

### Socrates on Weakness of Will

One of Socrates' most puzzling views is the claim that *akrasia*, or weakness of will, is impossible. We seem to encounter weakness of will on a daily basis. What is weakness of will? Standardly, cases of weakness of will are characterized as those in which one knows or believes that a certain action is better than another, but voluntarily does the worse action instead.

There seem to be many ways this can occur. For example, in spite of my judgment that it is best not to have a cigarette with my morning coffee, I may nevertheless go ahead and indulge, while continuing to think that the best thing for me to do would be to forego it. Or, in spite of my judgment that it would be far better for me to give my lottery winnings to charity rather than use the sum to buy a second home in the Caribbean, I may nevertheless use it to buy the second home, and give none of it to charity.

Yet Socrates seems to maintain that such incontinent actions are not possible. He claims, e.g., in the dialogue *Protagoras* (See also *Apology* 25d8-26a7):

No one who either knows or believes that something else, which is in his power to do, is better than what he is doing, subsequently does the other, when he can do what is better. (358b7-c1)

and:

No one freely goes for bad things or things he believes to be bad; it's not...in human nature to be prepared to go for what you think to be bad in preference to what is good. And when he is forced to choose one of two evils, nobody will choose the greater when he can have the lesser. (358c6-d4)

#### I. Desire and action in the Meno

Do you think, Meno, that anyone, knowing that bad things are bad, nevertheless desires them? --I certainly do.

What do you mean by desiring? Is it to secure for oneself? --What else?

Does he think that the bad things benefit him who possesses them, or does he know they harm him? There are some who believe that the bad things benefit them, others who know that the bad things harm them.

And do you think that those who believe that bad things benefit them know that they are bad?

No, that I cannot altogether believe.

It is clear then that those who do not know things to be bad do not desire what is bad, but they desire those things that they believe to be good but that are in fact bad. It follows that those who have no knowledge of these things and believe them to be good clearly desire good things. Is that not so? --It is likely.

--Plato, *Meno* 77b2-e4

Although Socrates' view that weakness of will is impossible conflicts with our ordinary experience of life, his view seems to rest on very plausible assumptions. To see his point, it is important to consider more closely the relationship between desire and action. In the *Meno* (77c-78b), (and elsewhere), Socrates argues for the claim that:

*No one desires bad things; all people desire good things.*

On the face of it, this appears to commit Socrates to the claim that:

- 1) If an action is bad/evil, then one will not desire to do it; and if an action is good, then one will desire to do it.

We can then add a further premise to connect this to action:

- 2) If I desire to perform or avoid an action, and it is in my power to do so, then I will do so.

From this it appears that we can conclude:

- 3) If an action is good, and it is in my power to perform it, then I will perform it; and if an action is bad, and it is in my power to avoid it, then I will avoid it.

But if these conclusions are correct, then if an action is bad, and I perform it anyway, it must not have been in my power to avoid it, i.e., I must have done it involuntarily. So bad action is never voluntary.

This argument lacks plausibility as it stands because the premise from the *Meno* Socrates relies on (i.e., (1) above) isn't obviously true. Moreover, (3) seems to be a much stronger conclusion than the thesis that weakness of will is impossible: the argument leads to the conclusion that *no wrong action* is voluntary, but this goes beyond saying that *action I believe to be wrong* is involuntary. Perhaps we should look for a different version of the argument that points to the latter rather than the former conclusion.

## II. The *Meno* argument: second thoughts

To start, let's consider the argument (1-3) as it stands. Should we accept the conclusion that no one desires bad things, and everyone desires good things? It would seem not. For example, suppose that I desire to eat a mushroom quiche I made from some mushrooms I found on a walk in the woods. Unbeknownst to me, the mushrooms are poisonous. Eating the quiche would be a bad thing, yet it is something I desire. Socrates considers this objection in the *Meno*. He points out the distinction to Meno between:

- a) Desiring bad things, knowing or believing they are bad.
- b) Desiring bad things, not knowing or believing they are bad.

He argues that if one desires something bad, not knowing it is, then the object of one's desire is something *believed good*, even if *actually bad*. In short, he argues that:

- If one believes an action to be good, then one will desire to do it; and  
If one believes an action to be bad, then one will not desire to do it.

The point here is to link our evaluations with our desires. Consider our practices of evaluation. Often we make initial evaluations of courses of action, and find that several actions might be good, or that one action appears better than another. As we proceed in our evaluations, weighing all the pros and cons, we finally come to a judgment about what's the best thing for us to do in the circumstances. Socrates is proposing that this evaluation leads us to action: we want to do what we judge to be overall best in the circumstances. As he puts it in the quote mentioned: "it is not in human nature to go for what you think to be bad in preference to what is good." But once we've modified his claim to read:

*No one desires what **one judges overall** to be bad; one desires what **one judges overall** to be good.*

The claim seems much more plausible. And if we revise the other premises accordingly, then we are faced with the paradox of denying weakness of will. E.g.,

- 1\*) If one believes that an action is bad/evil, then one will not desire to do it; and if one believes the action is good, then one will desire to do it.
- 2\*) If one desires to perform or avoid an action, and it is in one's power to do so, then one will do so.

From this it appears that we can conclude:

- 3\*) If one believes that an action is good, and it is in one's power to perform it, then one will perform it; and if one believes that the action is bad, and it is in one's power to avoid it, then one will avoid it.

This argument is more plausible than our initial (1)-(3) and doesn't entail that wrongdoing is always involuntary. But this argument still yields a conclusion which denies weakness of will, which is also problematic. What are we to do? All of the premises seem plausible, but they are inconsistent with the common sense claim that:

Sometimes I judge an action to be overall best, and nevertheless fail to perform it.

### III. Possible responses:

- As a matter of fact, weakness of will is impossible. Ordinary cases are not what they may seem, i.e., (3\*) is unproblematic.
- The premise that we always desire what we judge best is false; judgment and desire is not linked in this way, i.e., (1\*) is false.
- The premise that we always act on our desires (where possible) is false, i.e., (2\*) is false.

#### Divide and Conquer?

So far we have not distinguished different ways that something can be bad. On one hand things which are harmful, or against our self-interest are bad for us; on the other hand things may be judged morally good or bad. If we are sensitive to this distinction we may see that the argument can be divided into two different conclusions:

If I judge that an action is best for me, and it is in my power to perform it, then I will perform it.

If I judge that an action is morally best, and it is in my power to perform it, then I will perform it.

Consider how this distinction might affect Socrates' crucial premise:

*No one desires what they judge to be (morally) bad.*

*No one desires what they judge to be harmful (bad for them).*

The first is clearly less plausible than the second, since people may make judgments not only on the basis of moral right and wrong, but factoring in other things they value. I.e., our desires may be responsive to other concerns besides moral value. But this in fact suggests that there are two paradoxical ideas at stake:

#### Prudential paradox:

No one desires what's bad or harmful (for oneself).

All those who pursue what's bad or harmful (for themselves) do so involuntarily.

#### Moral Paradox:

Virtue is (moral) knowledge, i.e., one is virtuous if and only if one has moral knowledge).

All those who act immorally and do what is wrong do so involuntarily, i.e., out of ignorance.

### IV. Where are we now?

i) We are still left with the prudential paradox if we grant the claim that people don't desire what they judge to be bad for them and they act on their desires.

ii) Are there reasons to take seriously the moral paradox? What if we assume that doing the right thing is in our self-interest?

MIT OpenCourseWare  
<https://ocw.mit.edu>

24.01 Classics of Western Philosophy  
Spring 2016

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.