state of affairs of type S is good" =def.

"One ought to bring about states of affairs of type S."

Or perhaps:

"A state of affairs of type S is good" =def.

"One ought not to destroy states of affairs of type S."

Suppose instead that it is judgments or propositions the goodness or badness of states of affairs that are fundamental. Then judgments or propositions about the rightness or wrongness of actions might be analyzed as follows:

"Action S is morally right" =def.

"Among the possibilities open to one, action S is the one that produces the best balance of good states of affairs over bad states of affairs."

"Action S is morally wrong" =def.

"Among the possibilities open to one, action S is not the one that produces the best balance of good states of affairs over bad states of affairs."

We'll consider the pros and cons of these alternative views later on. But, offhand, do you have a clear preference for either of these views?

4. Three Main Types of Ethical Theory: Consequentialist Theories, Deontological Theories, and Virtue Theories

Different views as to which type of ethical statement is the most fundamental give rise to different sorts of ethical theories. So let us consider the three main possibilities have taken seriously.

(1) Consequentialist Theories

Consequentialist theories maintain that the fundamental ethical judgments involve claims about what states of affairs are intrinsically good and intrinsically bad. The idea then is that an action's being wrong can be analyzed along the lines just mentioned, that is, in terms of its not being the action that, among those that are open to one, produces the best balance of good effects over bad effects.

At this point, there is a crucial idea that needs to be introduced, namely, the idea of something's being **intrinsically good**, or **intrinsically bad** – or, alternatively, of something's being **good in itself**, or **bad in itself**. Consider, for example, the experiences that one typically has when one visits a dentist. They are not usually experiences that one would choose to undergo unless one thought that things would be better later on than they would be if one did not see the dentist. Visiting the dentist is therefore **instrumentally good**; it is **good as a means** either to something else that is desirable, or as a means to something else that is even

more undesirable, such as a lot of future pain. The idea then is that if something is desirable as a means, if something is instrumentally valuable, then it must be because is either brings about some state of affairs that is **good in itself**, **good independently of its consequences**, or because it is a means of avoiding about some state of affairs that is **bad in itself**, **bad independently of its consequences**. One has then the crucial ideas of states of affairs that are **intrinsically good** and **intrinsically bad**, and this gives rise to the following very important ethical questions:

(1) What types of states of affairs are **intrinsically good**?

(2) What types of states of affairs are **intrinsically bad**?

Class Discussion of these Two Questions

What are some possible answers to question (1)?

Pleasure? Friendship? Freedom to act? Knowledge, perhaps of certain types? Great works of art? Development of a morally good character? A relationship with God?

What are some possible answers to question (2)?

Pain? Ignorance of important truths? Development of a morally bad character? Lack of a relationship with God?

Given the idea of intrinsically good and intrinsically bad states of affairs, let us now return to considering consequentialist approaches to morality.

A Famous Consequentialist Theory: Utilitarianism

The most famous consequentialist theory is utilitarianism. This theory comes in different versions. Some versions maintain that the only thing that is **intrinsically good**, or **good in itself**, is **pleasure**, and the only thing that is intrinsically bad is pain. So to determine the moral status of an action, what one considers the total **quantity** of pleasure that the action produces, and the total **quantity** of pain that it produces. The better the balance of the former over the latter, the better the action is, and the action with the best balance of pleasure over pain is the action that one should perform.

If this view is right, then other things that are valuable are only **instrumentally** valuable. So things like friendship, knowledge of various sorts, great works of art, and so, are valuable only because they give rise to pleasure, or to a reduction in pain.

This version of utilitarianism was advanced by Jeremy Bentham (1748-1832). Bentham entered Oxford at the age of 12, and graduated at the age of 15, and then went on to study law. Rather than practicing law, however, he worked on the tasks of developing a better legal system, and of reforming both criminal and civil law. His work had a very great impact upon legal theory.

Bentham's approach to ethics was then both adopted, and modified, by many philosophers. Two of the earliest were James Mill (1773-1836), and his son, John Stuart Mill (1806-1873). The latter is the author of *Utilitarianism,* certainly the best-known exposition of utilitarianism, and still widely read today.

In his book *Utilitarianism*, John Stuart Mill argued that one could maintain, as Bentham had, that the only thing that is intrinsically good is pleasure, and the only thing that is intrinsically bad is pain, without agreeing with Bentham that all that mattered was the **quantity** of the pleasure or pain. Thus Mill said,

It is quite compatible with the principle of utility to recognize that some kinds of pleasure are more desirable and more valuable than others. In estimating ·the value of· anything else, we take into account quality as well as quantity; it would be absurd if the value of pleasures were supposed to depend on quantity alone.

This immediately leads, of course, to the question that Mill immediately went on to ask, and answer:

'What do you mean by "difference of quality in pleasures"? What, according to you, makes one pleasure more valuable than another, merely as a pleasure, if not its being greater in amount?' There is only one possible answer to this.

Pleasure P1 is more desirable than pleasure P2 if: all or almost all people who have had experience of both give a decided preference to P1, irrespective of any feeling that they ought to prefer it.

If those who are competently acquainted with both these pleasures place P1 so far above P2 that they prefer it even when they know that a greater amount of discontent will come with it, and wouldn't give it up in exchange for any quantity of P2 that they are capable of having, we are justified in ascribing to P1 a superiority in quality that so greatly outweighs quantity as to make quantity comparatively negligible.   (Jonathan Bennett's translation, p. 6)

Then a little later (p. 7), there is Mill's famous remark:

It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool or the pig think otherwise, that is because they know only their own side of the question. The other party to the comparison knows both sides.

Mill's answer to the question of what it means to say that one pleasure is **qualitatively better** than another one is interesting.   But one can also ask what it is for one pleasure to be **quantitatively** greater than another, and the question is whether there is any way of answering that question other than by also appealing to what people who are familiar with both pleasures would choose.

It's not exactly clear, then, that Mill's distinction between quantitative and qualitative differences in pleasures makes sense. If not, then one could view the above passages from Mill as really providing an account of what it is for one pleasure to be **quantitatively greater** than another.

Two other, fairly closely related versions of utilitarianism are worth mentioning. First of all, one can jettison the ideas of pleasure and pain in favor of the ideas of happiness and unhappiness, and hold that the only thing that is **intrinsically good** is **happiness**, and that the only thing that is **intrinsically bad** is **unhappiness**. One then needs to offer an account of what happiness and unhappiness are, if this is not to be explained simply in terms of pleasure and pain.

 A final version of utilitarianism maintain that the only thing that is **intrinsically good** is getting what one wants, **having one's desires satisfied**, and that the only thing that is **intrinsically bad** is not getting what one wants, **not having one's desires satisfied**.

But, it might be objected, aren't pleasure and pain very closely related to the satisfaction and non-satisfaction of one's desires? The answer is that this isn't so.

Many people, especially in America, for example, have a very strong desire to be extremely wealthy. They would be willing, for example, to spend, say, two years in solitary confinement if they would receive a few million dollars at the end of those two years. Yet many such people may virtually never think about the fact that they are not extremely wealthy, and, even when they do, will not be in pain. So not having a very strong desire satisfied is not identical with, nor even closely related to, being in pain.

This form of utilitarianism is often referred to as **preference satisfaction** utilitarianism, in contrast to what is referred to as the **hedonistic** utilitarianism of Bentham. The preference satisfaction utilitarian agrees with hedonistic utilitarian that pleasure is intrinsically good and pain is intrinsically good, but he or she maintains that pleasure is intrinsically good pleasure is, by definition, simply a sensation that one intrinsically desires, while pain is, by definition, simply a sensation whose absence one intrinsically desires. So pleasure is intrinsically good precisely because one is getting what one intrinsically wants, while pain is intrinsically bad precisely because one is getting what one intrinsically does not want. But pleasure and pain are not the only things that are intrinsically significant, since one can have desires for things other than sensations.

Given an account of what states are intrinsically good and what states are intrinsically bad, the utilitarian claims that the rightness of an action is related to the good or badness of its consequences. Though here, too, there is a split between two rather different approaches, referred to as

(1) **Act utilitarianism,** and

(2) **Rule utilitarianism**.

Act utilitarianism maintains that the rightness and wrongness of an action is a matter of how the consequences of the action compare with the consequences of the other actions that the person could have performed: the right action is the one with the best consequences, the best balance of good effects over bad effects.

Rule utilitarianism, in contrast, maintains that the consideration of consequences is relevant, not in choosing an individual action, but in deciding upon **moral rules**. The correct moral rules are those rules that are such that if people follow them, the best consequences will result overall, over the long term. The right action to perform in a given situation, accordingly, is **not** the action that will have the best consequences **in that situation**: the right action, rather, is the action that conforms to the moral rules that are the best moral rules - in the sense just explained.

Non-Utilitarian Forms of Consequentialism

A distinctive characteristic of utilitarianism in all of its forms is that it is held that there is **only one** thing that is intrinsically good and **only one** thing that is intrinsically bad. If one maintained, instead, that there were **two or more** things that were intrinsically good or **two or more** things that were intrinsically bad, then one have to say how to weight those different things against one another. So if one held, for example, that both pleasure and friendship are intrinsically good, the question would arise as to how valuable friendship was in comparison with pleasure, and it might well be difficult to find a **non-arbitrary** answer to that question.

Could there be a non-utilitarian form of consequentialism that there is only one thing that is intrinsically good and only one thing that is intrinsically bad? The difficulty would be that it is surely true either that non-sadistic pleasures are intrinsically good, or that the satisfaction of harmless desires is intrinsically good, and the question is how one could account for this in terms of some other type of thing that is intrinsically good.

The upshot is that it seems likely non-utilitarian forms of consequentialism will have to be what one might call **non-monistic** theories, according to which either that there are **two or more** things that are intrinsically good or that there are **two or more** things that are intrinsically bad, or both, and so will face the comparison issue mentioned earlier.

(2) Deontological Theories

The most important alternative to consequentialist theories is what are known as "deontological" theories. (The term "deontology" comes from a Greek word that means, "what is binding".) Deontological theories maintain that statements about what one ought to do cannot be explained in terms of statements about the goodness and badness of consequences.

In support of this claim, advocates of deontological approaches often argue that individuals have **rights**, and that it is wrong to violate those rights simply on the grounds that doing so will lead to a better balance of good states of affairs over bad states of affairs.

A possible argument for deontological theories, and against consequentialist theories involves the case of a brilliant doctor – call her Mary, and it runs as follows:

Mary is stranded on a desert island, and that she has five patients who need various organ transplants if they are to live: one needs a heart transplant, one a liver transplant, one a lungs transplant, one a kidney transplant, and one a pancreas transplant. Unfortunately, no organs are available, so it looks as if Mary cannot do anything to save her five patients. But then it occurs to her that she has a healthy patient, and that that person could be a source of the needed organs. So Mary kills the healthy patient, and uses that person's organs to save the lives of the five patients who would otherwise die.

The question was whether Mary acted wrongly or not? If she did, then it would seem that at least act consequentialist positions must be false, since Mary by choosing to kill one patient to save five, chose the action with the better consequences.

Many people are inclined to say that what Mary did was wrong. But as we also saw, when one shifts to some cases that appear to be related – such as the trolley case where one can save five people by steering a trolley onto another track, where it will kill only one person, people sometimes have intuitions that conflict with their intuitions in the doctor case.

To return to the characterization of deontological theories, such moral theories generally feel that some or all of the following concepts are of ethical importance:

(a) the idea of individual **rights**;

(b) the idea that someone **deserves**, or does **not deserve** something;

(c) the idea of **fairness**;

(d) the idea of a reasonably **equitable distribution**.

(a) has been illustrated by the case of the doctor, but could be illustrated by many other cases as well. Consider, for example, massive redistribution of the world's wealth.

For (b), consider the case of deciding who should get a TV set that is to be given away: a friendly grandmother, or a mass murderer. (Suppose that it would give the murderer more happiness than the grandmother.)

For (c) and (d), consider a parent who consistently gives nicer presents to one of her children than she gives to the others. (Again, we can suppose that more happiness is produced by this way of doing things.)

Some Varieties of Deontological Theories

One of the big divides in the case of deontological theories is that between those that postulate a number of unrelated principles concerning what actions are right and what actions are wrong. A philosopher named W. D. Ross (1877-1971), for example, maintained that people had a number of what he referred to as "prima facie duties". Thus, in his book *The Right and the Good* (1930), Ross listed the following seven prima facie duties:

(1) The obligation to be faithful, including the obligation to keep one's promises;
(2) The obligation of reparation – the obligation to repay others when one has harmed them;
(3) The obligation of gratitude;
(4) The obligation of non-maleficence – that is, not to harm others;
(5) The obligation to treat others fairly and justly;
(6) The obligation of beneficence – that is, to help and benefit others; and
(7) The obligation to improve oneself.

As in the case of non-monistic consequentialist theories, when one holds that there are several different factors that are morally significant, the question arises as to how they are weighted. So Ross is faced with the question of saying how these various duties compare with one another in strength.

A famous deontological theory that attempts to avoid this problem is that offered by the German philosopher (1724-1804). Kant set out his moral theory in different places, one of which you have read parts of: *The Foundations of the Metaphysics of Morals*. There Kant attempt to set out a principle from which all moral duties can be derived. His first formulation of that principle is as follows:

"**Act as though the maxim of your action were to become, through your will, a universal law of nature**." (Jonathan Bennett's translation, p. 24)

Then, later on, Kant offers two other formulations of what he takes to be the fundamental principle of morality:

"**Act in such a way as to treat humanity, whether in your own person or in that of anyone else, always as an end and never merely as a means**." (Jonathan Bennett's translation, p. 29)

"**Act only so that your will could regard itself as giving universal law through**

its maxim." (Jonathan Bennett's translation, p. 32)

One problem here is that while the first and the third formulations seem fairly closely related, the second formulation looks very different, and it is not easy to see how one can get from the first formulation to the second, and Kant's attempt to do so does not seem to me at all satisfactory.

The main problem, however, is concerned with how one can derive any specific duties from any of these three formulations. Thus John Stuart Mill, in *Utilitarianism*, commenting on Kant's approach, says,

This remarkable man, whose system of thought will long remain one of the landmarks in the history of philosophical thought, lays down in that treatise a universal first principle as the origin and ground of moral obligation:

Act in such a way that the rule on which you act could be adopted as a law by all rational beings.

But when he begins to derive any of the actual duties of morality from this principle he fails, almost grotesquely, to show that there would be any contradiction—any logical impossibility, or even any physical impossibility—in the adoption by all rational beings of the most outrageously immoral rules of conduct. All he shows is that the universal adoption of such rules would have consequences that no one would choose to bring about. (Jonathan Bennett's translation, p. 3)

Thus Kant attempts to show, by appealing to his first formulation, that **suicide** is morally wrong, and claims that one could not will that everyone should commit suicide. But surely there is, as Mill would say of this particular case, nothing that is either logically impossible or physically impossible about willing that everyone commit suicide. All that is true is that that everyone's committing suicide would be a very undesirable outcome.

(3) Virtue Theories

The third main type of ethical theory that is being discussed by philosophers today is what is called "virtue theory." This type of theory goes back to the beginning of philosophy in Athens, and you have read parts of Aristotle's *Nicomachean Ethics*, where Aristotle sets out a virtue theory approach.

The distinction between virtue theories and deontological theories does not seem as great as that between either of these theories, on the one hand, and consequentialist theories on the other. For it would seem that for any moral rule, one can postulate both a corresponding, virtuous trait of character, and a corresponding vice. Similarly, for any virtuous trait of character, and the corresponding vice, one can postulate a corresponding moral rule. So it seems that there is a one-to-one mapping from one to the other. One has for example:

Moral Rule	Virtue	Vice
One should not lie.	Being honest	Being a liar

But if this can be done, then the implications of a given virtue theory with regard to what one ought to do will **coincide** with the implications of the corresponding deontological theory. So the difference between deontological theories and virtue theories does not seem all that important. By contrast, consequentialist theories, such as utilitarianism, often have quite radical conclusions about what actions are right and what actions are wrong.

Still, one can ask about whether, if one rejects consequentialism, one should go with a virtue theory or a deontological theory, and it is not easy to see what to say here. One thing is that virtue theories are **not monistic** theories: a number of virtues are postulated – and in some cases, quite a large number – so that a deontological theory that derives everything from a single moral rule – or at least a very small number – will be intellectually more illuminating.

Aristotle, in his *Nicomachean Ethics*, did offer a general characterization of virtues, holding that every virtue was a mean between two extremes, the extremes being vices. Thus he described courage as a mean between cowardice on the one hand, and rashness (or foolishness) on the other. But I don't think that this characterization, if it were right, would really help one to pick out the virtues, since I think a trait can be intermediate between two other traits, without being a virtue. Consider, for example, the extremes of wanting to stick to the same routine, and wanting as much variety as possible. Neither of these traits seems to me to be a vice, nor does it seem that there is a virtue that lies between them. So being a mean between two extreme is **not a** *sufficient* **condition** of being a virtue. **Nor** does it appear to be **a** *necessary* **condition**. For consider the virtues of being **honest**, or being **faithful.** What, in each of these cases, are the two vices that the virtue in question is intermediate between?