

Remark 1. In our Kant handouts pagination follows the numbers on the margins, which are the pages of the standard Prussian Academy edition of Kant's works.

THE GOOD WILL. The good will is declared good in itself, and not because of what it achieves. But first, what is the will? This is a trick question. Kant says a bit later that somehow we already have the concept of good will, and he doesn't get to talking about it explicitly until much later in Chapter III. We shall get there in due course. For now, let us simply say:

(1-1) Will = a kind of internal cause propelling a person to action towards a certain purpose.

This is not very clear, and I agree. But further, let us also say: the will is transparent to the agent. That is, the agent cannot will something unless he is also able to formulate what it is that he wills. Then we could say that the will is manifested in actions and can be individuated by agent's reports.

Example 2. Suppose you observe me raising my hand. You ask me why I did that. Possible answers:

- (A) I wanted to raise my hand.
- (B) I wanted to slap you on the face.
- (C) I wanted to catch a bullet about to hit that dog.
- (D) It went up by itself, I am as puzzled as you are.

The will, if there is any, can be individuated by these statements. In the first case, we say, the will is to raise my hand, and the purpose is achieved. In the second and the third the purposes, described by the sentences, are not reached. In the last case, however, the will associated with my motion is absent. I raised my hand unwillingly, which is in fact the same as to say that *I* did not raise my hand, that I did not *act* at all.

Kant claims then that the good will is not judged through its achievements. This means that we consider a situation where the will is to achieve a certain purpose, but there is a failure. Hence the case (C). The action was ineffective: the dog is dead. Now he claims that this fact is irrelevant, as far as the moral value of my will is concerned. The claim is better put in terms of an action, not will. Thus:

(1-2) The moral value of an action is determined by the agent's will, rather than by its actual effects.

That is, the moral value of the action is independent of whether the action achieved the purpose encoded in the will or not.

REASON AND WILL. Kant admits that the principle (1-2) may appear surprising. He proposes to defend it by using the following statement:

(1-3) Reason governs the will.

This offers a vital clue regarding the nature of the elusive notion of will. So a few comments. We often speak of actions we perform 'unwillingly'. What we mean often is that we go against our best judgement, being controlled by instinct or desire. The will is that motive that is aligned with reason (determined by it), though the action may ultimately be performed under a different kind of motive.

Example 3. I am on a diet aiming to lose weight. My dieting is a rational enterprise. However, one day I am unable to resist a cheesecake. So if I am still committed to my diet, I eat the cheesecake 'unwillingly'. I was tempted by the cheesecake and yielded to my desire.

Some philosophers (notably Hobbes) take a different view. The will is *any* motive determining the action. As long as it is possible to talk of actions at all, your actions are determined by your will.

REASON IS NO GUIDE TO HAPPINESS. Having accepted the principle (1-3), Kant now mounts a quaint quasi-evolutionary argument. Assume that the basic goal of humans is preservation (=welfare=happiness). Should reason be charged with ensuring our survival? No: first, a more reliable guide should be instinct. Second, reason often delivers the opposite of happiness. Intelligent people, precisely because of their refined intelligence, tend to be less happy than simpletons.

This argument is problematic. First, there is a sleight of hand. Preservation (*Erhaltung*) is the same as survival. But how on earth is survival the same as well-being? And how is well-being the same as happiness, understood clearly in a psychological sense?

Second, it may be that reason *is* adapted for preservation and even happiness, but not perfectly so. Perhaps it should work in tandem with other capacities. And perhaps we are still improving, evolving.

Third, it is not clear why an intelligent man should grant ignorant men more happiness. This relates to the first problem, the conflation of welfare and happiness. Perhaps ignorant men are psychologically happier, but

394

397

394

395-6

395

395-6

their well-being is still lower. In any event, intelligent men of course survive better (in happiness or not) than ignorant ones. This again shows that aligning ‘preservation’ with happiness is simply baseless.

Remark 4. Kant’s assumption regarding survival is probably directly borrowed from the Stoics. See the opening paragraphs of Cicero’s *On Duties*. However, the Stoics do not identify preservation with well-being, let alone with ‘happiness’.

THE PROPER ROLE OF PRACTICAL REASON. Kant now claims that since reason *is* given us practically, as influence on will, its only role can be the production of good will. Again, it cannot be relied on to produce good effects (with the aid of good will). Instead, it is charged with producing a will that is good in itself. This argument is of course as suspect as the previous one, for similar reasons. Nevertheless we learn that reason, when properly employed, is aligned with good will.

But now, what *is* good will. Kant claims that the concept of duty contains in it the concept of good will. We are now off to explicating this concept.

FOUR KINDS OF ACTIONS. Kant’s task, as he sees it, is to explicate the way duty motivates agents. To this end he contrasts several kinds of actions. (a) There are evil actions against duty. Kant puts them aside: it is clear that duty did not enter there into the motivation. (b) There are actions that conform to duty, but driven by other motives altogether. This is the case of a shopkeeper: honest dealing is in accordance with duty, but is not, in his case, motivated by duty. Thus it seems that duty has a dual role to play. It can be a standard of behaviour: if everyone is actually charged fairly, then this behaviour conforms to duty. And it can also be a motive.

Question 5. Give an example showing that these roles are independent of each other.

(c) Then there are actions performed out of an *immediate inclination*. Kant gives three examples. Suicide: To preserve my life is a right behaviour (in accordance with duty), and suicide is wrong. Kant will give a reason for that later on. Supposing this is the case, normally I carry on with my life out of inclination. That I have not committed suicide today is a good thing, but nothing very extraordinary. Similarly, beneficence: I can feed a stray dog out of pity or out of love. This action is commendable (it conforms to the duty of protecting animals, for example), but has no moral worth. The moral worth resides in the motivation, and its motivation was wholly psychological, void of the influence of duty. Then there is the duty to secure one’s welfare: of course, most people are in this business anyway.

(d) Finally, we can turn to actions performed out of duty. Modify the latter three examples: I may be injured or in pain, or depressed, but still carry on with my life. This time I act not out of duty-free inclination (inclinations move me to suicide), but solely out of duty. Or I can feed a dog, whilst experiencing no particular emotions towards it, being motivated entirely by my duty. Or I can suffer from a disease, so that I should have no expectation of achieving happiness. Yet now I can strive to achieve happiness (whose element is health) entirely out of duty.

KANT’S PROPOSITIONS. Kant declares his ‘second’ proposition on the nature of duty to be this: an action out of duty has its moral worth in its maxim, and not in its purpose. But what is the *first* proposition? There is a big mystery about this. But if we look at the examples just given, we could formulate perhaps thus:

First proposition An action conforming to duty has moral worth iff its maxim produces it by necessity, with or without an inclination.

The second proposition is derived from the examples too.

Example 6 (Second proposition). Consider feeding a dog. I can feed it out of love or out of duty. The purpose in both cases is the same. But, as we saw, moral worth is not the same. Why? Because the maxims in two cases differ.

The third proposition states that action out of duty is an action out of *respect* for law. Kant claims that it follows from the other two, but there is no question that we have no formal derivation here.

RESPECT. The notion of respect (or ‘reverence’—see later the discussion of Kant’s *Doctrine of Virtue*) pops up out of nowhere. Its role is clear: to give us a duty-bound, purpose-neutral motive distinct from non-moral inclinations. In a lengthy footnote Kant acknowledges the obscurity of this notion. Let us pick up some of its characteristics.

(a) Respect is a feeling—better, a kind of feeling, or an analogue of feeling. It is not caused by the world outside! Unlike a feeling of fear, say, it is generated entirely from within. (b) Specifically, it is generated by *law*. How so? Because I recognise myself as being subordinate to law. (c) Respect is a constraint on self-love. For through it, I see myself as bound by law. (d) The law I have respect for is one within me. It is not respect for the law outside, for then this respect would have been generated outside of me. (e) It is possible to respect individuals, but really, this respect is again only for law. How so? We respect individuals not for their unique characteristics, but for their instantiation of general characteristics, such as following a moral law or developing a talent. For example, I may respect (revere) Michael Kohlhaas—not because of his peculiar character, but rather because he instantiated (was an instance of) the sense of justice, integrity, courage etc.

THE LAW OF THE WILL. Kant now gets to the formulation of the law (in italics) that governs the good will. This is a preview of the categorical imperative developed later on. It implements the earlier ideas that duty, moral motivation must be self-governed: ‘I ought to act. . . I could also will. . . my maxim’.

It is vital to avoid some misunderstandings. Consider Kant’s own example of promising (simplified; to be taken up again later). Why ought I to keep promises? One answer: out of prudence (wholly selfish motives). I fear that if I break my promise, I will be punished by other people. Clearly this has nothing to do with duty, given the discussion earlier.

Another answer: out of concern for general welfare. If I break my promise, you break yours, and so forth, where will we be? Chaos and destruction. This is the interpretation sometimes favoured by utilitarians (Mill included). But this is not at all what Kant is saying (and should be saying). Duty, we said, cannot be concerned with purposes and uncertain outcomes.

But then, what does this locution ‘should become a universal law’ mean? Kant explains as follows. If my act becomes a universal policy, then the problem is not with the concrete outcome. The problem is, there will be no promises at all. So Kant sees a *contradiction* here, not pain or calamity: if universalised, my act of promising yields abolition of promises. But since, when I break my promise, I also will to break my promise, I cannot at the same will its universalisation—for there would be no promises to break then. My maxim fails to universalise, and therefore, it fails as a moral maxim.