4

CONSEQUENTIALISM AND KANTIAN ETHICS

This chapter will discuss two influential traditions of ethical thought. The first one of these is consequentialism. The most famous version of consequentialism is **utilitarianism**,[1] which was famously defended by Jeremy Bentham and John Stuart Mill. Bentham summarized this view by claiming that "it is the greatest happiness of the greatest number that is the measure of right and wrong (**Bentham, 1777**)."[2] This **greatest happiness principle** is not helpful because it doesn't tell whether (i) you should go for the *greatest amount of happiness* or (ii) attempt to make *as many people happy as possible*. There are situations in which you just can't do both. 4.1

It is better to begin from what is special about all **consequentialist ethical theories** (Mulgan, 2007; Bykvist, 2010; Driver, 2012).[3] Consequentialism is best understood as a family of ethical views that share the same structure. The next section will explain what this structure is and the different versions of consequentialism that share it. The second section will then discuss John Stuart Mill's famous argument in favor of utilitarianism. This argument is important because it clearly reveals the important moral ideals that continue to motivate utilitarians. 4.2

After consequentialism we will look at **Immanuel Kant**'s ethical theory, which is based on the **Categorical Imperative**.[4] This principle requires that you should "never . . . act otherwise than so that [you] could also will that [your] maxim should become a universal law (**Kant, 1785, 4:402**)."[5] We will find out the best way to understand this principle and look at Kant's arguments to the conclusion that this is the fundamental principle of morality. 4.3

This Is Ethics: An Introduction, First Edition. Jussi Suikkanen.

4.4 Philosophers have recently come to see the debates about ethical theories in a new light (see Parfit, 2011). Philosophers used to see consequentialism and **Kantian ethics** as competing theories about what is right and wrong. As a result the defenders of these theories looked for cases in which their opponents' view seems to get things wrong about what you should do. The end of this chapter will go through some of these cases.

4.5 However, nowadays philosophers don't really care about these sorts of cases. This is because both consequentialism and Kantian ethics are very flexible. When there is an alleged counterexample to these views their defenders can always reformulate their view so as to avoid the problem. This is possible because the central parts of these theories can be understood in so many different ways. As a result the defenders of these views seem to be able to agree on what is right and wrong. The last section of this chapter will explain all of this in more detail.

Consequentialism

4.6 In order to understand consequentialism you need to know first that there are two distinct families of moral qualities. First of all, there are **evaluative qualities**. You talk about them when you use words such as "value," "good," "bad," "desirable," "worthwhile," and "utility." You can, for example, say that it is good that so many people help the poor.

4.7 In addition to these evaluative qualities, there are also **deontic qualities**. You talk about them when you use words such as "ought," "must," "should," "right," "permissible," "wrong," "reasons," "duty," and "obligation." You can use these words to describe how people are to act. If you say that Ben ought to keep his promise, you are telling him what to do.

4.8 The fundamental idea of **consequentialism** is that you can only make sense of the deontic qualities in terms of the evaluative ones. Therefore, according to the core idea of consequentialism, what is good is always prior to and more important than what you ought to do and what is right or wrong. This is to say that consequentialists think that truths about good and bad explain all the truths about right and wrong.

4.9 Consequentialist theories therefore always have two separate elements. Let us call the first one of these **the evaluative element**, which is a **theory of value**.[6] Such a theory attempts to capture which things are good. Theories of this type play a crucial role in all forms of consequentialism.

Imagine that you are in a situation in which you have the following options: 4.10

A. Call a friend.
B. Watch television.
C. Travel to France.

Suppose that, for some reason, these are your only options. The value theory element of a consequentialist view allows you to rank these options in terms of how good their consequences are.

We can use a very silly theory of value to illustrate how this works. This silly (and blatantly false) theory of value says that nothing except how many rubber ducks there are in the world has value. Rubber ducks and only rubber ducks are good. Let us also imagine that whether you do A, B, or C affects the number of rubber ducks there will be. Traveling to France makes some people want to produce more rubber ducks whereas calling a friend makes them want to destroy few. As a result this theory of value means that traveling to France is your best option and calling a friend your worst one. 4.11

In order to be able to always rank your options like this in terms of how good their consequences are, all consequentialist theories must thus give an account of what is good. They have to describe what makes outcomes better – like the rubber ducks did in the previous example. Instead of silly rubber duck views, real consequentialists obviously want to offer you more plausible and realistic theories of value which can tell us what kind of outcomes are genuinely good. Note that when you compare how good your options are it does not matter how far in the future the relevant goods are created or who gets them. All consequences matter equally just as long as they are better consequences than the consequences of your other alternatives. 4.12

The second element of all forms of consequentialism is the so-called **deontic**[7] element. It uses the evaluative ranking of the options to specify which options it is right and wrong to choose. One deontic element could say that it is always wrong to do the act that is ranked the lowest. This deontic element would mean in the previous case that it is wrong for you to call a friend. All forms of consequentialism therefore tell us what is right and wrong by relying on the evaluative rankings of the options. This explains why, according to these views, what is good is always prior to what 4.13

is right (Rawls, 1971, p. 24). As we'll see, consequentialists also disagree about which sort of deontic element to use in their theory.

Utilitarianism

4.14 **Utilitarianism**[8] illustrates nicely the basic structure of consequentialism. According to this theory, happiness, and only happiness, has value. Utilitarians therefore rank your options by how much happiness they bring about. Again, it does not matter who experiences this happiness and when. Utilitarians then add that an action is right if and only if it has the best consequences – that is, when the action brings about more general happiness than the other actions which you could do in your situation.

4.15 In many cases this view has intuitive consequences. Why is it, for example, wrong to rape someone? Utilitarians have a plausible answer to this question. It is true that the rapist will get some fleeting happiness from the act of rape. However, the rape will also make its victim deeply unhappy. It will also affect the happiness of the victim's friends and family, and more generally will create a negative atmosphere. On balance then, rape has horrible consequences, which makes it the wrong thing to do according to utilitarians.

4.16 Utilitarianism also has more controversial consequences. Consider **William Godwin**'s famous example: you can save either a skillful surgeon or your mother (**Godwin, 1793, vol. I, book II, ch. 2**).[9] In this case, you will produce more general happiness if you save the surgeon. If you don't save your mother, you and your family will be unhappy and your mother will no longer experience any happiness. However, if you don't save the surgeon then a whole lot more people will be just as unhappy and unable to experience happiness – all the other people whom the surgeon's skills could help. For this reason the utilitarians rank the option of saving the surgeon above saving your mother. This means that according to utilitarianism, it is right for you to save the surgeon (the option with the best consequence) and wrong for you to save your mother.

4.17 Utilitarianism is not the only version of consequentialism. All forms of consequentialism have an evaluative element and a deontic element. When consequentialists formulate their views, they can pick and mix different combinations of evaluative and deontic elements from the following alternatives. The following sections will introduce some of the most important versions of consequentialism. These versions will be motivated by the problems of the standard form of utilitarianism.

Deliberation procedure vs. criterion of rightness

Some consequentialists put their theory forward as a **deliberation procedure** and others as a **criterion of rightness** (**Bales, 1971**).[10] A deliberation procedure, funnily enough, is what you use to decide what to do. If a consequentialist says her theory is the correct deliberation procedure, then this means that you have to think about, and rank, the consequences of your available options when you're deciding what to do. 4.18

The trouble is that understanding consequentialism like this undermines the whole point of the theory. Deliberating in this consequentialist way requires too much time – you will usually have lots of options open to you, and remember you have to think about the long-term consequences of what you do. If you went about things in this way, you'd never get anything done. In this way consequentialism is self-undermining: if you use it to decide what to do, the consequences are bad. 4.19

The alternative is to think that consequentialism is a criterion of what is right. It only says that it is always right to do what has the best consequences. This leaves it open how you should go about deciding what to do. It could be that you can get better consequences by following simple, tried-and-tested moral rules of thumb than by working out the details of every decision you have to make. 4.20

Direct vs. indirect forms of consequentialism

Act-consequentialism is the best example of **a direct consequentialist theory**. Consider the previous case in which you can either lie or tell the truth. Act-consequentialism says that what is right for you to do in this situation depends only on how good the consequences of these actions will be. If lying has good consequences, then is right for you to lie. The rightness and wrongness of your acts is thus *directly* a result of how good the consequences they have are. Many people think that this view makes bad actions such as lying to other people too easily permissible. 4.21

Indirect forms of consequentialism that attempt to avoid this problem are slightly more complicated (Adams, 1976; Hooker, 2000). On these views we first evaluate the consequences of something other than acts. Whether an act is right or wrong then depends on those other things. **Rule-consequentialism**[11] is a good example of indirect consequentialism. According to this theory, you need to first consider the consequences different rules would have if they were adopted by everyone. Whether an 4.22

act is right then depends on whether it is permitted by the rules that have the best consequences.

4.23 So, in the previous case, a rule that allows you to lie would probably have catastrophic consequences. If everyone lied, you couldn't trust anyone! This is why the best rules forbid you to lie. This means that rule-consequentialism tells you to tell the truth even if this would have bad consequences in the given situation. Some people think that for this reason rule-consequentialism better matches what we intuitively think is right and wrong.

Utilitarianism vs. richer conceptions of value

4.24 As explained above, utilitarians think that only happiness is good and they understand happiness in terms of the balance of our pleasurable and painful experiences (**Mill, 1861, ch. 2, para. 2**).[12] So, according to their view, right actions always maximize the amount of pleasure and minimize the amount of pain. In many cases, this theory of value seems to have embarrassing consequences. Imagine living in a society in which the vast majority experiences pleasure from the misfortunes of a small minority. According to basic utilitarianism, in this situation it would be right for the majority to discriminate against the minority. This would, after all, make the society a happier place overall.

4.25 This hedonistic theory of value is not the only option for consequentialists. Firstly, if you like the idea of utilitarianism but you don't like hedonism, you could think that only well-being is good and then understand well-being in terms of desire satisfaction. Your view would then say that it is always right to maximize the general amount of desire satisfaction. You could also accept a **perfectionist**[13] form of consequentialism according to which it is always right to maximize the realization of human capabilities (Hurka, 1993).

4.26 Even more value pluralist forms of consequentialism could claim that:

1. Some basic acts like breaking of promises, killing, and lying are bad in themselves.
2. Some things like human interaction, knowledge, virtue, and aesthetic pleasures are good in themselves.

These views would then go on to argue that it is always right for you to act in ways that minimize the number of bad acts done, and maximize the

amount of intrinsically good things. This consequentialist theory would not care about how happy people are.

All these different theories of value will create very different evaluative rankings of options. As a result, they will lead to very different conclusions about what is right and wrong. This is important because many objections to consequentialism assume very simplistic theories of value, which are only optional for the consequentialist. The more sophisticated theories of value can help consequentialists to avoid these objections. For example, in the case of the previous case where the majority enjoys discrimination, a form of consequentialism that says that discrimination is bad in itself will tell you not to discriminate, which is exactly what you want to hear. 4.27

Actual vs. expected value

Some consequentialist views rank options in terms of **the value of their actual consequences** and the rest in terms of the **expected value** of their consequences.[14] Consider a case in which you offer a friend of yours a glass of **water**.[15] Sadly, your friend dies as a result of this because someone had, unbeknownst to you, poisoned the water. Did you do anything wrong? 4.28

The actual consequences of your friendly act were horribly bad. The consequences of not offering your friend a drink would not have been equally bad. Therefore, if the actual consequences of your acts matter, then you did something wrong in this case. Most people think that this is an absurd view – you didn't know that the water was poisoned. To avoid this problem, many consequentialists point out that your act had very good expected consequences. You didn't expect your friend to die. They then say that according to more plausible versions of consequentialism, it is always right to do what has the best expected consequences. 4.29

The expected value of an act is determined by: 4.30

i. what potential outcomes the act has;
ii. how likely these different outcomes are; and
iii. how much value they contain.

The better the potential outcomes of a given act are and the more likely these good outcomes are, the more expected value the act has (for how the expected value of an option is calculated more precisely, see paragraphs 10.82–10.85 below). The likely outcome of offering someone water is that they will enjoy drinking it. It is highly unlikely that they will die. Because

of this, even if the actual outcome of your act – that your friend died – was bad, your act of offering her water still had a lot of expected value. The expected value form of consequentialism therefore fits the intuition that you don't do anything wrong if you harm people by accident.

Maximizing vs. satisficing

4.31 Consequentialists can also accept different deontic elements. So far we have only discussed **maximizing** versions of consequentialism. They claim that it is always right to do whatever has the best consequences, and that it is wrong to do anything else. Maximizing forms of consequentialism thus claim that you should always maximize the good.

4.32 In contrast, **satisficing** versions of consequentialism claim that you should always do what has *good enough* consequences (**Slote, 1984**).[16] Imagine a situation in which your options are ranked in terms of how good their consequences are. In this case a satisficing theory does not require you to pick the best option but rather only an option that has at least a certain level of good consequences. If going to Amsterdam has almost as good consequences as going to Paris (which would be best), then this view allows you to choose either option. Some people think that satisficing versions of consequentialism are more plausible because they leave us more freedom to choose between many morally permissible options.

Mill's Argument for Utilitarianism

4.33 Now you know what the core idea of consequentialism is and you are familiar with the main versions of consequentialism. Why should you be a consequentialist in the first place? This section explains **John Stuart Mill**'s famous answer to this question.[17] Many people think that his argument in favor of utilitarianism is a complete disaster. However, there is more to Mill's argument than first meets the eye. Mill's argument is important because it reveals the fundamental moral ideals on which consequentialist thinking is based.

4.34 Mill's argument can be quoted here almost in full:

> The only proof capable of being given that an object is visible, is that people see it… In the like manner… the sole evidence it is possible to produce that anything is desirable, is that people do actually desire it… No reason

> can be given why the general happiness is desirable, except that each person, so far as he believes it to be attainable desires his own happiness. This, however, being a fact, we have not only all the proof which the case admits of, but all which it is possible to require, that happiness is a good; that each person's happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of all persons. (**Mill, 1861, ch. 4, para. 3**)[18]

As you saw above, all consequentialist theories have two elements. Mill only addresses the first one in his argument. He assumes without an argument that it is always right to do what has the best consequences. His argument's main aim is to show that only happiness can make the consequences of an act good: that "happiness is desirable, and the only thing desirable, as an end" (Mill, 1861, ch. 4, para. 2). He also wants to show that everyone's happiness is equally desirable. 4.35

The first stage of Mill's argument to this conclusion compares the visibility of objects to the desirability of outcomes. Mill appears to think that, because an object is visible when it is seen, an outcome is similarly desirable when it is desired. Given that you probably desire your own happiness, this enables Mill to conclude that your own happiness must be desirable for you. 4.36

In the second stage of the argument Mill directly states, on the basis of the previous claim, that everyone's general happiness must also be desirable as an end. As you can see, Mill suggests in the quote above that *because* your happiness is good for you therefore general happiness must be a desirable end for all of us. 4.37

The final part of the argument attempts to argue that only general happiness is a desirable end (Mill, 1861, ch. 4, paras 4–7). This part of the argument is a response to an objection and it takes place just after the quoted passage. You could object that people in fact desire many other things than happiness. You can desire food and conversation, for example. Doesn't it follow from Mill's own argument that these things too must therefore be desirable? 4.38

Mill's response to this objection is that you always first desire food and conversation as a means to happiness. However, once you eat food and take part in conversation, you will come to realize how pleasant these experiences are in themselves. As a consequence, these goods become an element of your happiness. In desiring food and conversation, you are then in the end desiring the pleasant experiences that constitute your happiness. 4.39

Because of this, you do not really desire anything else except happiness, and thus only happiness is desirable.

The problems with Mill's argument

4.40 We can now list Mill's mistakes. Firstly, Mill merely assumes without any argument half of what he is supposed to prove. The assumption that it is always right to do what has the best consequences is not supported by any part of the argument whatsoever. However, unless you show all parts of a theory to be true, you do not successfully prove that the theory is correct.

4.41 The second step of Mill's argument relies on a problematic analogy between visibility and desirability (**Moore, 1903, ch. 3, sects 39–40**).[19] It is true that the things you see are visible. This is because to be visible is to be something that is capable of being seen. However, it is not true that to be desirable is to be something that people can desire. You can desire to count blades of grass but this does not make doing so desirable. Instead, to be desirable is to be something that is *worth* desiring. As the previous example illustrates, there are many things we desire that are not worth desiring and thus are not desirable. The second step of Mill's argument is thus equally flawed.

4.42 The third step of his argument is even worse (**Sidgwick, 1907, p. 388**).[20] Mill just bluntly asserts that because your happiness is good for you, the general happiness of everyone must therefore be a desirable goal for each of us. However, this is just what Mill's opponents want to deny. Egoists who do not care about others, for example, claim that only their own happiness is a desirable goal for them. Mill gives no reason for the egoists to change their mind. If you read the quote above carefully from the egoist's perspective, nothing it in addresses your views.

4.43 The last step of Mill's argument is not any better either (**Moore, 1903, ch. 3, sect. 43**).[21] Perhaps it is true that we can only desire experiences that are pleasant and that these experiences are also desirable. Even in this case, you could still desire these experiences for other reasons than the pleasures they provide you. Consider **Picasso's Guernica**.[22] Even if you get pleasure from looking at this magnificent painting, many of us want to spend time looking at it for other reasons too. It is desirable to look at this painting because it illustrates the horrors of the Spanish Civil War in an acute way.

Saving Mill's argument

At this point, you might think that Mill's argument is hopelessly flawed. 4.44
Despite this, Geoffrey Sayre-McCord convincingly argues that Mill's argument still contains some of the most important ethical ideas that support consequentialism (**Sayre-McCord, 2001**).[23]

It is true that Mill doesn't argue for the other half of utilitarianism. This is 4.45
not because he was a bad philosopher. Mill did not argue for the claim that right actions have the best consequences because this was believed to be self-evidently true when he was writing (**Moore, 1903, ch. 1, sect. 17**).[24]
Imagine that you have two options: one that is very good and another that is not good at all. Isn't it obvious in this case that you should choose the better option? What possible reason could there be for making things worse?

The second step of Mill's argument is not too bad either. If you read the 4.46
long quote above carefully, you notice that Mill never claims that to be desirable is to be what people desire. All he says is that what people desire is *good evidence* for what is desirable. This has to be correct. Consider all the things that you desire. Most of those things have to be good, right? One plausible explanation for this is that we desire things just because we take them to be good in some way.

The third step of the argument is the most important one. This is the 4.47
step where Mill moves from the idea that your happiness is good for you to the conclusion that everyone's happiness is a desirable goal for all of us.

If you want to be charitable to Mill, you should read this quick step in 4.48
the following way. Your happiness is desirable for you – it is a worthwhile goal for you. We can then compare your happiness to my happiness. If you do this, you must recognize that they seem to be exactly alike. In fact, the only thing that distinguishes your happiness from mine is that your happiness belongs to you and mine to me. But, if your happiness is otherwise just like mine and your happiness is something worth pursuing, then shouldn't my happiness too be worth pursuing even if it is not yours?

This reasoning captures the important moral insight behind Mill's argu- 4.49
ment. **When you consider two things and you can't find any relevant differences between them, you must consider them to be equally good**.[25]
If there are no significant differences between your and my happiness, and your happiness is desirable, then mine must be too. In fact, given that there are no important differences between your happiness and other people's happiness, then everyone's happiness must be equally desirable – just as Mill argued.

4.50 We can therefore understand the main crux of Mill's argument as a challenge. If you don't want to be a utilitarian, you must either deny that your own happiness is desirable or you must explain what makes your happiness different from other people's happiness. This difference cannot be that your happiness belongs to you whereas mine belongs to me given that your car is not better than mine merely because it is yours. Unless you can respond to this challenge, Mill's reasoning requires you to accept utilitarianism.

4.51 This important insight also reveals the fundamental ethical ideals behind consequentialism. The argument Mill is making here is that my happiness can't be any more important than anyone else's happiness. Everyone's happiness must therefore count equally. In this way, Mill is grounding his utilitarianism on an appealing idea of the basic moral equality of everyone. This idea that we should be impartial is in many ways appealing. It supports the idea that you should not treat people differently on the basis of their gender, ethnicity, wealth and social class, sexual orientation, and so on. When Mill was writing, this idea was very radical, and even today many people continue to resist the idea that we all matter equally, morally speaking.

Kantian Ethics

4.52 Even if Mill's argument makes utilitarianism appealing, there are also good alternatives. You already know what these alternatives must be like. Non-consequentialist ethical theories have to specify what is right and wrong without relying on what is good. The opponents of consequentialism thus need an alternative starting point. This section introduces one of the main alternatives to consequentialism: Kantian ethics. The beginning of the section explains how Kant thought that you could use the so-called **universalization test** to determine what principles it is right and wrong to act on. The second part of the section will then investigate the main arguments for Kant's view. These are at the same time arguments for why we should follow the principles that pass Kant's test.

The good will

4.53 **Immanuel Kant** thought that we could start from our own wills (**Kant, 1785**; for **an overview of Kant's ethics**, see **Johnson, 2008**; see also O'Neill, 1975; Hill Jr., 1992; Herman, 1993; Korsgaard, 1996; Timmermann,